11-2022

Hydrologic and Environmental Thresholds in Stream Fish Assemblage Structure Across Flow Regimes

J. Tyler Fox

Daniel D. Magoulick
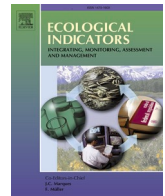
Original Articles

Hydrologic and environmental thresholds in stream fish assemblage structure across flow regimes

J. Tyler Fox [a,*], Daniel D. Magoulick [a,b]

[a] *Arkansas Cooperative and Wildlife Research Unit, Department of Biological Sciences, University of Arkansas, Fayetteville, AR, USA*
[b] *U.S. Geological Survey, Arkansas Cooperative Fish and Wildlife Research Unit, Department of Biological Sciences, University of Arkansas, Fayetteville, AR, USA*

ARTICLE INFO

ABSTRACT

The characteristic pattern of variation in flow magnitude, frequency, duration, timing, and rate of change defines the flow regime of rivers and streams and is a key driver of ecosystem processes in fluvial ecosystems. Understanding how freshwater biotic assemblages change across gradients of hydrology and anthropogenic-source disturbance in different streamflow regimes is crucial to managing for sustainable environmental flows and watershed conservation. We compiled long-term (1916–2016) occurrence records for fishes collected in the Ouachita-Ozark Interior Highlands and West Gulf Coastal Plain streams, together with hydrologic metrics calculated from daily streamflow data measured at USGS stream gauging stations (n = 111), to examine important drivers and thresholds for fish assemblage turnover in groundwater (GW), runoff (RO), and intermittent (INT) flow regimes. We also examined the importance of spatial gradients (latitude, longitude, elevation, drainage area) and anthropogenic-source stressors (Hydrologic Disturbance Index; HDI) for fish assemblage turnover using a gradient forest modeling approach. Watershed fragmentation was of high importance for fish assemblage turnover in RO and INT streams, while changes in dam storage were more important for fishes in GW streams. Hydrologic metrics describing seasonal and stochastic properties of daily streamflow (Mag6) were most important for fish assemblage turnover in INT streams. Timing of high flow events had significantly higher importance compared to flow magnitude, duration, and frequency metrics, especially for fish assemblages in GW and INT streams. The frequency and timing of low flow events had high importance for fish assemblage turnover across all stream flow classes, while the magnitude of low flows and the magnitude and rate of change of average flows was most important for INT stream fish assemblages. In addition to benefiting multi-species conservation and management actions through identification of local and regional flow-ecology relationships generalized across different flow regimes, the results of this study provide a better understanding of complex nonlinear threshold effects, which is critical to anticipating changes in aquatic ecosystems and communities.

1. Introduction

Across trophic levels, variations in the magnitude, frequency, timing, duration, and rate of change of streamflow play a significant role in mediating the behavior, morphology, and life-histories of aquatic organisms at multiple spatial scales (Poff et al. 1997, Bunn and Arthington 2002, Lytle and Poff 2004, Matthews and Marsh-Matthews 2017). Natural variations in streamflow are important in linking and mediating a number of important ecosystem processes (Palmer and Ruhi 2019), including stream channel geomorphology (Poff et al. 2006), floodplain connectivity (Junk et al., 1989), aquatic habitat availability (Freeman et al. 2001, Kennard et al. 2007), sediment and solute transport (Junk et al. 1989, Bunn and Arthington 2002), and rates of instream microbial decomposition of organic material (Yeung et al. 2018). In addition to having an important influence on morphological divergence and filtering of life history traits (Poff 1997, Bruckerhoff et al. 2019), predictable seasonal high and low flow events function to regulate species competition (Hasegawa and Yamamoto 2010), spawning and hatching (Freeman et al. 2022), juvenile abundance, growth, and survival (Freeman et al. 2001), predator–prey interactions (Power et al. 1996), and can help limit the spread of invasive species (Marchetti and Moyle 2001). Given the fundamental role of streamflow in maintaining the structure and functioning of riverine ecosystems, understanding how distinct hydrologic characteristics of stream flow regimes, along with

watershed-scale disturbance influence aquatic species and the key interactions supporting them, is essential for sustainable conservation and management of native freshwater biodiversity (Palmer and Ruhi 2019, George et al. 2021, Freeman et al. 2022, Ziegeweid et al. 2022).

Ecological responses caused by the interaction of species life-history traits with hydrologic and environmental alterations resulting from natural and anthropogenic factors can lead to abrupt changes in aquatic communities over time (Dodds et al. 2010, Palmer and Ruhi 2019). As global consumptive demand for freshwater grows, the extraction, impoundment, and diversion of surface and groundwater for anthropogenic uses has resulted in the partial or complete alteration (i.e. regime shifts) of the natural flows of many rivers and streams (Lytle and Poff 2004, Döll et al. 2009, Carlisle et al. 2011, Mims and Olden 2013), and may contribute to shifts of stream fish assemblages into alternative stable states (Pelletier et al. 2020). For example, increasing urbanization and impervious surface area in watersheds can lead to fish assemblage shifts that favor habitat generalist species (Roy et al. 2005), while higher runoff and storm flows can decrease juvenile survival for many fish species (Matthews and Marsh-Matthews 2017). Many species may experience detectible changes in occurrence as the result of hydrologic and environmental disturbance (Carlisle et al. 2011), and threshold analysis can be used to identify the point at which species and assemblages begin to respond to key hydrologic characteristics (Dodds et al. 2010, King and Baker 2010). While responses to anthropogenic change may be linear, aquatic communities often respond nonlinearly to environmental disturbance or increasing anthropogenic land use intensity (Rosenfield 2002, Allan 2004). Nonlinear relationships between environmental drivers and ecological responses arise when interactions among ecosystem components are not directly proportional and a high rate of change occurs abruptly over a narrow range along an environmental gradient (Allan 2004, Davies and Jackson 2006). The frequency and strength of nonlinear responses are predicted to increase in aquatic ecosystems under future land use and climate change scenarios, particularly in systems already approaching existing thresholds (e.g. temperature, flow) or where communities are exposed to multiple stressors (Rosenfeld 2017, Walker and Walters 2019). There is a crucial need to better understand how species and assemblages respond to hydrological variations and watershed disturbance related to anthropogenic land uses, especially where ecological breakpoints are likely to occur.

Hydrologic metrics characterizing primary flow regime components (e.g duration, frequency, magnitude, timing, and rate of change, Poff et al. 1997) along with their seasonality and variability may be calculated directly from long-term observed hydrologic timeseries data (Carlisle et al. 2017, Eng et al. 2017), or using process-driven models which simulate streamflow and contributing hydrologic processes across catchment areas (Mohammadi et al. 2021). Data-driven machine learning methods have proven to be an efficient tool for modeling nonlinear hydrologic processes with high accuracy (Mohammadi 2021). In this study, we provide a quantitative description of the relationships between stream fish assemblage composition and turnover, hydrology, spatial, and watershed-scale variables describing selected anthropogenic stressors across different stream flow regimes. We used a nonparametric multivariate gradient forest modeling approach (Ellis et al. 2012) to identify hydrologic and environmental thresholds where abrupt changes in stream fish assemblage composition occur. Gradient forest aggregates multiple species-specific random forest machine learning models (Breiman, 2001) to make predictions of variable importance in driving assemblage change along a predictor gradient. It does so by building monotonic functions for predictors through the partitioning of species occurrence data (presence-absence or abundance) to provide information about patterns of species- and assemblage-level turnover (Nieto-Lugilde et al. 2018). This incremental fitting approach differs from other community modeling methods (e.g. generalized dissimilarity modeling), making gradient forest particularly well suited to the analysis of large datasets representing complex

relationships between environmental predictors and species' responses with high predictive power (Ellis et al. 2012).

In addition to identifying regional flow-ecology relationships for fishes, generalized across different flow regimes, we tested the hypotheses that predictor cumulative importance varied significantly by metric category and stream flow class. Our results provide information on the importance of timing, frequency, duration, magnitude and rate of change of high, average, and low flow conditions, together with seasonal stochasticity and watershed-scale anthropogenic disturbance. Our results can further aide refinement and targeting of habitat and multispecies conservation and management actions in aquatic ecosystems by providing a better understanding of complex nonlinear relationships, which is critical to anticipating changes in aquatic ecosystems and communities.

2. Methods

2.1. Study area

Our study area was defined based on the spatial extent of available fish sampling records and included rivers and streams in the National Hydrography Dataset Plus Version 2 (NHDPlusV2) Database within the Ouachita-Ozark Interior Highlands and West Gulf Coastal Plain regions of Arkansas, Missouri, and eastern Oklahoma, U.S.A., bounded by Level III ecoregions (Fig. 1). North American ecoregions are designated at varying levels of coarseness ranging from Level I-IV (finest scale) and denote areas where similar ecosystems and environmental resources are concentrated. The study area is both hydrologically and geographically diverse, encompassing eleven Level III ecoregions with most occurring in more than one state. Leasure et al. (2016) identified seven natural flow regimes in the study region including: Groundwater stable (GS), Groundwater (GW), Groundwater flashy (GF), Perennial runoff (PR), Runoff flashy (RF), Intermittent runoff (IR), and Intermittent flashy (IF). These seven distinct flow regimes can be grouped into three broad categories – Groundwater (GW), Runoff (RO), and Intermittent (INT) streams. Groundwater streams tend to be larger perennial systems with low flow variability and high constancy from significant groundwater recharge and are predominately found in the Ozark Highlands. In general, the interaction between surface and groundwater is lower in the Osage Plains, Boston Mountains and St. Francois Mountains but increases in the Springfield Plateau,

Salem Plateau and the Mississippi Alluvial Plain (Adamski 1995). Runoff streams represent the most abundant and spatially-distributed flow regime across the study area but are most prevalent in the Boston Mountains. Runoff streams have lower base flow, average more low flow spells than groundwater streams, and those receiving some groundwater recharge are rarely reduced to zero flow and therefore may be considered perennial. More flashy runoff streams may experience zero flow days, but compared to intermittent streams they recede slower, have fewer days of no flow, and less daily flow variability. Streams with intermittent flow regimes are found throughout the region, but particularly in the Ouachita Mountains, and have relatively small drainage areas (<1,000 km^2) averaging 2 to 8 weeks, up to several months of zero-flow per year. High flows in these intermittent streams also tend to recede much faster than in the other flow regimes (Leasure et al. 2016).

