JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

AMERICAN WATER RESOURCES ASSOCIATION

elfgen: A New Instream Flow Framework for Rapid Generation and Optimization of Flow–Ecology Relations

Joseph Kleiner, Elaina Passero, Robert Burgholzer, Jennifer Rapp, and Durelle Scott

Research Impact Statement: New framework developed for modeling the upper limits of biodiversity as a function of streamflow using principles of the river continuum concept and widely available hydrologic and ecological data.

ABSTRACT: Effective water resource management requires practical, data-driven determination of instream flow needs. Newly developed, high-resolution flow models and aquatic species databases provide enormous opportunity, but the volume of data can prove challenging to manage without automated tools. The objective of this study was to develop a framework of analytical methods and best practices to reduce costs of entry into flow–ecology analysis by integrating widely available hydrologic and ecological datasets. Ecological limit functions (ELFs) describing the relation between maximum species richness and stream size characteristics (streamflow or drainage area) were developed. Species richness is expected to increase with streamflow through a watershed up to a point where it either plateaus or transitions to a decreasing trend in larger streams. Our results show that identifying the location of this "breakpoint" is critical for producing optimal ELF model fit. We found that richness breakpoints can be estimated using automated low-supervision methods, with high-supervision providing negligible improvement in detection accuracy. Model fit (and predictive capability) was found to be superior in smaller hydrologic units. The ELF model ("elfgen" R package available on GitHub: https://github.com/HARPgroup/elfgen) can be used to generate ELFs using built-in datasets for the conterminous United States, or applied anywhere else streamflow and biodiversity data inputs are available.

(KEYWORDS: flow–ecology; instream flow; ecological limit function (ELF); river continuum concept (RCC); water supply management; ecological limits of hydrologic alteration (ELOHA); species–discharge relationship (SDR); percent of flow (POF).)

INTRODUCTION

Flow–Ecology in Water Supply Management

The challenge of managing water resources to meet human water supply demands, preserve ecological integrity, and protect designated use has spurred ecological flow research over the past 60 years. Commonly implemented techniques of the Ecological Limits of Hydrologic Alteration (ELOHA) (Arthington et al. 2010) framework include classification of flow regimes (Henriksen et al. 2006; McManamay et al. 2012; Novak et al. 2016), hydrologic models of pre and postalteration, and

---

Paper No. JAWRA-19-0132-P of the *Journal of the American Water Resources Association* (JAWRA). Received September 4, 2019; accepted July 28, 2020. **Discussions are open until six months from issue publication**.

Office of Water Supply (Kleiner, Burgholzer), Virginia Department of Environmental Quality Richmond, Virginia, USA; Department of Civil and Environmental Engineering (Passero), Colorado State University Fort Collins, Colorado, USA; U.S. Geological Survey (Rapp), Virginia and West Virginia Water Science Center Richmond, Virginia, USA; and Department of Biological Systems Engineering (Scott), Virginia Tech Blacksburg, Virginia, USA (Correspondence to Kleiner: joseph.kleiner@deq.virginia.gov).

Citation: Kleiner, J., E. Passero, R. Burgholzer, J. Rapp, and D. Scott. 2020. "elfgen: A New Instream Flow Framework for Rapid Generation and Optimization of Flow-Ecology Relations." *Journal of the American Water Resources Association* 1–18. https://doi.org/10.1111/1752-1688.12876.

correlation analyses of aquatic ecological response data at paired flow sites to develop flow alteration–response relations (Poff et al. 2010). ELOHA studies typically require pre and postalteration biological, streamflow, and/or land-use and flow control structure information. This often results in a limited number of sample sites that can be analyzed due to data availability. Space-for-time substitution, that is, comparing flow–ecology data at most and least altered sites, is often used as a surrogate for pre and postalteration biological responses. However, due to variation in aquatic communities, it can be difficult to identify comparable sites for a space-for-time comparison. Ungaged streams also present a challenge due to the resources and effort required to create effective hydrologic models.

Historically, many studies have catalogued evidence that support the use of stream size based characteristics (drainage area [DA], streamflow, or stream order) as a primary organizing factor in determining richness or density in aquatic communities (Schlosser 1987; Beecher et al. 1988; Angermeier and Schlosser 1989; Walters et al. 2003; Filipe et al. 2010; Pracheil et al. 2013; Vander Vorste et al. 2017). The term species–discharge relationship (SDR) is commonly used to describe the increase in ecological diversity with streamflow and is often represented as a continuous curve with some suggesting a piecewise model (Xu et al. 2016). Additional previous research has found that fish and benthic macroinvertebrate communities exhibit negative responses to decreases in streamflow magnitude (Dewson et al. 2007; Döll and Zhang 2010; Poff and Zimmerman 2010; Armstrong et al. 2011; Carlisle et al. 2012; Rolls et al. 2012; Gido et al. 2013; Kennen et al. 2014; Knight et al. 2014; Brooks and Haeusler 2016; Rapp and Reilly 2017). In light of these most basic stream size–ecology relations, studies have begun to advocate the use of SDRs to predict changes in fish species richness resulting from decreases in streamflow (Jowett et al. 2005; Xenopoulos et al. 2005; Xenopoulos and Lodge 2006; Döll and Zhang 2010; Iwasaki et al. 2012).

Using mean flows as a single indicator may appear to be overly simplistic given the flow-variability concepts that underpin contemporary instream flow science (Poff et al. 1997). However, due to the cost, time, and analytical power required to produce flow–ecology relations with traditional ELOHA models, some ecological flows practitioners have begun to advocate for "percent of flow" (POF) management schemes, which limit flow alterations to a certain percent of instantaneous flow (Richter et al. 2012). POF management schemes are a particularly good match for mean flow-based SDRs, because they often distribute flow alterations more equally across the hydrograph, thus preserving the natural variability in the flow regime while reducing its magnitude. While the use of SDR to predict declines in biodiversity due to consumptive water use shows promise, investigators have noted issues of scale dependence, and watershed fish-assemblage specificity as barriers to generalizability of SDR curves. Additionally, when extrapolating from SDR assessments, previous investigators have cautioned against assuming a continuous relation of increasing richness from the smallest to the largest streams (McGarvey and Ward 2008).

The River Continuum Concept

River continuum concept (RCC) hypothesizes that species richness throughout a watershed is expected to increase with stream size characteristics up to a point where it either plateaus or transitions to a decreasing trend in larger streams (Vannote et al. 1980). We refer to the location of this shift in the flow–ecology relation as the "RCC breakpoint." To date, few studies have considered the impact of an RCC breakpoint when attempting to characterize relations between streamflow and species richness. In this study, we introduce the use of RCC principles to help bound areas in a watershed where a continuous relation between increasing flow and species richness is valid. Central questions regarding the RCC are whether the RCC breakpoint occurs in a majority of stream systems (and at what range of streamflows it occurs) and if it ultimately represents a means of classifying streams that will produce improved SDR model fit.

Ecological Limit Function

Ecological limit functions (ELFs) have been posited as a way to characterize the observed ecological response of an organism or group of organisms to a specific limiting factor (Arthington et al. 2010; Poff and Zimmerman 2010; Knight et al. 2014; Rosenfeld 2017). ELF models build on concepts outlined in SDR approaches but are unique in their focus on ecological maxima as it relates to one or more stream characteristics. We developed an ELF model (Figure 1) using R (R Core Team 2018). A quantile and natural log-linear regression-based model is used to define the ELF. The approach we developed begins by plotting richness with a stream size based characteristic such as streamflow, and identifying the RCC breakpoint. The RCC breakpoint is used to isolate the data that exhibit an increasing trend in maximum richness with streamflow. A user-supplied quantile value is then applied, and quantile regression is used to characterize the upper portion of this data subset (Koenker and Bassett 1978). Finally, the ELF model consists of a natural log-linear regression through the upper subset of data. The ELF represents the functional relation between streamflow and the maximum number of taxa that can be supported given prevailing biotic and abiotic conditions within a hydrologic unit.

The ELF model is described by the following equation:

$$y = m \times \ln(x) + b, \tag{1}$$

where m and b are, respectively, the slope and intercept of the ELF, x refers to the stream size metric used, and y is the maximum richness value for a given stream size metric. Regression statistics derived from the ELF equation include R^2, adjusted-R^2, p-value, and number of data points n in the upper subset identified using quantile regression (Figure 1). These statistics are used to assess model performance and ELF goodness of fit.

ELF Generation Framework (elfgen)

Our objective was to develop a framework of analytical methods and best practices to facilitate extensive flow–ecology analyses. Our methodology is novel in its synthesis of the existing concepts of ELOHA, SDR and ELF with the integration of RCC as an organizing factor to optimize traditional flow–ecology modeling approaches. The use of widely available hydrologic and ecological datasets is at the core of our approach. We developed automated routines for RCC breakpoint detection, and analytical techniques to quantify the effects of using an RCC breakpoint in flow–ecology analyses. We also examined the application of ELF models across spatial scales: from large 6-digit hydrologic unit code (HUC 6) to small (HUC 10) units as well as Level III ecoregions. Best practices for applying the elfgen framework are discussed, and potential use for informing POF management schemes by assessing ecological response to consumptive withdrawals is considered. The primary outcome of this study is a decision support tool for ELF generation and analysis in the form of the elfgen R package for describing the relation between maximum species richness and stream size characteristic (available on GitHub: https://github.com/HARPgroup/elfgen). Our findings and the elfgen tool we have developed provide a consistent framework for ELF analysis and can be used in a wide range of applications across the field of water resources management. We illustrate its utility and provide guidance for its usage through an illustrative application which is discussed next.

APPLICATION METHODOLOGY

Study Area

The greater Virginia study area was chosen to facilitate the development of the ELF model (Figure 2). Extensive availability of both fish and benthic macroinvertebrate richness data make this an ideal

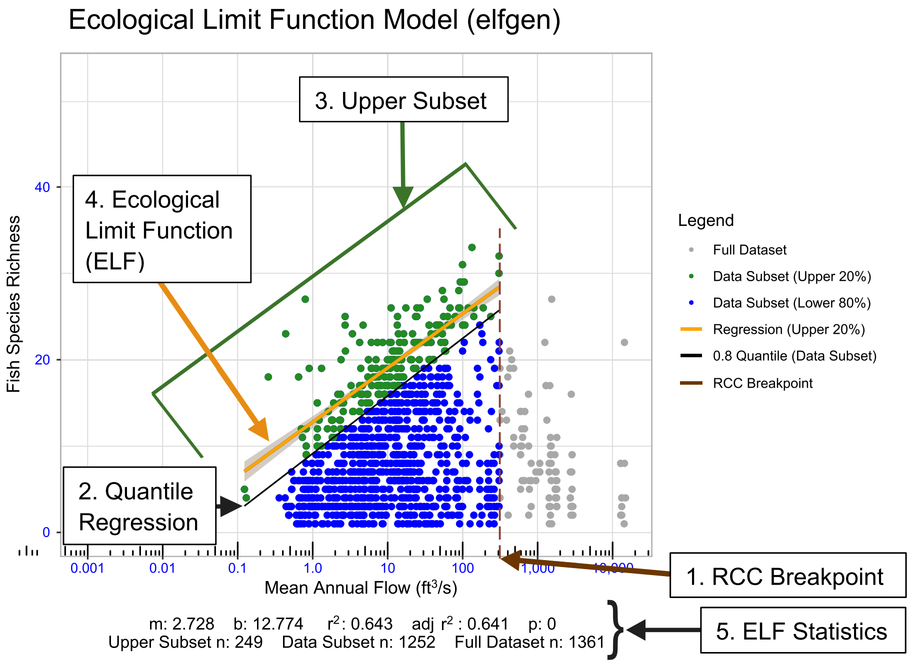

FIGURE 1. Diagram of "elfgen" Ecological Limit Function (ELF) model — Example relation between Mean Annual Flow (MAF) and fish species richness for the Potomac River 6-digit hydrologic unit code (HUC 6) (020700). ELF model derived using the following methodology: (1) river continuum concept (RCC) breakpoint identified and used to isolate dataset of interest; (2) Quantile regression used to characterize the (3) upper subset of data; (4) ELF generated using natural log-linear regression of upper subset; (5) Regression statistics derived from ELF equation.