2.2. Biological data

We complied long-term (1916–2016) fish survey records from the USGS National Gap Analysis Project (Aquatic GAP) (Scott et al. 1993, Sowa et al. 2007), the Missouri Department of Conservation (MDC), the Oklahoma Natural Heritage Inventory (ONHI), and the Oklahoma Conservation Committee (Table 1). Preprocessing of species datasets consisted of standardizing species names to correct for misspelled or duplicate records and to remove uncertain taxonomic identifications. Records for hybrids were also removed and subspecies designations
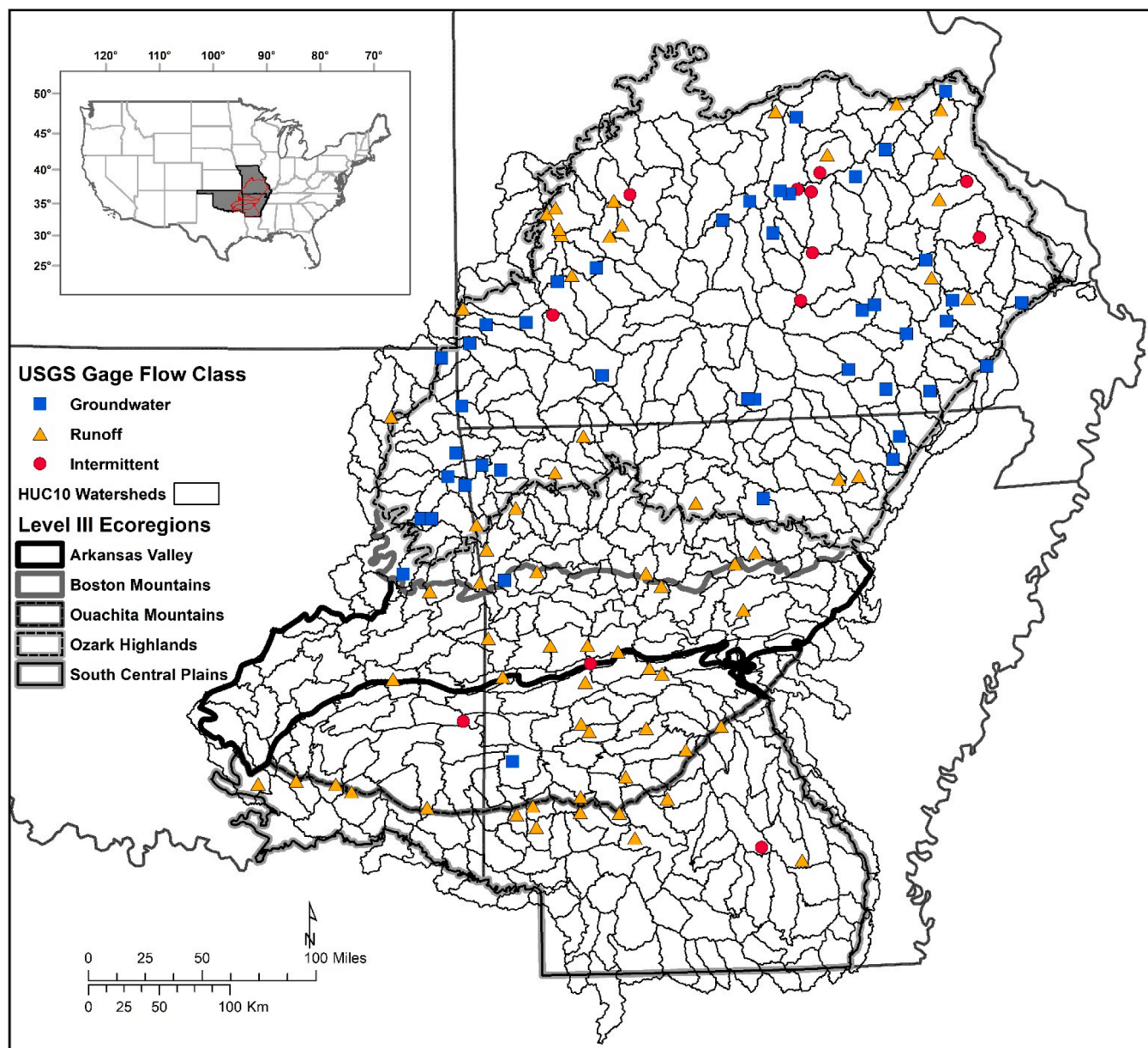
Fig. 1. Map of the study area showing the locations and flow classes of USGS stream gages, along with level 10 hydrologic units (HUC10) and level III ecoregions.

Table 1

Data sources and information for fish occurrence data spatially subset for use in the gradient forest analyses.

State	Records	Date Range	Citation
Arkansas Game and Fish Commission	70,855	1927–2009	Personal/Agency Communication
Missouri Department of Conservation	117,373	1923–2012	Personal/Agency Communication
Oklahoma Natural Heritage Inventory	35,908	1916–2016	Personal/Agency Communication
Oklahoma Conservation Committee	13,807	2000–2016	Personal/Agency Communication

were truncated to species level before converting the dataset to a presence-absence table for analysis. We used species occurrence data in place of abundance to help control for differences in fish sampling methodologies, effort, and gear (Guo and Olden, 2014; Fox and Magoulick 2019). Analysis of changes in fish species occurrence places a greater emphasis on coarser-scale regional species patterns, in contrast to abundance data, which tends to emphasize finer-scale local patterns (Allen and Starr 1982).

Fish records were grouped according to flow regime into Groundwater (GW; n = 319 stream reaches), Runoff (RO; n = 396), and Intermittent (INT; n = 63) stream classes in order to increase sample size and focus on broad categories of response. Since our focus was on patterns associated with representative fish assemblages, stream sampling sites with fewer than five recorded species were excluded from the analysis to help control for the effects of surveys targeting individual fish species. Fish species with low representation, defined as 5 percent or less of the maximum number of fish records from streams in each flow class were also excluded from the analysis. The resulting dataset consisted of n = 17,375 records for 187 fish species collected from 778 stream reaches across the study area (Table 2). The NHDPlusV2 stream network database was used to assign each georeferenced fish record a unique 14-digit USGS stream reach code using linear referencing in ArcMap v.10.5 (Environmental Systems Research Institute, Redlands, California), with a snapping distance of 200 m. Stream reach codes form the basis of the

Table 2

Sampling locations and fish collection records for flow classes with averages and standard deviations for upstream drainage area, elevation, daily flow, and Hydrologic Disturbance Index (HDI). GW = groundwater, RO = runoff, INT = intermittent.

Flow Class	Stream Sites	Fish Records	Drainage Area (km^2)	Elevation (m)	Daily Flow (cfs)	HDI
GW	319	7186	2390 (±1969)	200 (±77)	979 (±786)	13 (±6)
RO	396	9025	1428 (±1245)	185 (±76)	687 (±662)	14 (±5)
INT	63	1164	70 (±118)	273 (±72)	200 (±363)	8 (±4)

NHD stream linear referencing system and allows georeferenced observations to be linked to specific points along a stream network. The fish assemblages in our analysis included all fish species sampled in a single stream reach, assigned to one of the three representative stream flow regime classes (GW, RO, INT) across the study area. As such, the fish assemblages are more representative of regional-scale fish biodiversity rather than local-scale biodiversity. This is an important distinction given that, while local fish assemblages may show considerable variation over time, turnover at regional scales tends to be more stable over both short (e.g 1 year) and long (e.g. decadal) time scales (Zbinden 2020). It is also at the regional scale that variations in stream community structure generally begin to reflect important hydrological patterns, influences, and constraints (Poff 1997).

2.3. Hydrologic metrics

Daily stream flow measurements (cfs, ft^3/s) recorded by 111 USGS stream gauging stations located across the study area (Fig. 1) were compiled from the National Water Quality Information System. USGS gages included in the analysis had minimum periods of record of 15 complete years (mean of 37 years) with <7 days of contiguous missing data. Daily streamflow time series data were processed in the R open source programming environment using the "EflowStats" package, a reimplementation of the Hydrologic Index Tool (HIT) (Henriksen et al. 2006) to determine average daily flow, peak annual flow, and flood recurrence threshold. These statistics were used to calculate 171 HIT metrics (Appendix S2), grouped into five categories: magnitude (n = 94), frequency (n = 14), duration (n = 44), timing (n = 10) and rate of change (n = 9) in flow events, which together define the natural flow regime and ecological health of river systems (Richter et al. 1996, Poff et al. 1997, Olden and Poff 2003, Poff et al. 2010, Poff and Zimmerman 2010). An additional six flow statistics describing the seasonal and stochastic characteristics of streams (Mag6) were calculated for each stream gage, including: the coefficient of daily variation (tau2); skewness of daily streamflow (tau3); kurtosis of daily streamflow (tau4); the autoregressive lag-one correlation coefficient (AR1); the amplitude; and the phase shift of the seasonal signal (Archfield et al. 2014). The first three metrics (tau2-4) describe the magnitude of stream flow, while the AR1 coefficient describes the persistence of stream flow from one day to the next, and therefore is a proxy for the duration and rate of change of flow events. The amplitude and phase shift of the sinusoidal seasonal signal were estimated based on a spectral analysis of daily stream discharge time series (Warner 1998) and are functions of climate and catchment attributes, which together affect the timing of stream flow events (Archfield et al. 2014). Hydrologic data for respective USGS stream gages were range standardized prior to analysis by dividing each metric by their absolute values to transform the metrics to comparable scales (Bond and Kennard 2017).