study area for testing the application of the ELF framework. The study area extended to all of Virginia's major hydrologic boundaries: Mid-Atlantic (HUC 02), South Atlantic-Gulf (HUC 03), Ohio (HUC 05), and Tennessee (HUC 06). Included were data from surrounding states (Tennessee, West Virginia, Maryland, Kentucky, and North Carolina) in order to characterize these major drainage basins shared with Virginia. In lieu of flow regime classifications typical of ELOHA, we sought to characterize flow–ecology relations within the study area for the entire Virginia dataset, and contained nested hydrologic units (Angermeier and Winston 1999), and Level III ecoregion units (Omernik and Griffith 2014).

Ecological and Hydrologic Datasets

Ecological data for this study were sourced from the VAHydro-EDAS dataset (Tetra Tech Inc 2012). This dataset was chosen due to its wealth of both fish and benthic macroinvertebrate richness data throughout the Virginia study area. A second data source, the IchthyMaps database (Frimpong et al. 2016), contains ecological data for the entire conterminous United States (U.S.) and is used as the default data source in the elfgen R package. All hydrologic data (including Mean Annual Flow [MAF], 12 Mean Monthly Flows [MMF], and DA) used in this study and the accompanying elfgen R package are sourced from the National Hydrography Dataset Plus Version 2 (NHDPlus V2) and Enhanced Unit Runoff Method (EROM) datasets (McKay et al. 2012). For our methodology, each stream segment in the NHDPlus V2 dataset is assigned modeled flow data from the EROM dataset. See Appendix A-D for additional notes on datasets used.

ELF Model Assumptions

The methods in this study assume that the upper subset of data (defined here as the samples higher

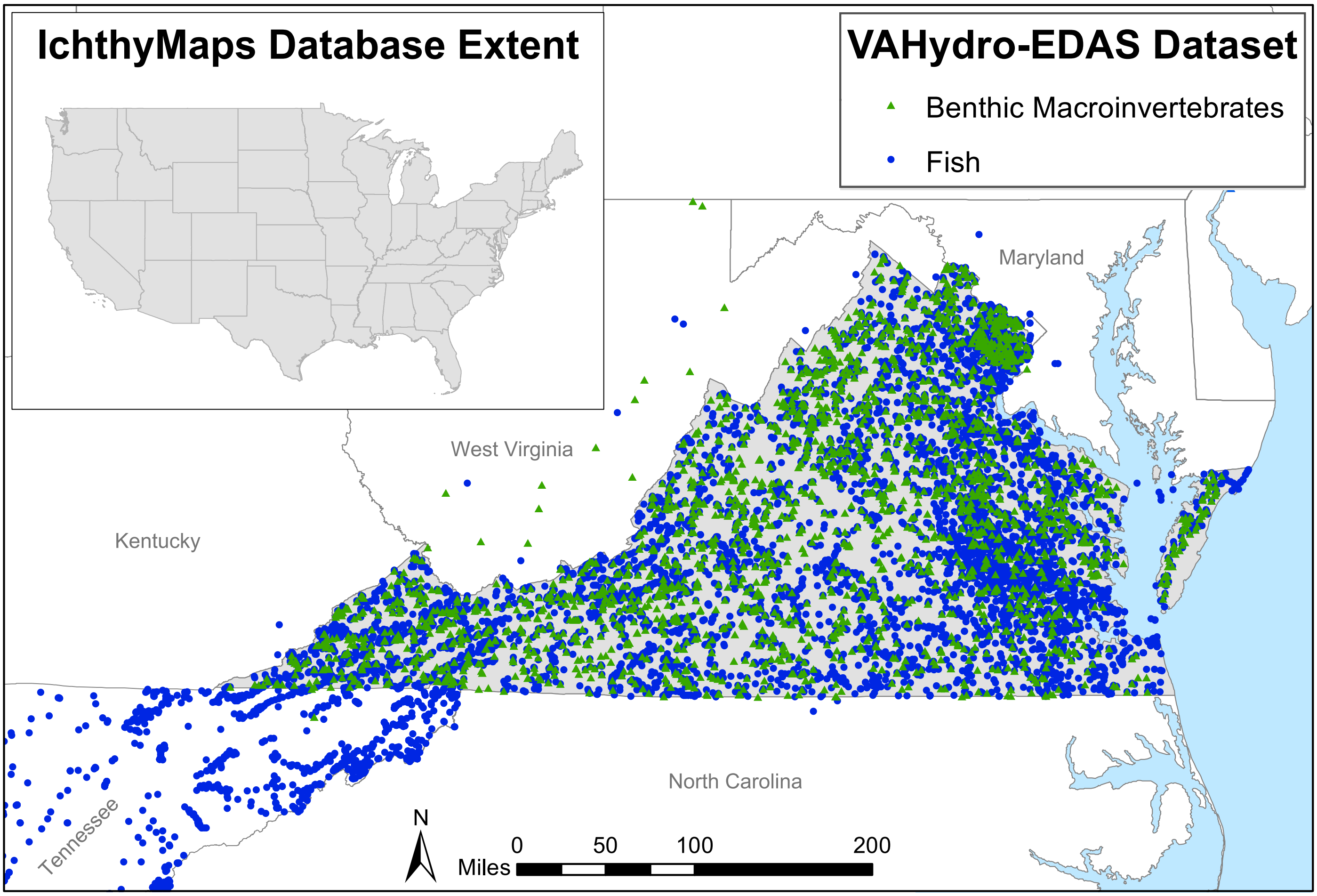

FIGURE 2. Map of ecological monitoring stations — IchthyMaps database covering fish samples for the entire conterminous United States, and VAHydro-EDAS dataset covering fish and benthic macroinvertebrate sample data for the greater Virginia study area. EDAS, Ecological Data Application System.

than the 80th percentile) is representative of the maximum species richness possible across all stream sizes. The ELF model we developed utilizes quantile regression for characterizing this upper subset, as it is often used in ecological models for its ability to estimate the effects of limiting factors such as streamflow (Cade et al. 1999; Schmidt et al. 2012; Knight et al. 2014). It should be noted that the elfgen R package accepts user inputs of any quantile of interest, and the 80th percentile value was chosen to facilitate batch generation of ELFs. The 80th percentile generally provided a sufficient number of data points in the upper subset for performing regression analyses, while still reliably tracing the upper edge of the wedge-shaped data distribution within our study area.

Identification of RCC Breakpoint

Relative diversity patterns described by the RCC (Vannote et al. 1980) indicate a peak in species richness occurring in intermediate streams. This phenomenon has been widely reported in ecological literature for both fish and macroinvertebrates (Minshall et al. 1985; Ward 1998; Paller et al. 2006) and implies there is a threshold in stream size after which richness can be expected to plateau or decrease with increasing streamflow. The RCC breakpoint, can be thought of as dividing a river into two domains, left and right of the breakpoint, with the left portion exhibiting richness increases with increasing streamflow.

Initial data exploration examined richness peaks and breakpoint patterns in plots of richness with MAF and DA. To provide a baseline, all available fish samples for Virginia were plotted in a single scatterplot. Initial inspection revealed a wedge-shaped distribution of richness consistent with the RCC paradigm and other ecological limit studies (Terrell et al. 1996; Vaz et al. 2008; Schmidt et al. 2012; Zorn et al. 2012; Knight et al. 2014). At the HUC 6 scale, ELFs were then constructed both including and excluding samples from larger streams (MAF > 200 ft^3/s, 5.66 m^3/s) for comparison. By excluding samples from larger streams, ELFs more closely traced the visual upper limits of species richness than those that included data points from streams of all MAFs. We proceeded to develop a set of methods for identifying RCC breakpoints in order to construct ELFs using samples from smaller streams only. In larger streams, beyond the RCC breakpoint, it can be assumed that the ELF model is no longer optimal for characterizing the flow–ecology relationship.

Due to the size of the datasets to be analyzed, we sought rapid manual and robust automated methods of RCC breakpoint identification. We explored three candidate techniques: the “Fixed,” “Ymax,” and “Piecewise Iterative” methods.

Fixed Method. The Fixed Method utilizes a user specified breakpoint value. This "fixed breakpoint" is determined through visual inspection of the flow–ecology relation by the user. An example plot of an ELF produced with the Fixed Method using a fixed value of 200 ft^3/s (5.66 m^3/s) is shown in Figure 3b.

Ymax Method. The Ymax Method treats the maximum observed species richness value as the RCC breakpoint. This approach begins by locating the point with the highest y-value in the full dataset, and then utilizing the associated *x*-value as the breakpoint (Figure 3c). The theory behind this method centers on the RCC, which assumes a peak in richness occurring in intermediate streams. Identifying the point with the single greatest richness value provides a general approximation of the RCC breakpoint.

Piecewise Iterative Method. The Piecewise Iterative Method (PWIT) (Lemoine 2012) uses an iterative algorithm to identify shifts in the relation between maximum richness and stream size metric. This method requires a user to identify an initial bounding range between two x-values (“bound low” = *blo*, “bound high” = *bhi*), that is, believed to contain a breakpoint. For our analyses, visual inspection of the full dataset was used to select the bounding range. The PWIT function uses an upper subset of data to search for potential breakpoints. The *x*-values of all data points in the upper subset between *blo* and *bhi* are set as potential breakpoint locations. Natural log-linear regression equations are then developed on either side of each potential breakpoint. The iteration that produces the piecewise regression with the lowest root mean squared error is identified as the breakpoint. An example plot of an ELF produced with the PWIT Method is shown in Figure 3d. This method, if applied over a high-resolution dataset could be prohibitively computationally intensive. However, the RCC breakpoint of a larger HUC should apply to those nested HUCs contained within it. Therefore, in this study we used the PWIT Method to determine a breakpoint in a HUC 6, then used that location to inform a fixed breakpoint location for contained HUC 8 and HUC 10 units in order to minimize reliance on individual user knowledge of bounding ranges for smaller hydrologic units.