2.4. Hydrologic disturbance Index

Additional model predictors included the Hydrologic Disturbance Index (HDI), which is part of the Geospatial Attributes of Gages for Evaluating Streamflow, version II (GAGESII) developed by Falcone et al. (2010). The HDI includes seven variables describing important anthropogenic stressors within each USGS gage's watershed (Table 3). HDI values for the USGS gages included in our analysis ranged from 4 to 27, with an arithmetic mean and median of 14 (sd = 5.7; Table 2). In comparison, HDI values range from 1 to 42 across the contiguous United States, with a national median value of 15 (Falcone et al., 2010, Falcone, 2011).

2.5. Georeferencing fish assemblage datasets

The NHDPlusV2 stream network was used as the basis for spatially joining the fish assemblage records to the HIT and Mag6 metrics and HDI variables from the nearest USGS stream gage up to a network distance of 10 km, using linear referencing in ArcMap v.10.5. This distance was conservatively chosen based on observations that ecological data and hydrologic metrics were readily transposable up to 25 km, after which prediction uncertainty rapidly increased (Bond and Kennard 2017).

2.6. Gradient forest model development

Gradient forest models were implemented in the R packages "gradientForest" and "extendedForest", which also rely on the "randomForest" package (Ellis et al., 2012; R Core Team, 2013). For each stream flow class (GW, RO, INT), separate gradient forest models were parameterized and run for HDI and Mag6, high, average, and low flow HIT metrics. Gradient forest models aggregate the results of Random

Table 3

List and definitions of Hydrologic Disturbance Index (HDI) variables included in the gradient forest analysis. See Appendix 2 for a detailed list and definitions of the Hydrologic Indices Tool (HIT) metrics.

HDI Variable	Definition	Unit
FRESHW_WITHDRAWAL	Freshwater withdrawal from 1995-2000 county-level estimates	m^3/year/km^2
MAJ_DDENS_2006	Major dam density (> 50 feet in height (15m) or having storage > 5,000 acre feet)	#/100 km^2
X2006_STOR	Change in all dam (reservoir) storage (not just major dams), 1950 to 2006	megaliters/km^2
sum_percent_canals_artif	Stream kilometers coded as "Canal", "Ditch", "Pipeline", or "Artificial Path" in NHDPlusV2.	percent
DIS_ADJ_NEAREST_MAJ_NPDES	Adjusted distance to nearest major NPDES site.	max raw distance minus actual distance
ROADS_KM_SQ_KM	Road density from Census 2000 TIGER roads	km of roads per watershed km^2,
FRAGUN_BASIN	Fragmentation Index of "undeveloped" land in the watershed (e.g. all land which is not urban nor agriculture)	1 - % undeveloped pixels completely surrounded by other undeveloped pixels. High numbers = less undeveloped and unfragmented land cover in basin. Based on Riiters et al., 2000 "interior" pixels calculation

Forest models (Breiman, 2001) for individual species represented in assemblages at each sampling site and fitted with n = 999 classification trees using the default settings for the correlation threshold used in the conditional importance calculations for hydrologic and environmental variables. Gradient forest predictor importance is calculated based on a drop in model performance under random permutation using a conditional approach to account for the inflation of variable importance measures for correlated predictors. Under conditional permutation, the predictor variables were permuted within blocks of observations defined by splits in a given tree only if the correlation with another variable was higher than a specified threshold ($r > 0.05$) up to a maximum number of splits (log2(n × 0.368 ⁄ 2)), where n equals the number of sites included in the analysis (Ellis et al. 2012). Model splits were compacted in bins (n = 201) to prevent memory problems associated with the analysis of large species datasets. While the use of binning in species distribution and other ecological studies has been criticized on the basis it can discard important information (Gray et al. 2006), the large number of trees and bins used in the implementation of the gradient forest models limits the potential for information loss (Ellis et al. 2012). We assessed the overall importance for a predictor using a weighted R^2 by averaging across the results for all species in an assemblage. An increase in predictor importance relies on the deviance reduction (i.e. decrease in residual sums of squares) when each variable is permuted in turn (Pitcher et al. 2012). Split values were cumulatively summed along each predictor gradient to construct a monotonic nonlinear turnover function describing the magnitude and rate of assemblage composition change, weighted by the model accuracy and specific importance of a variable for each species (Ellis et al. 2012, Pitcher et al. 2012). Steeper slopes or steps of the response curves indicate values along each gradient where variables have a higher cumulative importance in species turnover and assemblage change.

Model performance was assessed using an analogue of a linear regression R^2 for classification trees, R_c^2 (1 - OOB misclassification rate / base error rate), where base error rate equals $2p(1 - p)$ and p is the prevalence of the species. In this sense R_c^2 is analogous to a linear regression R^2 because it is 0 when the model has no predictive power and 1 when the model predicts perfectly (Ellis et al. 2012). The gradient forest model retains only species having $R_c^2 > 0$ and these species contribute to the combined estimates of assemblage turnover, while species with no predictive power are excluded from the model. Normalizing standardized splits for all species by the model error rate and accumulating provides the empirical nonlinear functions for assemblage cumulative importance curves relating compositional change along the gradient of each explanatory variable. Statistical differences ($\alpha < 0.5$) in the equality of the distributions of cumulative importance values among the different metric categories and between flow regime classes were tested using non-parametric Fisher-Pitman permutation tests implemented in the R "coin" package (Zeileis et al. 2008). We approximated the conditional null distribution of the test statistic to obtain p-values via Monte Carlo resampling (n = 9999) using asymptotic approximation of the exact distribution of the unique sums of squares (Type III SS) within an ANOVA framework.

3. Results

Gradient forest models retained a total of 155 fish species for GW streams, 153 species for runoff (RO) streams and 83 species for intermittent (INT) streams. Average model prediction errors (1-relative error rate) for individual fish species were $R_c^2 = 0.44$ (SD = 0.15) for groundwater (GW) streams, $R_c^2 = 0.40$ (SD = 0.14) for RO, and $R_c^2 = 0.48$ (SD = 0.22) for INT streams (See Appendix S1 Figures S1-S3 for summary of model fit for fish species). Fisher-Pitman permutation tests indicated significant differences in the distributions of cumulative importance values among spatial, HDI and Mag6 variables ($\chi^2 = 12.305$, p-value = 0.0012), high flow metrics ($\chi^2 = 29.803$, p-value < 0.0001), and low flow metrics ($\chi^2 = 29.802$, p-value < 0.0001). Pairwise

comparisons of metric categories showed spatial variables (latitude, longitude, elevation, and upstream drainage area) had significantly greater cumulative importance (p-value = 0.003) compared to HDI variables and Mag6 metrics (Fig. 2). The timing of high flows had significantly greater cumulative importance (p < 0.0001) compared to high flow magnitude, duration, and frequency (Fig. 2), while the frequency and timing of low flows had significantly greater importance (p < 0.001) compared to low flow magnitude and duration (See Appendix S1: Figs. S4-S6 for a summary of predictor importance for fish distributions).

Examining differences between stream flow classes showed the distributions of cumulative importance values for spatial and HDI metrics were not statistically significant for GW, RO, and INT flow classes. However, the cumulative importance of Mag6 variables was significantly higher for INT streams compared to GW (p = 0.005) and RO streams (p = 0.011) (Fig. 2). The timing of high flows was of lower importance for fish assemblage turnover in RO streams (p = 0.049) compared to INT streams, while the magnitude of high flows had significantly higher importance for INT stream fish assemblages (p = 0.02) compared to those in RO streams. The magnitude (p = 0.03) and rate (p = 0.02) of mean flows were of significantly higher importance for fishes in INT streams compared to RO streams (Fig. 2). In the case of low flow metrics, flow magnitude had significantly higher cumulative importance for INT stream fish assemblages (p = 0.003). See Appendix S1: Tables S1-S4 for a full summary of the statistical comparisons.

3.1. Spatial, HDI and Mag6 variables

Spatial variables including latitude, longitude, elevation, and upstream drainage area had the highest importance for all three flow regimes, but the specific relationships varied among the different metric categories and flow classes. Latitude was most important for fish assemblage compositional change in GW and RO streams and was of higher importance compared to longitude for all three flow classes (Figs. 2 and 3). Upstream drainage area (km²) was most important for fish assemblage turnover in INT streams, with an apparent threshold at around 150 km² (Fig. 3). Elevation was most important for turnover in RO streams and least important in GW streams. Focusing on HDI variables, change in dam storage from 1950 to 2006 (X2006_STOR) had the highest importance for fish assemblages in GW streams, with a substantial threshold apparent at a value of 90 megaliters/km² and another at 250 megaliters/km². Basin fragmentation was of highest importance in RO and INT streams, with thresholds evident along the basin fragmentation gradient for INT stream fish assemblages at low (20 %), moderate (35 %) and high (55 %) values. In RO streams, the cumulative importance of fragmentation increased at a more gradual rate with a threshold along the gradient apparent at 70 % of watershed area. Mag6 metrics were of highest importance for INT stream fish assemblage change and were of least importance in RO streams. The phase of the seasonal signal was the most important Mag6 metric across all flow regimes and had the highest importance for INT stream fish assemblages (Figs. 2 and 3).