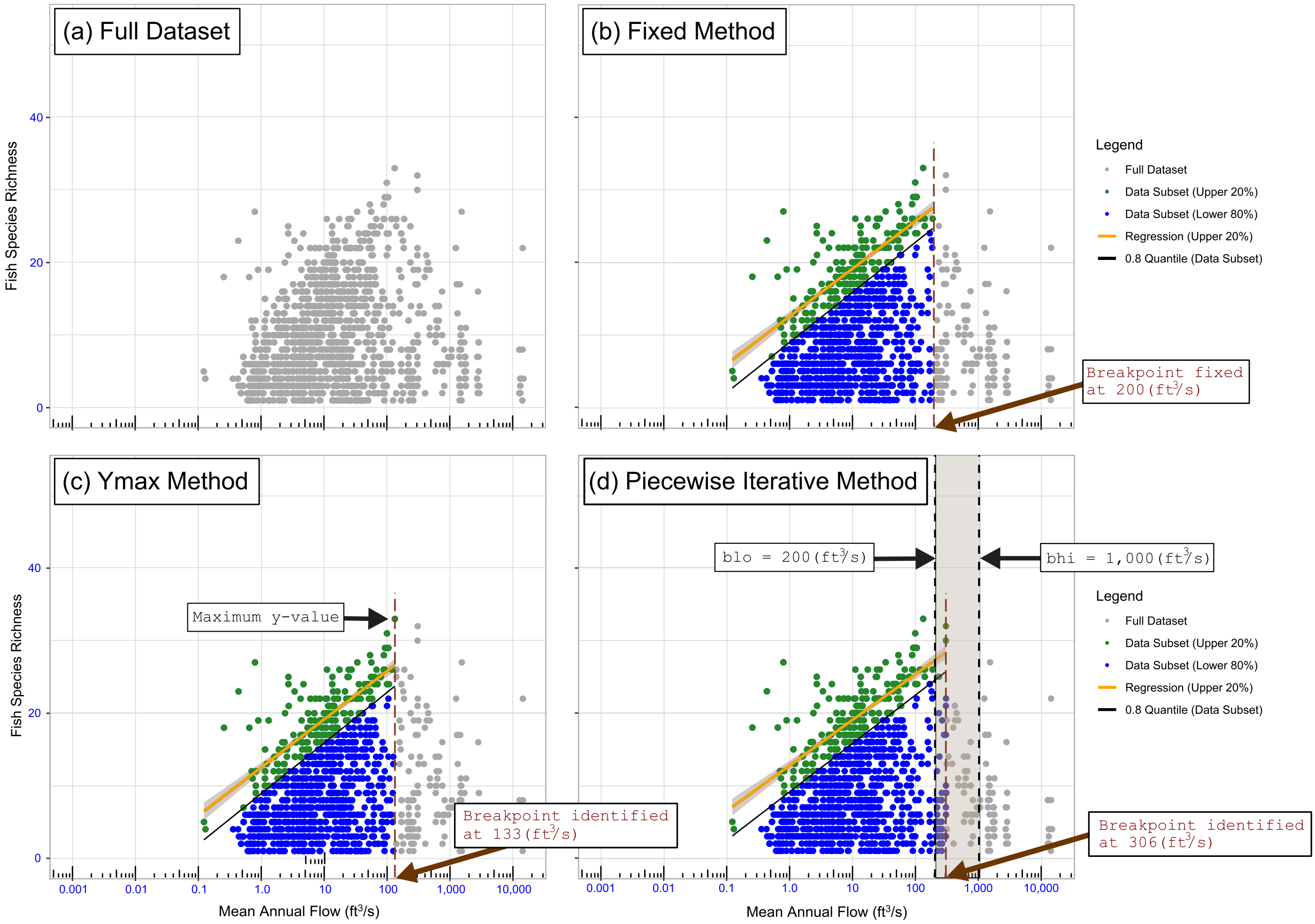

FIGURE 3. RCC breakpoint identification methods. (a) Example relation between MAF and fish species richness for the Potomac River HUC 6 (020700). Three techniques developed for determining RCC breakpoint locations include the (b) Fixed Method (user-specified, single breakpoint value), (c) Ymax Method (breakpoint identified corresponding to maximum y-value in each geospatial unit), and (d) Piecewise Iterative (PWIT) Method (breakpoint identified for each geospatial unit between user-specified range of *x*-values). These automated methods vary in the level of user supervision required.

Application of elfgen

Evaluating the elfgen framework required batch generation of ELFs for fish and benthic macroinvertebrate richness with DA, MAF and MMF at varying hydrologic units. To evaluate the utility of the three RCC breakpoint identification methods, batches were run of all the HUC 6, HUC 8 and HUC 10 units within the study area, and resulting slope and goodness of fit compared. A comparison of performance between the methods can be found in the Results and Discussion section that follows.

In the interest of producing an objective, minimally supervised method of ELF characterization, the elfgen framework was applied using specific analytical approaches distinguished by their levels of user supervision. ELFs resulting from the varying supervision levels were compared on the quantity of statistically significant regressions and goodness-of-fit criteria. Model runs that did not employ a breakpoint are considered unsupervised (See "None" in Table 1) and were included as a baseline for comparison. The initial method of identification by visual inspection using the Fixed Method (breakpoint set at 530 ft^3/s, 15.0 m^3/s) is referred to as low supervision, due to the use of a single breakpoint value applied to all hydrologic units and the minimal user input required. Ymax Method runs are considered medium supervision due to this method's ability to automate breakpoint detection, with some user supervision needed in cases of high richness outliers at small streamflows. The PWIT Method runs are considered high supervision due to the high level of user input required for selecting custom *blo* and *bhi* inputs for every HUC 6 in the study area.

Flow–ecology relations are watershed and location specific, and it is critical to be able to characterize richness-flow trends at varying hydrologic unit scales. To explore the elfgen framework for nested hydrologic units, we generated ELFs for HUC 8 and HUC 10 units contained within each HUC 6 investigated. Figure 4 helps to highlight ELF model trends at varying hydrologic unit scale (regression statistics presented in Table 2). By producing an ELF for the entire state of Virginia (Figure 4a), we can get a reference point to the overall flow–ecology trend observed using all the monitoring data in Virginia, without considering any watershed-specific organizing factors. Ecoregions were also explored as grouping units for instances where sufficient ecological data are lacking at the hydrologic unit level. ELFs produced using Level III ecoregions and HUC 6 units were compared due to their similar relative size within the study area.

RESULTS AND DISCUSSION

The Breakpoint Factor

The example in Figure 5 highlights the improved ELF model performance when considering an RCC breakpoint, and the potential consequences of failing to implement a breakpoint in flow–ecology analyses. As presented in the figure, the strength of the model regression is significantly improved when utilizing a breakpoint (Figure 5b). Not only does the ELF model more closely trace the visual upper limits of biodiversity, but the ELF statistics also emphasize the improved goodness of fit. Adjusted-R^2 for this example improves from 0.079 when not using a breakpoint to 0.673 when considering a breakpoint. Table 1 also highlights improvements in ELF fit that result from employing an RCC breakpoint, with the median Adjusted-R^2 of 0.49 without a breakpoint, increasing to 0.6 when employing a breakpoint. In addition to improved model fit, the number of hydrologic units with a statistically significant ELF increases by 10%–17% depending on the breakpoint method used. Therefore, factoring in RCC breakpoints prior to running flow–ecology models allows for more accurate interpretations of hydrologic unit ecological trends.

TABLE 1. Comparison of adjusted-R^2 values and count of resulting statistically significant ELFs ($p < 0.01$, $m > 0$, adjusted-$R^2 > 0$) for fish species richness, HUC 8 units, organized by supervision level. The "Overall" column represents the median adjusted-R^2 of ELFs generated with all explanatory variables; drainage area (DA), MAF, and 12 Mean Monthly Flows (MMF).

	Median adjusted-R^2			
Level of supervision	**Overall**	**DA**	**MAF**	**HUC count**
None (no breakpoint)	0.49	0.50	0.48	30
Low (Fixed Method)	0.60	0.65	0.61	35
Medium (Ymax Method)	0.60	0.65	0.62	33
High (PWIT Method)	0.60	0.66	0.57	34

The Breakpoint Factor — Comparison of Identification Methods

All three breakpoint identification methods: PWIT, Ymax, and Fixed proved capable of providing RCC breakpoints that greatly improved model fit over ELFs generated without a breakpoint. The PWIT Method performed best at the HUC 6 resolution (due to HUC 6 units in Virginia that generally encompass entire watersheds), and successfully discerned geographic variation in breakpoint location. The Ymax Method performed well overall, but had a unique limitation: when HUC datasets contained high richness outliers at small streamflows, it selected RCC breakpoints that appeared inaccurate based on visual inspection. The Fixed Method, when implemented with a user-defined breakpoint of 530 ft^3/s (15.0 m^3/s) MAF (identified by inspection from a plot of richness vs. DA for all data in the study), was simplest to use overall and performed well in a wide variety of situations. It was hypothesized that the PWIT and Ymax methods would produce a high degree of local breakpoint specificity, and this proved to be true. However, local breakpoint specificity did not always result in better model fit, with the Fixed Method providing results comparable to PWIT on the whole.

Compared to RCC breakpoints from the PWIT Method, the 530 ft^3/s (15.0 m^3/s) value appeared to over-estimate the breakpoint in eastern Virginia HUCs, and underestimate in western HUCs (these evaluations were based on visual inspection of RCC breakpoint location). The eastern side of the study area included the New River and Atlantic drainage basins and the western side included tributaries of the Tennessee drainage basin, a known biodiversity hotspot (Jenkins and Burkhead 1994; Ostby et al. 2014; Zipper et al. 2014). While this did not appear to affect the results of this study substantially, we suggest that a single fixed breakpoint value should be used with caution in areas of high physiographic variation. It should be noted that in certain situations data were organized in a way (such as a single vertical line, or only for a limited range of streamflow) that no breakpoint could be identified with the PWIT Method.