3.2. High flow metrics

In GW streams, both the Julian date of annual maximum (th1) and variability in Julian date of annual maxima (th2) were among the most influential flow metrics (Fig. 2). A threshold was evident aligned with a late-January arrival of maximum flows (th1), earlier than the typical period of maximum flows in study area streams from February to May (Fig. 4). The cumulative importance of variability in the date of annual maximum flow (th2) for fish assemblages in GW streams increased across the gradient of temporal variability, with a breakpoint equivalent to a Julian date in early March. The skewness of annual maximum flows (mh19), variability across maximum monthly flows (mh13), and variability of annual maximum of 90-day moving average flows (dh10) were
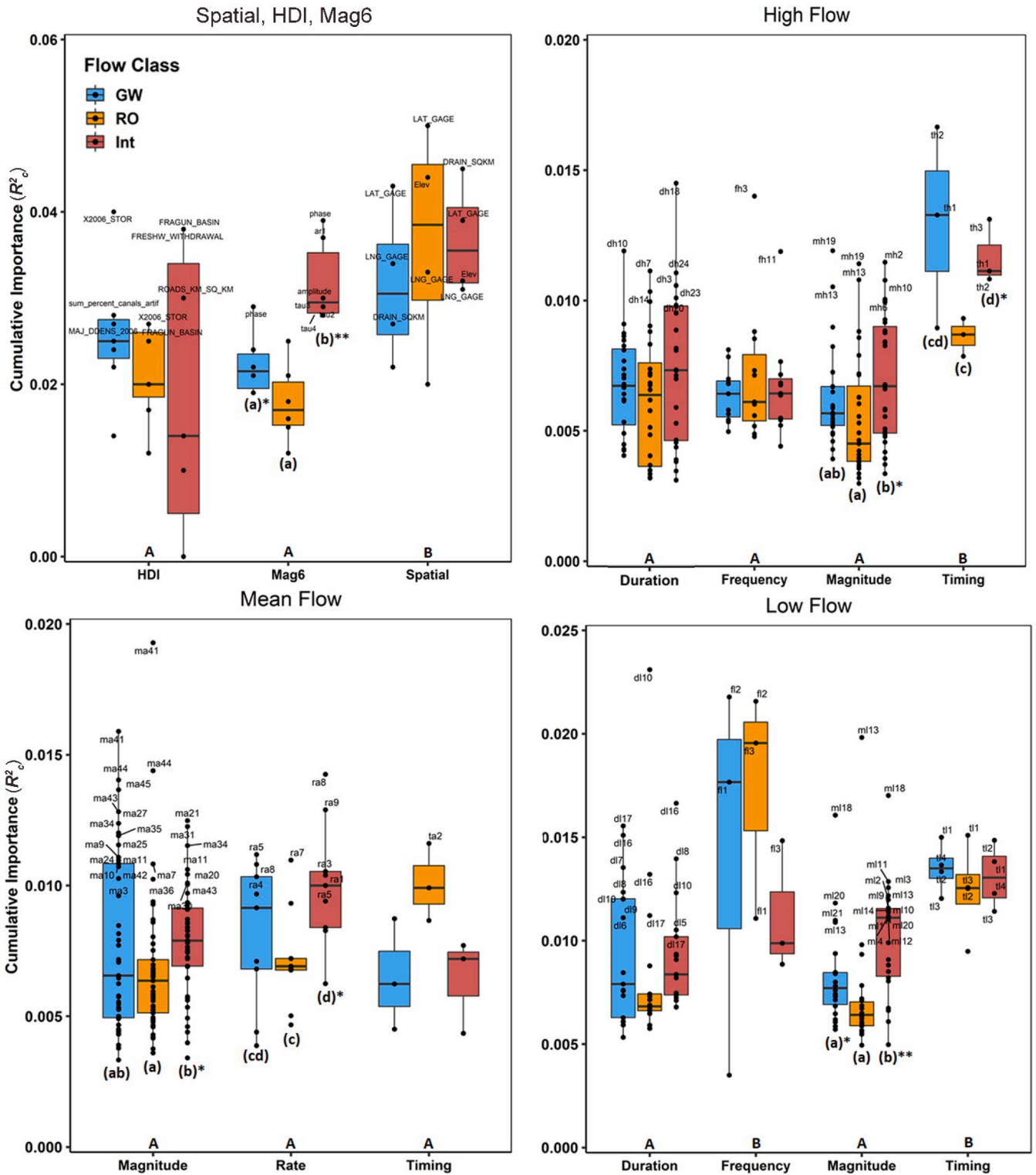
Fig. 2. Box plots summarizing the distributions in gradient forest model cumulative importance values for variables and metric categories for groundwater (GW), runoff (RO) and intermittent (INT) flow regimes. The y-axis labels differ among the plots and points are labeled for metrics with cumulative importance values ≥ 0.025 (Hydrologic Disturbance Index; HDI and Mag6) and ≥ 0.1 (High, Mean, Low flow). Lowercase letters in parentheses show statistical relationships between stream flow classes. Uppercase letters along the x-axis compare the equality of the distributions of cumulative importance values among the different metric categories. Shared letters indicate no significant difference (α > 0.5) and only categories and flow classes with significant differences are labeled. Significance of variable coefficients: p ≤ 0.001***, p ≤ 0.01**, p ≤ 0.05*. See Appendix S1: Tables S1-S4 for a full summary of the statistical comparisons. Box plots depict the minimum, first quartile, median, third quartile, and maximum, with outliers depicted as single points.
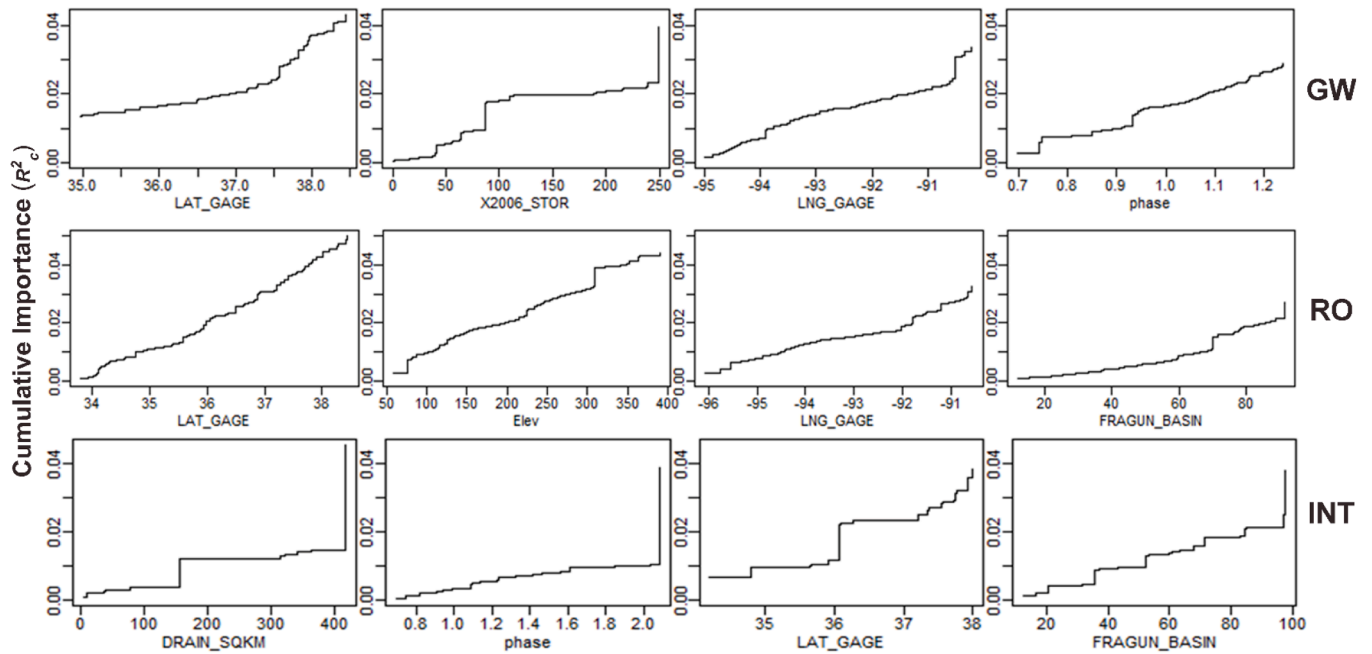
Spatial, HDI, Mag6