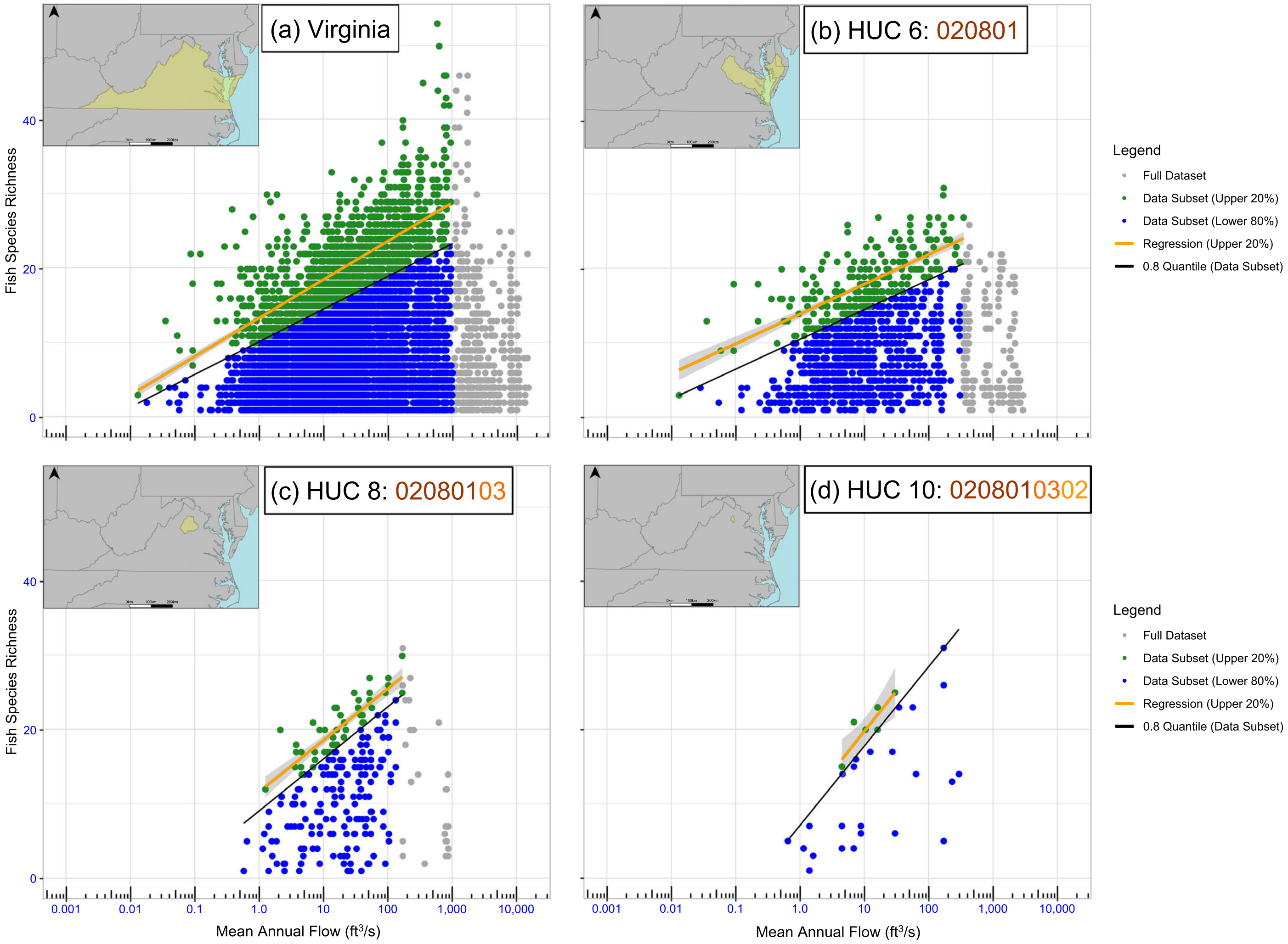

FIGURE 4. ELF models at varying HUC scale. (a) ELF produced using all data points in Virginia (included for reference). (b) ELF at HUC 6 resolution — Lower Chesapeake. (c) ELF for a contained HUC 8 — Rapidan-Upper Rappahannock. (d) ELF for a contained HUC 10 — Carter Run-Rappahannock River. As HUC resolution becomes finer, the ability to characterize location-specific flow–ecology trends improves. All ELFs were constructed using fish species richness and MAF.

TABLE 2. ELF model regression statistics at varying hydrologic unit scale.

		ELF statistics			
Region	**Area (km^2)**	***m***	***b***	**adjusted-R^2**	***p***
Virginia	110,786	2.3	13.4	0.6	0.0000
HUC 6 — Lower Chesapeake	22,347	1.7	14	0.5	0.0000
HUC 8 — Rapidan-Upper Rappahannock	4,032	3.0	11.7	0.7	0.0000
HUC 10 — Carter Run-Rappahannock River	343	4.7	8.9	0.8	0.0055

The Breakpoint Factor — Impact of User Supervision

High and medium supervision selection of RCC breakpoints produced better-fitting ELFs for HUC 6 units than the low supervision method. All methods delivered comparable results for HUC 8 units (Table 1). The number of HUC 8 ELFs produced was roughly equivalent across the three supervision methods for ELFs generated with each of the individual 14 explanatory variables (DA, MAF, and 12 MMF) (Figure 6). Overall, little monthly variation was observed at the watershed unit scale evaluated. All methods performed similarly in the fall months, which produced some poorest-fitting ELFs (Figure 6). The highest goodness of fit and greatest number of fish richness ELFs produced were for DA and May

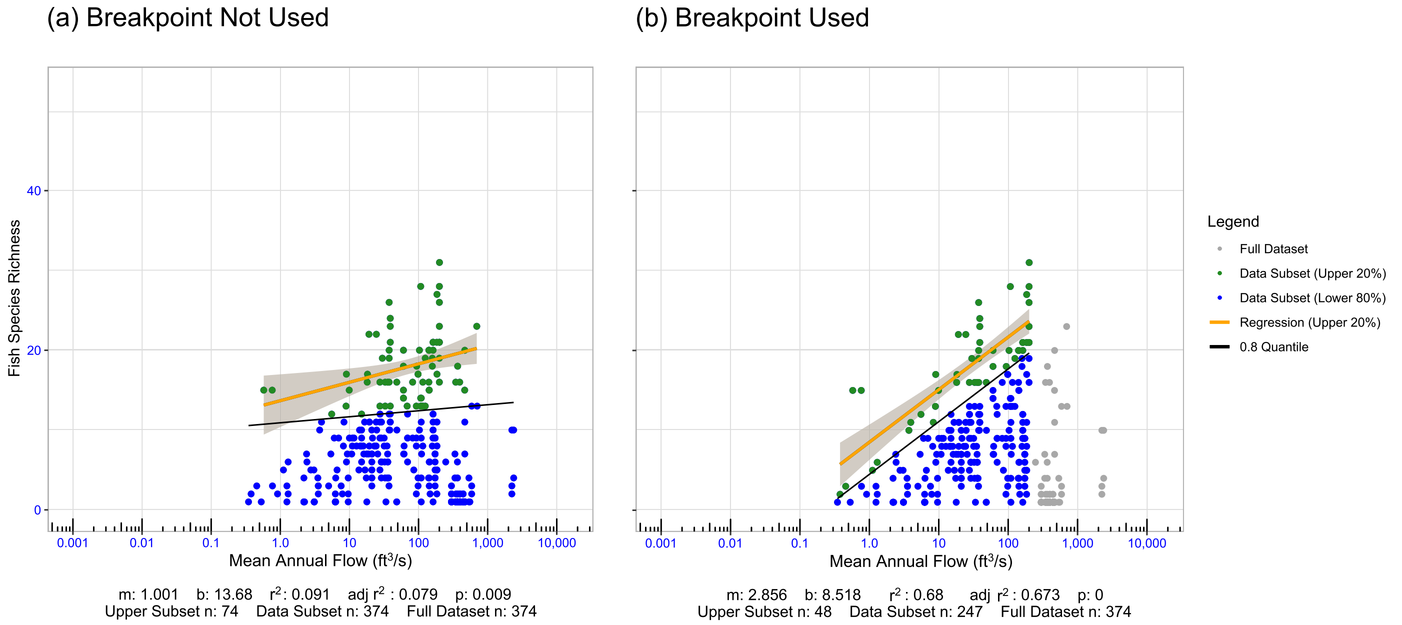

FIGURE 5. Example highlighting improved ELF model performance when using a breakpoint — Relation between MAF and fish species richness for the Upper Dan HUC 8 (03010103) shown. (a) flow–ecology trend using entire dataset, (b) RCC breakpoint identified and used to isolate data exhibiting an increasing trend in richness with streamflow. The ELF model in (b) more closely traces the visual upper limits of biodiversity and has improved goodness of fit compared with the ELF produced using the entire dataset (a). When neglecting to consider a breakpoint, there is potential to falsely characterize the flow–ecology relation in a HUC.