Fig. 3. Gradient forest cumulative importance curves for spatial, Hydrologic Disturbance Index (HDI), and Mag6 variables constructed using the split information contained in the individual species models and indicating relative patterns of compositional change of fish assemblages in groundwater (GW), Runoff (RO), and Intermittent (INT) streams. Plots for each predictor (and in Figs. 4-6) show cumulative R_c^2 weighted importance distributions standardized by density of observations, averaged over all species. Steeper slopes or steps of the response curves indicate values along the gradient of each predictor where variables gain higher importance in assemblage turnover. For example, upstream drainage area (km^2) was most important for fish assemblage turnover in INT streams, with an apparent threshold at around 150 km^2 indicating a pronounced transition between fish assemblages across intermittent stream sites.
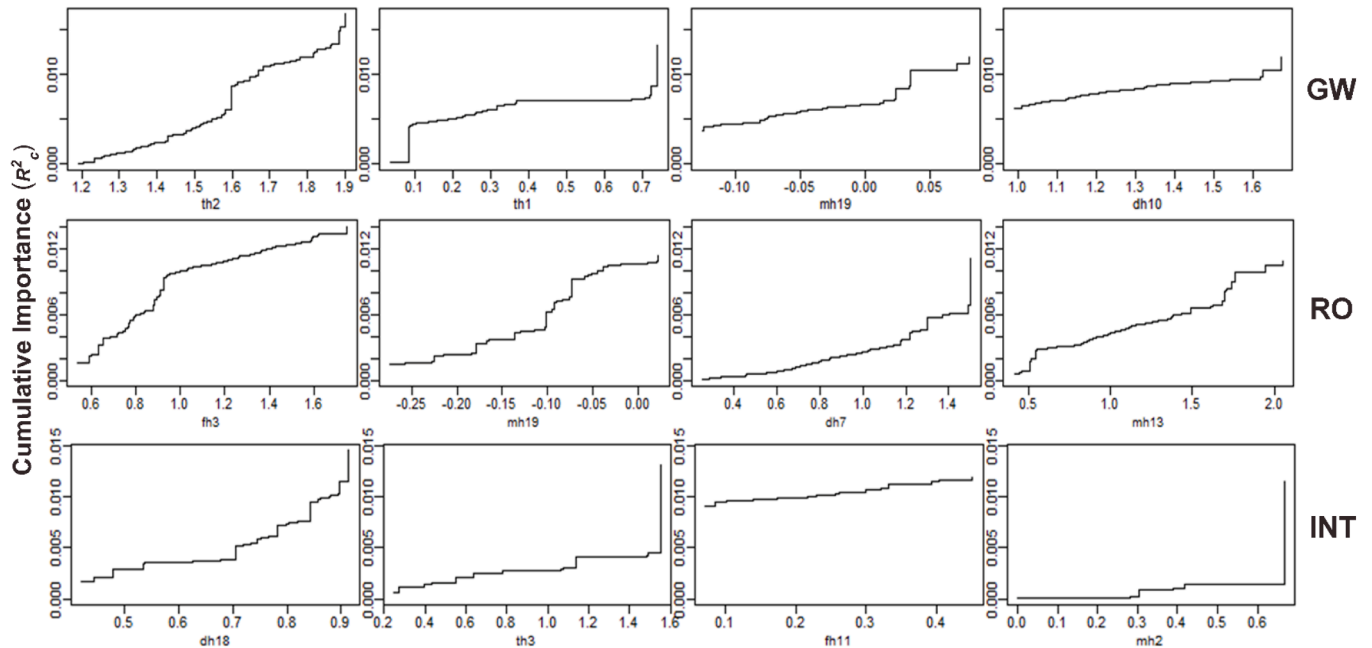
High Flow

Fig. 4. Gradient forest cumulative importance curves for high flow metrics. Timing of high flow events including the Julian date of annual maximum (th1) and variability in Julian date of annual maxima (th2) had the highest importance for fish assemblage turnover in groundwater (GW) streams. High flood pulse count (fh3) was the most important high flow metric in runoff (RO) streams, and high flow duration (dh18) was most important in intermittent (INT) streams.

also important in GW streams. High flood pulse count (fh3) was the most important high flow metric for fish assemblage turnover in RO streams and increased in importance rapidly up to a range adjusted value of 1.0, or an average of approximately 75 days per year with high flows above a

threshold equal to three times the median flow for the entire record (see definitions in Appendix S2). The average value of fh3 for RO streams was 94 days, compared to 55 days in GW streams and 96 days in INT streams. High flow duration (dh18) was the most important predictor of

compositional change for INT stream fish assemblages. Cumulative importance of dh18 at the assemblage-level increased in a stepwise fashion with a notable increase in the slope of the curve at a value of 0.7, equivalent to an average high flow duration of 18 days above a threshold equal to three times the median flow value for the entire flow record.

3.3. Mean flow metrics

Average annual runoff by drainage area (ma41, cfs/sq. mile) and metrics describing the magnitude of annual flow variability (ma44, ma43) and skewness of annual flows (ma45) had the highest cumulative important for fish assemblage turnover in GW streams (Fig. 2). Annual runoff (ma41) was also the most important variable for fish assemblage turnover in RO streams, with the greatest rate of increase between values of 0.1–0.2, or 0.98 and 1.7 cfs/sq. mile (Fig. 5). Also of high importance in RO streams was the magnitude of annual flow variability (ma44) and predictability of average flows (ta2), which is composed of two independent, additive components: constancy (e.g. temporal invariance) and contingency (a measure of periodicity). Interestingly, while ta2 was one of the most influential metrics in RO streams, it was not identified as an important metric for fish assemblages in either GW or INT streams. Instead, metrics describing the rate of change of average flow events, the annual number of flow reversals (ra8) and the variability of flow reversals (ra9) were of highest importance to assemblage turnover in INT streams, in addition to the magnitude of mean flows in October (ma21), and flow variability in August (ma31).

3.4. Low flow metrics

Frequency metrics describing the variability of low pulse count (fl2) and count of low flow pulses (fl1) had the highest importance in GW streams in addition to variability in base flow (ml18), and variability in low pulse duration (dl17) (Fig. 2). In the case of fl1, the average number of low flow events in GW streams below a threshold equal to the 25th percentile for the entire flow record increased in importance between the values of 0.2–0.3, or an average of 6 to 9 low flow events per year (Fig. 6). The variability of low pulse count (fl2) had high assemblage-

level cumulative importance for fish assemblage turnover in both GW and RO streams, although the fl2 metric importance increased more rapidly and remained high across its range in GW streams. In RO streams, variability of annual minimum of 90-day moving average flow (dl10), variability of low pulse count (fl2), the coefficient of variation across minimum monthly flow values (ml13), and the frequency of low pulse spells (fl3) were the most important metrics. The fl3 metric cumulative importance increased sharply in RO streams at a value of 0.3, or an average of around 11 low flow events pre year below a threshold equal to 5 percent of the mean flow value for the entire flow record. The variability in base flow (ml18) had the highest importance for fish assemblages in INT streams, followed by low flow pulse duration (dl16), variability in Julian date of annual minima (tl2) and frequency of low pulse spells (fl3).

4. Discussion

Identifying and generalizing regional flow-ecology relationships across different stream flow regimes is a crucial step towards anticipating future changes in aquatic ecosystems and communities. We applied a gradient forest ensemble machine learning approach to better understand how fish assemblages vary across gradients of hydrology, geography, and anthropogenic-source hydrologic disturbance. Our results provide insight into the potential drivers and key hydrologic metrics related to spatial turnover of fish assemblages in different stream flow regimes to identify where ecological thresholds are likely to occur.

Spatial variables had a high importance for fish assemblage change across all three flow regimes. While latitude was more important in groundwater (GW) and runoff (RO) streams, upstream drainage area was of highest importance for fish assemblage turnover in intermittent (INT) streams. Bruckerhoff et a. (2019) also found spatial factors described a large proportion of variance in fish traits within flow regimes, and watershed area was an important predictor of fish assemblage traits in intermittent streams. In addition to being strongly related to fish species richness, and especially darter (Percidae) species diversity (Matthews and Robison 1998), drainage area has been identified as an important determinant for predicting low flow metrics along with
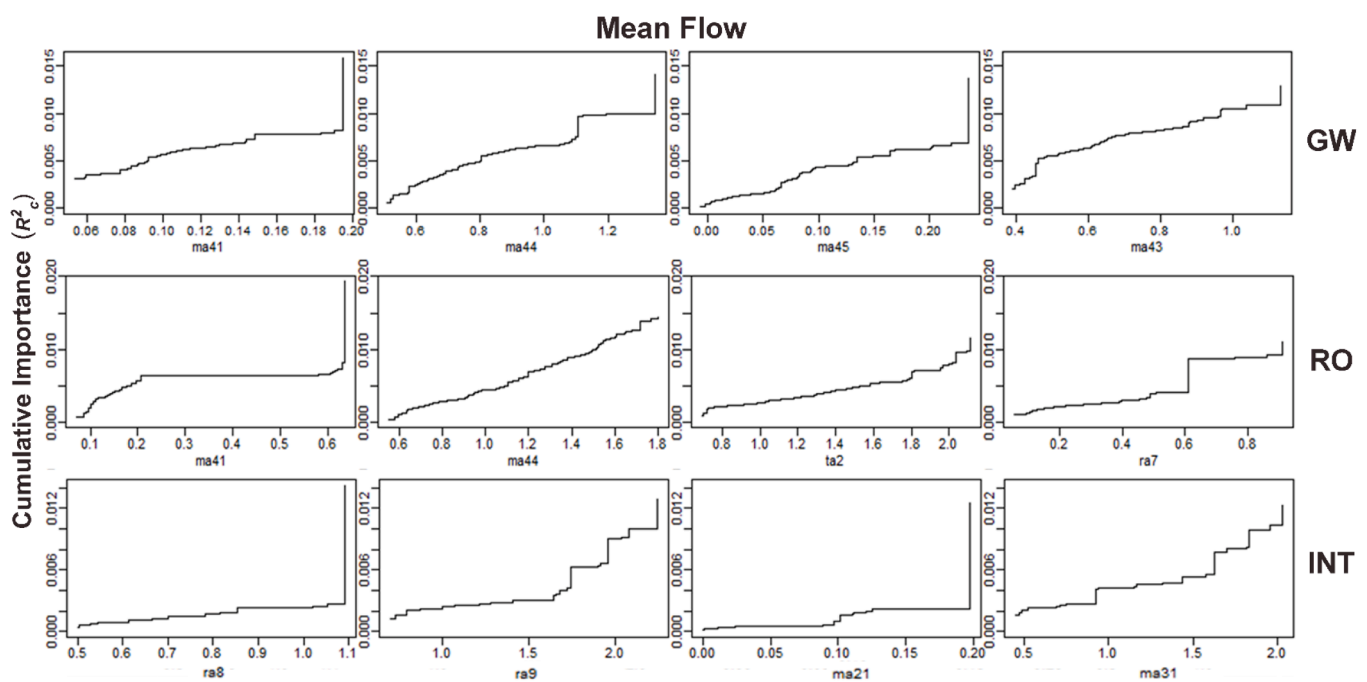

Fig. 5. Gradient forest cumulative importance curves for mean flow event metrics. Average annual runoff by drainage area (cfs/km^2, ma41) and annual flow variability (ma44) had the highest cumulative importance for fish assemblage turnover in groundwater (GW) and runoff (RO) streams. Annual number of flow reversals (ra8) and the variability of flow reversals (ra9) had the highest importance in intermittent (INT) streams.
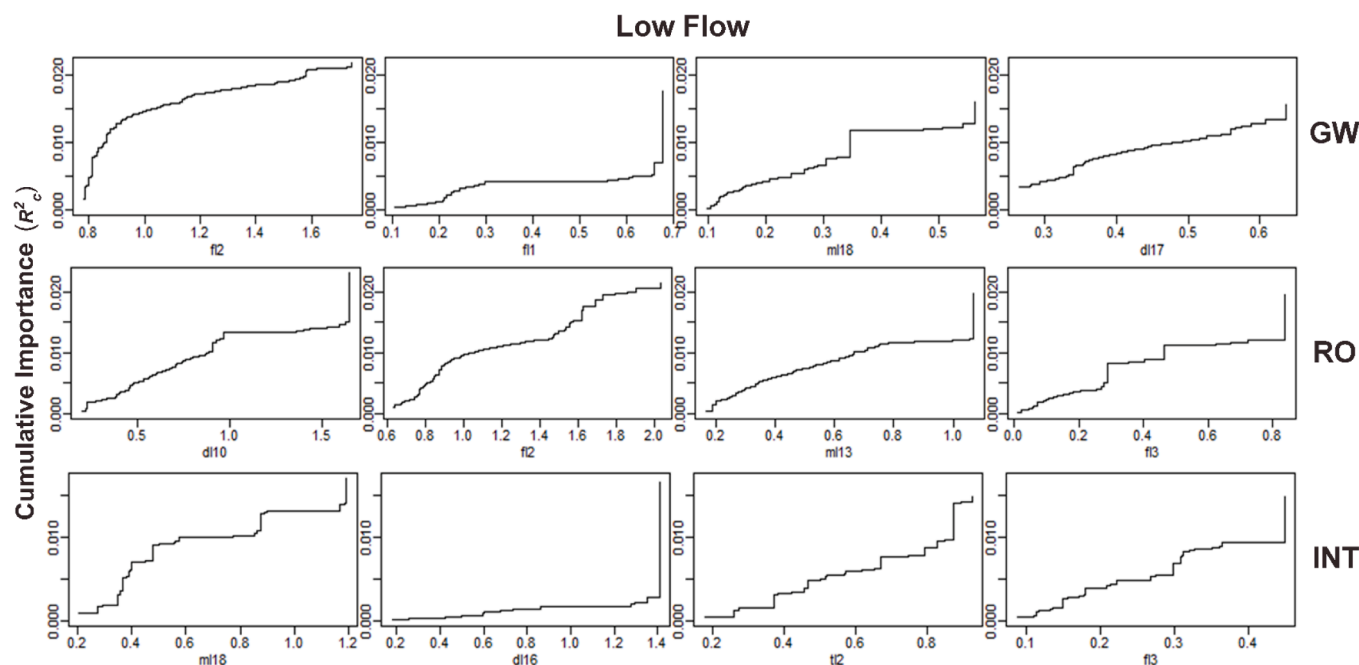
Low Flow

Fig. 6. Gradient forest cumulative importance curves for low flow metrics constructed using the split information contained in the individual species models and indicating relative patterns of compositional change of fish assemblages in groundwater (GW), runoff (RO), and intermittent (INT) streams. Frequency of low flows, including the variability (fl2) and number of low flow pulses (fl1) had the highest importance for fish assemblage turnover in GW streams. Variability in the annual minimum of 90-day moving average flows (dl10) was the most important metric for RO streams, and variability in base flow (ml18) in INT streams.

streamflow intermittency (Reynolds et al. 2015). Our analyses also showed both fragmentation of undeveloped land and water withdrawals were of high importance for intermittent stream fish assemblages. Watershed fragmentation was also an important gradient for fish assemblage composition in runoff streams, although assemblage-level turnover functions showed a different pattern of increase over the fragmentation gradient in runoff streams, compared to intermittent streams. For example, intermittent stream fish assemblage turnover exhibited multiple thresholds at low (20 %), moderate (35 %) and high (55 %) watershed fragmentation values (Fig. 3), while in runoff streams the importance of fragmentation increased more gradually, with a single threshold apparent at 70 %, indicating an important influence arising at a considerably higher percentage of developed (i.e. urban and agricultural) land area. Given that intermittent streams in our study area (and in general) have the smallest drainage areas of the three flow classes, protecting intermittent stream watersheds from anthropogenic-source disturbance should be considered a high priority for management and conservation activities.

Variables describing seasonal and stochastic stream flow properties (Mag6) were significantly more important for intermittent stream fish assemblages compared to groundwater and runoff streams. The phase shift of the seasonal signal, representing the time lag in the arrival of peak flow across study years, was the most important Mag6 variable for all flow regimes but especially for intermittent stream fish assemblages (Fig. 2). Wetter catchments and those with impervious bedrock tend to partition more rainfall and runoff into fast flow, leading to a higher amplitude and a smaller phase shift, while drier stream catchments, as well as those with a larger proportion of carbonate sedimentary geology (i.e. limestone and dolostone karst), generally have larger and more variable phase shifts and lower amplitudes (Gnann et al. 2020). In our study area, the amplitude of the seasonal signal had a higher relative importance in runoff streams, compared to groundwater and intermittent streams where amplitude was considerably less important.

Many streams in our study area tend to have strong flow seasonality, particularly those located in the Ozark Highlands, which experience frequent extreme high flows peaking in the spring and early summer

(February-May), followed by stream drying beginning in June and becoming most severe in September and October. Our analysis showed the mh19 (skewness in annual maximum flows) and mh13 (variability across maximum monthly flows), were among the most important metrics for fish in groundwater streams (Fig. 2). Lynch et al.'s (2019) study of flow alteration in Ozark Highlands groundwater flashy streams also found that variability across monthly maximum flows (mh13) were highly influential for fish assemblage relationships in flood years. Although timing of high flows had lower importance for fish assemblage turnover in runoff streams, timing of average flows, including the predictability of average flows (ta2), was considerably higher for fishes in runoff streams compared to groundwater and intermittent streams. Leasure et al. (2016) also observed a high ecological importance of ta2 in runoff flashy streams, while Kennen et al. (2009) found that ta2 was the most influential average flow metric in differentiating runoff streams from intermittent and groundwater streams in Missouri. The significance of the ta2 metric for fish assemblage turnover, as well as for characterizing runoff flow regimes, indicate predictability of average flows may be especially important for fish in streams experiencing higher variability in daily flows, but fewer zero flow days compared to intermittent streams.

Fish assemblage stability tends to be greater in streams experiencing lower environmental variability (Ross et al. 1985, Matthews and Marsh-Matthews 2017, Magoulick et al. 2021). The high importance of magnitude and rate of change of average flows in groundwater streams supports the idea that fish assemblages are typically less tolerant of large and rapid changes in discharge in more stable groundwater systems, compared to those in more variable runoff-dominated streams (Poff and Ward 1989, Magoulick et al. 2021). Stable streams with lower environmental variability also tend to have more equilibrium and specialized fish assemblages, which are typically less tolerant of hydrologic variability, while stream systems with higher flow variability tend to be dominated by more tolerant, generalist species (Poff et al. 1994). Common generalist species in our study area streams include Western Mosquitofish (*Gambusia affinis*), Red Shiner (*Cyprinella lutrensis*), Green Sunfish (*Lepomis cyanellus*) and Longear Sunfish (*Lepomis megalotis*), all

of which have advantageous reproductive traits like short interbrood intervals, enabling them to have several broods a year (Matthews and Marsh-Matthews 2017). These generalist species were rarely identified in our analysis as contributing significantly to overall patterns of assemblage compositional change, likely owing to their broad distribution in streams throughout the Ouachita-Ozark Interior Highlands. However, in groundwater streams the Julian date of annual maximum (th1) and arrival of maximum flows in January (as opposed to later in February-May) was identified as important for many species including *C. lutrensis*. Red shiners are known to continue to produce broods as late as November, with large numbers of extremely small (<14 mm total length) young-of-the-year found in Ozark streams during the coldest months of winter (Matthews and Marsh-Matthews 2017). Such early high flow events in Ozark groundwater streams may prove to be an important factor in sustaining high overwinter survival rates and rapid growth of young *C. lutrensis* observed in the early spring (Matthews and Marsh-Matthews 2017). In contrast, the timing of high flow events was of lower importance for fish assemblages in runoff streams than the duration of high flows. This indicates the possibility that fishes in these runoff streams may be less affected by future changes in the timing of peak discharge due to the comparative regularity of high flows from rapid precipitation runoff in these systems. Turnover of intermittent stream fish assemblages was strongly related to high flow duration (dh18) and seasonal predictability of non-flooding (th3) (Fig. 2). The duration and timing of high flow events may be especially important for intermittent stream fish, potentially as cues for initiating reproduction, migration and dispersal, as well as for juvenile recruitment, predation, and competition.

Decreased streamflow is associated with shifts in community composition in addition to declines in fish survival, growth, and abundance (Walters 2016). In our study streams, the frequency and timing of low flows had a significantly greater importance for fish assemblage turnover compared to the magnitude and duration of low flows (Fig. 2). Lynch et al. (2019) found that alteration of low flows had less of an influence on fish and macroinvertebrate assemblage relationships in groundwater flashy streams than alteration of high and average flows. Our results indicate that alteration of low flow magnitude may be especially important for intermittent stream biota, while changes in the frequency of low flow is of greater importance in groundwater and runoff-dominated systems. In a study of factors affecting fish assemblage structure in Ozark Highland streams during seasonal drying, Dekar and Magoulick (2007) also found that the magnitude of low flow events were critical in intermittent streams, with fish densities negatively related to pool refugia area and depth. The relative influence of these factors, however, varied among species and between years (Dekar and Magoulick 2007), highlighting the importance of local habitat attributes, species traits, and the nature of the low flow events themselves in driving fish assemblage change (Walters 2016).

In our analyses, darters (*Etheostoma* spp.) were frequently identified as contributing significantly to overall patterns of assemblage compositional change across all flow regimes, and have also previously been shown to be important for predicting hydrologic disturbance in Ozark and Ouachita Highlands and Gulf Coastal Plains streams (Fox and Magoulick 2019). Darters and other benthic aquatic species are particularly vulnerable to the effects of fragmentation and hydrologic alteration due to anthropogenic disturbance, particularly substrate instability and sedimentation of gravel substrate and riffle and benthic habitat (Tipton et al. 2004, Barbarossa et al. 2020). Unlike most fishes, darters in the genus *Etheostoma* lack a swimbladder, allowing them to maintain position in fast flowing water but also potentially limiting their dispersal abilities (Evans and Page, 2003). Consequently, as benthic, riffle-dwelling species adapted to fast flowing water, darters may be more sensitive to changes in flow velocity and low flows in particular, compared to species which are pool-dependent (Buchanan et al. 2017).

4.1. Limitations

Our empirical modeling approach used taxonomic records to quantify flow-ecology relationships and potential environmental thresholds in different stream flow regimes, providing a basis for future research on regional flow ecology relationships. We used species occurrence to control for differences in fish sampling methodologies, effort, and gear and to emphasize regional species patterns. Occurrence data are inherently less informative than abundance data, particularly as they relate to local changes in species distribution or density. On the other hand, abundance-based models may suffer from higher levels of uncertainty which cannot be alleviated by including additional samples (Waldock et al. 2022). Community turnover is just one measure of ecological response to hydrology and watershed disturbance, and in some cases, may be comparatively insensitive relative to other biological metrics such as population size or energy flux (Rosenfeld 2017). Regional, historical, and biotic factors may be as or more important than abiotic conditions in determining fish assemblage structure, along with location in the drainage basin and proximity to other streams (Angermeier and Winston 1998, Magoulick 2000). Similarly, our focus on hydrologic metrics may overlook the influence of other important abiotic factors (e.g. water temperature, dissolved oxygen, suspended sediment, organic chemical contaminants). Many of the hydrologic metrics are highly derivative, which renders their interpretability more difficult. Furthermore, substantial bias may be present in the calculation of hydrologic metrics, particularly those describing seasonality and low-flow events in intermittent streams having a high proportion of 0-flow days (Eng et al. 2017). While many of the hydrologic metrics we identified as important to fish turnover cannot be directly or easily regulated, these threshold relationships can nonetheless help managers establish how flow variability affects fish assemblages in streams with different flow regimes, instead of broadly applying a minimum instream flows approach. For example, in dam-regulated systems, a flow regime-based gradient forest threshold approach can assist with developing multi-species ecological hypotheses as a foundation for experimental testing of re-operation plans for restoring environmental flows (Richter and Thomas, 2007). In the case of unregulated streams, specific actions and watershed-scale conservation efforts can be developed to limit and enhance ecosystem resilience to anthropogenic- and climate-induced hydrologic alteration, while maintaining the characteristics of natural flow regimes in these different stream types to sustain diverse native fish populations and other biotic assemblages.

5. Conclusions

Identifying hydrologic and anthropogenic-source drivers of fish assemblage turnover across river networks can aid natural resource managers in identifying planning and management strategies that sustain natural ecosystem structure and function and will form the basis for examining flow-ecology relationships for macroinvertebrates and other taxonomic groups (e.g. mussels, crayfish, and algae) at state, regional and national scales. The frequent presence of nonlinear patterns in fish assemblage turnover identified by our analysis indicates that conservation and management stakeholders will need to carefully consider how assumptions of linearity or nonlinearity may affect future environmental flow needs. In addition to benefiting habitat and multi-species conservation and management actions through identification of local and regional flow-ecology relationships generalized across flow regimes, the results of this study provide a better understanding of complex nonlinear relationships that are critical to anticipating changes in aquatic ecosystems and communities, while highlighting the need for expanded coverage of stream monitoring gages in intermittent and headwater streams. Future research linking long-term stream temperature data with hydrology and using functional traits in addition to assemblage structure can be used to predict how aquatic communities respond to hydrologic and thermal variations across stream flow

regimes.

Data Availability Statement.

Hydrology and watershed data are freely available through the National Hydrographic Dataset (NHD; usgs.gov/national-hydrography/national-hydrography-dataset) and GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow Database (water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml). Biological data are available from the authors upon reasonable request.

CRediT authorship contribution statement

J. Tyler Fox: Conceptualization. **Daniel D. Magoulick:** .

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2022.109500.

References

Adamski, J. C. 1995. Environmental and hydrologic setting of the Ozark Plateaus study unit, Arkansas, Kansas, Missouri, and Oklahoma. National Water-Quality Assessment Program.

Allan, J.D., 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annu. Rev. Ecol. Evol. Syst. 35 (1), 257–284.

Allen, T.F., Starr, T.B., 1982. Hierarchy: perspectives for ecological complexity Chicago. University of Chicago Press Illinois.

Angermeier, P.L., Winston, M.R., 1998. Local vs. regional influences on local diversity in stream fish communities of Virginia. Ecology 79, 911–927.

Archfield, 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 Res. Appl. 30 (9), 1166–1183.

Barbarossa, V., Schmitt, R.J.P., Huijbregts, M.A.J., Zarfl, C., King, H., Schipper, A.M., 2020. Impacts of current and future large dams on the geographic range connectivity of freshwater fish worldwide. Proc. Natl. Acad. Sci. 117 (7), 3648–3655.

Bond, N.R., Kennard, M.J., 2017. Prediction of hydrologic characteristics for ungauged catchments to support hydroecological modeling. Water Resour. Res. 53 (11), 8781–8794.

Breiman, L, 2001. Random forests. Machine learning 45, 5–32.

Bruckerhoff, L.A., Leasure, D.R., Magoulick, D.D., Strecker, A., 2019. Flow–ecology relationships are spatially structured and differ among flow regimes. J. Appl. Ecol. 56 (2), 398–412.

Buchanan, B.P., Auerbach, D.A., McManamay, R.A., Taylor, J.M., Flecker, A.S., Archibald, J.A., Fuka, D.R., Walter, M.T., 2017. Environmental flows in the context of unconventional natural gas development in the M arcellus S hale. Ecol. Appl. 27, 37–55.

Bunn, S.E., Arthington, A.H., 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. Environ. Manage. 30 (4), 492–507.

Carlisle, D.M., Wolock, D.M., Meador, M.R., 2011. Alteration of streamflow magnitudes and potential ecological consequences: a multiregional assessment. Front. Ecol. Environ. 9 (5), 264–270.

Carlisle, D.M., Grantham, T.E., Eng, K., Wolock, D.M., 2017. Biological relevance of streamflow metrics: Regional and national perspectives. Freshwater Science 36 (4), 927–940.

Dallas Evans, J.D., Page, L.M., 2003. Distribution and relative size of the swim bladder in Percina, with comparisons to Etheostoma, Crystallaria, and Ammocrypta (Teleostei: Percidae). Environ. Biol. Fishes 66 (1), 61–65.

Davies, S.P., Jackson, S.K., 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecol. Appl. 16 (4), 1251–1266.

Dekar, M.P., Magoulick, D.D., 2007. Factors affecting fish assemblage structure during seasonal stream drying. Ecol. Freshw. Fish 16 (3), 335–342.

Dodds, W.K., Clements, W.H., Gido, K., Hilderbrand, R.H., King, R.S., 2010. Thresholds, breakpoints, and nonlinearity in freshwaters as related to management. J. North Am. Benthol. Soc. 29 (3), 988–997.

Döll, P., Fiedler, K., Zhang, J., 2009. Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. Hydrol. Earth Syst. Sci. 13, 2413–2432.

Ellis, N., Smith, S.J., Pitcher, C.R., 2012. Gradient forests: calculating importance gradients on physical predictors. Ecology 93 (1), 156–168.

Eng, K., Grantham, T.E., Carlisle, D.M., Wolock, D.M., 2017. Predictability and selection of hydrologic metrics in riverine ecohydrology. Freshwater Science 36 (4), 915–926.

Falcone, J.A., 2011. GAGES-II: Geospatial attributes of gages for evaluating streamflow. US Geological Survey.

Falcone, J.A., Carlisle, D.M., Wolock, D.M., Meador, M.R., 2010. GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States: Ecological archives E091-045. Ecology 91, 621.

Fox, J.T., Magoulick, D.D., 2019. Predicting hydrologic disturbance of streams using species occurrence data. Sci. Total Environ. 686, 254–263.

Freeman, M.C., Bowen, Z.H., Bovee, K.D., Irwin, E.R., 2001. Flow and habitat effects on juvenile fish abundance in natural and altered flow regimes. Ecol. Appl. 11 (1), 179–190.

Freeman, M.C., Bestgen, K.R., Carlisle, D., Frimpong, E.A., Franssen, N.R., Gido, K.B., Irwin, E., Kanno, Y., Luce, C., Kyle McKay, S., Mims, M.C., Olden, J.D., LeRoy Poff, N., Propst, D.L., Rack, L., Roy, A.H., Stowe, E.S., Walters, A., Wenger, S.J., 2022. Toward Improved Understanding of Streamflow Effects on Freshwater Fishes. Fisheries 47 (7), 290–298.

George, R., McManamay, R., Perry, D., Sabo, J., Ruddell, B.L., 2021. Indicators of hydro-ecological alteration for the rivers of the United States. Ecol. Ind. 120, 106908.

Gnann, S.J., Howden, N.J., Woods, R.A., 2020. Hydrological signatures describing the translation of climate seasonality into streamflow seasonality. Hydrol. Earth Syst. Sci. 24, 561–580.

Gray, J.S., Bjorgesaeter, A., Ugland, K.I., 2006. On plotting species abundance distributions. J. Animal 75 (3), 752–756.

Guo, Q., Olden, J.D., Higgs, D.M., 2014. Spatial scaling of non-native fish richness across the United States. PLoS ONE 9 (5), e97727.

Henriksen, J. A., J. Heasley, J. G. Kennen, and S. Nieswand. 2006. Users' manual for the Hydroecological Integrity Assessment Process software(including the New Jersey Assessment Tools). U. S. Geological Survey.

Junk, W.J., Bayley, P.B., Sparks, R.E., 1989. The flood pulse concept in river-floodplain systems. Canadian special publication of fisheries and aquatic sciences 106, 110–127.

Kennard, M.J., Olden, J.D., Arthington, A.H., Pusey, B.J., Poff, N.L., 2007. Multiscale effects of flow regime and habitat and their interaction on fish assemblage structure in eastern Australia. Can. J. Fish. Aquat. Sci. 64 (10), 1346–1359.

Kennen, J. G., J. A. Henriksen, J. Heasley, B. S. Cade, and J. W. Terrell. 2009. Application of the hydroecological integrity assessment process for Missouri streams.

King, R.S., Baker, M.E., 2010. Considerations for analyzing ecological community thresholds in response to anthropogenic environmental gradients. J. North Am. Benthol. Soc. 29 (3), 998–1008.

Leasure, D.R., Magoulick, D.D., Longing, S.D., 2016. Natural flow regimes of the Ozark-Ouachita interior highlands region. River Res. Appl. 32 (1), 18–35.

Lynch, D.T., Leasure, D.R., Magoulick, D.D., 2019. Flow alteration-ecology relationships in Ozark Highland streams: Consequences for fish, crayfish and macroinvertebrate assemblages. Science of The Total Environment.

Lytle, D.A., Poff, N.L., 2004. Adaptation to natural flow regimes. Trends Ecol. Evol. 19 (2), 94–100.

Magoulick, D.D., 2000. Spatial and temporal variation in fish assemblages of drying stream pools: the role of abiotic and biotic factors. Aquat. Ecol. 34, 29–41.

Magoulick, D.D., Dekar, M.P., Hodges, S.W., Scott, M.K., Rabalais, M.R., Bare, C.M., 2021. Hydrologic variation influences stream fish assemblage dynamics through flow regime and drought. Sci. Rep. 11, 1–15.

Marchetti, M.P., Moyle, P.B., 2001. Effects of flow regime on fish assemblages in a regulated California stream. Ecological Applications 11, 530–539.

Matthews, W.J., Marsh-Matthews, E., 2017. Stream fish community dynamics: A critical synthesis. JHU Press.

Matthews, W.J., Robison, H.W., 1998. Influence of drainage connectivity, drainage area and regional species richness on fishes of the interior highlands in Arkansas. Am. Midland Natural. 139 (1), 1–19.

Mims, M.C., Olden, J.D., 2013. Fish assemblages respond to altered flow regimes via ecological filtering of life history strategies. Freshw. Biol. 58, 50–62.

Mohammadi, B., 2021. A review on the applications of machine learning for runoff modeling. Sustain. Water Resour. Manage. 7, 1–11.

Mohammadi, B., Moazenzadeh, R., Christian, K., Duan, Z., 2021. Improving streamflow simulation by combining hydrological process-driven and artificial intelligence-based models. Environ. Sci. Pollut. Res. 28 (46), 65752–65768.

Nieto-Lugilde, D., Maguire, K.C., Blois, J.L., Williams, J.W., Fitzpatrick, M.C., 2018. Multiresponse algorithms for community-level modelling: Review of theory, applications, and comparison to species distribution models. Methods in Ecology and Evolution 9, 834–848.

Olden, J.D., Poff, N.L., 2003. Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. River Res. Appl. 19 (2), 101–121.

Palmer, M., and A. Ruhi. 2019. Linkages between flow regime, biota, and ecosystem processes: Implications for river restoration. Science 365:eaaw2087.

Pelletier, M.C., Ebersole, J., Mulvaney, K., Rashleigh, B., Gutierrez, M.N., Chintala, M., Kuhn, A., Molina, M., Bagley, M., Lane, C., 2020. Resilience of aquatic systems: Review and management implications. Aquat. Sci. 82, 1–25.

Poff, N. L., B. D. Richter, A. H. Arthington, S. E. Bunn, R. J. Naiman, E. Kendy, M. Acreman, C. Apse, B. P. Bledsoe, and M. C. Freeman. 2010. The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards. Freshwater Biol. **55**:147-170.

Poff, N.L., Olden, J.D., Pepin, D.M., Bledsoe, B.P., 2006. Placing global stream flow variability in geographic and geomorphic contexts. River Res. Appl. 22 (2), 149–166.

Poff, N.L., Ward, J.V., 1989. Implications of streamflow variability and predictability for lotic community structure: a regional analysis of streamflow patterns. Can. J. Fish. Aquat. Sci. 46 (10), 1805–1818.

Poff, N. L., J. D. Allan, and J. D. Yount. 1994. Streamflow variability, fish community structure, and implications for climatic change. ERL-D, Office of Research and Development, US Environmental Protection Agency.

Pitcher, R.C., Lawton, P., Ellis, N., Smith, S.J., Incze, L.S., Wei, C.L., Greenlaw, M.E., Wolff, N.H., Sameoto, J.A., Snelgrove, P.V., 2012. Exploring the role of environmental variables in shaping patterns of seabed biodiversity composition in regional-scale ecosystems. Journal of Applied Ecology 49, 670–679.

Poff, N.L., 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the north american Benthological society 16, 391–409.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., Sparks, R.E., Stromberg, J.C., 1997. The natural flow regime. Bioscience 47 (11), 769–784.

Poff, N.L., Zimmerman, J.K., 2010. Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. Freshw. Biol. 55, 194–205.

Power, M.E., Dietrich, W.E., Finlay, J.C., 1996. Dams and downstream aquatic biodiversity: potential food web consequences of hydrologic and geomorphic change. Environmental management 20, 887–895.

R Core Team, 2013. R: A language and environment for statistical computing. https://cran.microsoft.com/snapshot/2014-09-08/web/packages/dplR/vignettes/xdate-dplR.pdf.

Reynolds, L.V., Shafroth, P.B., Poff, N.L., 2015. Modeled intermittency risk for small streams in the Upper Colorado River Basin under climate change. J. Hydrol. 523, 768–780.

Richter, B. D., G. A. Thomas. 2007. Restoring environmental flows by modifying dam operations. Ecol. Soc. **12**.

Richter, B.D., Baumgartner, J.V., Powell, J., Braun, D.P., 1996. A method for assessing hydrologic alteration within ecosystems. Conserv. Biol. 10 (4), 1163–1174.

Rosenfeld, J.S., 2017. Developing flow–ecology relationships: Implications of nonlinear biological responses for water management. Freshw. Biol. 62 (8), 1305–1324.

Rosenfield, J.A., 2002. Pattern and process in the geographical ranges of freshwater fishes. Glob. Ecol. Biogeogr. 11, 323–332.

Ross, S.T., Matthews, W.J., Echelle, A.A., 1985. Persistence of stream fish assemblages: effects of environmental change. Am. Nat. 126 (1), 24–40.

Roy, A.H., Freeman, M.C., Freeman, B.J., Wenger, S.J., Ensign, W.E., Meyer, J.L., 2005. Investigating hydrologic alteration as a mechanism of fish assemblage shifts in urbanizing streams. J. North Am. Benthol. Soc. 24 (3), 656–678.

Scott, J.M., Davis, F., Csuti, B., Noss, R., Butterfield, B., Groves, C., Anderson, H., Caicco, S., D'Erchia, F., Edwards Jr., T.C., 1993. Gap analysis: a geographic approach to protection of biological diversity. Wildlife monographs 3–41.

Sowa, S.P., Annis, G., Morey, M.E., Diamond, D., 2007. A gap analysis and comprehensive conservation strategy for riverine ecosystems of Missouri. Ecological monographs 77, 301–334.

Tipton, J.A., Bart, H.L., Piller, K.R., 2004. Geomorphic disturbance and its impact on darter (Teleostomi: Percidae) distribution and abundance in the Pearl River drainage, Mississippi. Hydrobiologia 527 (1), 49–61.

Waldock, C., Stuart-Smith, R.D., Albouy, C., Cheung, W.W.L., Edgar, G.J., Mouillot, D., Tjiputra, J., Pellissier, L., 2022. A quantitative review of abundance-based species distribution models. Ecography 2022 (1).

Walker, R.H., Walters, A.W., 2019. A mechanistic understanding of ecological responses to land-use change in headwater streams. Ecosphere 10, e02907.

Walters, A.W., 2016. The importance of context dependence for understanding the effects of low-flow events on fish. Freshwater Sci. 35 (1), 216–228.

Warner, R.M., 1998. Spectral analysis of time-series data. Guilford Press.

Yeung, A.C., Musetta-Lambert,, J.L., Kreutzweiser, D.P., Sibley, P.K., Richardson, J.S., 2018. Relations of interannual differences in stream litter breakdown with discharge: bioassessment implications. Ecosphere 9.

Zbinden, Z.D., 2020. Temporal dynamics of stream fish assemblages and the role of spatial scale in quantifying change. Ecology and evolution 10, 952–961.

Zeileis, A., Wiel, M.A., Hornik, K., Hothorn, T., 2008. Implementing a class of permutation tests: the coin package. J. Stat. Softw. 28, 1–23.

Ziegeweid, J.R., Johnson, G.D., Krall, A.L., Fitzpatrick, K., Levin, S.B., 2022. Quantifying relations between altered hydrology and fish community responses for streams in Minnesota. Ecol. Processes 11, 1–25.