MMF. None of the methods showed a significantly higher median adjusted-R^2 (Table 1) or tighter interquartile range of adjusted-R^2 (Figure 6) across the majority of the explanatory variables. However, with increased supervision, the minimum adjusted-R^2 values generally increased (Figure 6). This increase suggests watershed-specific breakpoints may reduce the number of poor fitting ELFs produced with lower supervision methods using generalized breakpoints; however, this increase was not statistically significant. At spatial scales finer than HUC 8, we did not find an improvement in the ELF fit (using HUC 10) to justify applying HUC 8 breakpoints to contained HUC 10 units.

The Resolution Factor — ELFs in Nested Hydrologic Units

It is possible to generate a single ELF for an area the size of Virginia (Figure 4a), but it is critical for water supply management to highlight the variation in flow–ecology relations within such a large spatial unit. For example, a subset of these data using only those data points contained within the Lower Chesapeake HUC 6 (Figure 4b) produces a substantially different distribution than that of the entire Virginia dataset, a flow–ecology relation that is specific to the Lower Chesapeake. Though not as large and diverse an area as the entire extent of Virginia, the heterogeneous nature of large hydrologic units, such as a HUC 6, results in an ELF that incorporates data from several stream systems, each with their own unique streamflow and ecology characteristics. To highlight the effect of hydrologic unit resolution, additional nested hydrologic unit trends are presented in Figure 4 at the HUC 8 and HUC 10 levels. It is important to note that neighboring HUC 10 units contained within the same HUC 8 may exhibit significantly different flow–ecology trends from each other. The ELF-modeled trends presented in Figure 4 vary not only visually, but statistically as well. As highlighted in Table 2, slope m and y-intercept b vary depending on HUC resolution. Perhaps more significant are the increases in model performance observed at finer resolution HUCs. The adjusted-R^2 for this set of example HUCs increases as we move from HUC 6 to HUC 8 to HUC 10. Characterizing the high degree of variability in flow–ecology relations across hydrologic unit scales is critical for developing an understanding of localized impacts of individual water supply management actions.

The Resolution Factor — Batch Analysis of HUCs and Ecoregions

The improvement in model performance achieved when using higher resolution HUCs is not isolated to the example presented in Figure 4, and is something

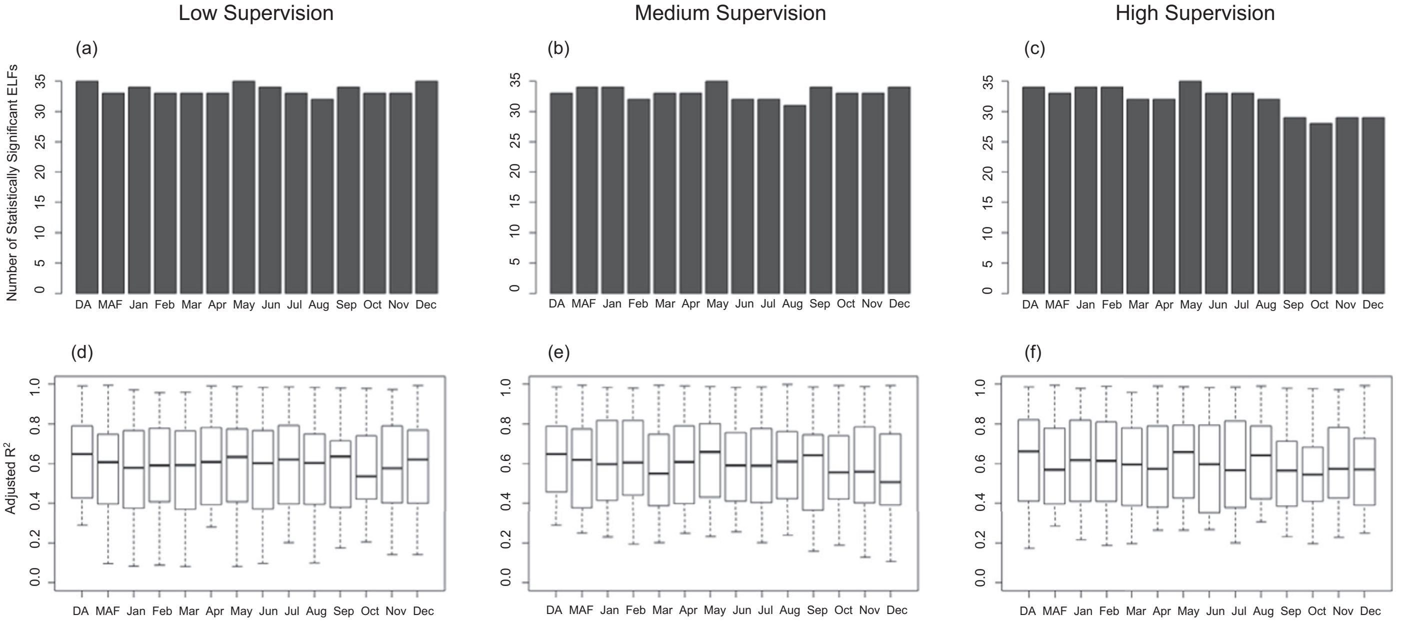

FIGURE 6. Number of statistically significant HUC 8 DA, MAF, and MMF ELFs using (a) low, (b) medium, and (c) high supervision methods of RCC breakpoint selection for fish species richness. Goodness of fit of HUC 8 DA, MAF, and MMF ELFs using (d) low, (e) medium, and (f) high supervision methods of RCC breakpoint selection for fish species richness. Plots include only those ELFs with adjusted-$R^2 > 0$, $m > 0$, and $p < 0.01$.

we have been able to characterize for HUCs across the study area. ELF goodness of fit improved as HUC size decreased for both MAF and DA (Figure 7a, 7b). Moving from HUC 6 down to HUC 10, the range of slopes also widened (Figure 7c, 7d). HUC 8 and HUC 10 units had more varied slopes and adjusted-R^2 values than HUC 6 units because each HUC 6 represents a summary or integration of four to eight nested HUC 8 units, and as many as 100 HUC 10 units. Higher goodness of fit with a wider range of slopes indicates the emergence of distinct relationships at smaller spatial scales. Using smaller spatial scales may homogenize other geomorphic factors (Angermeier and Winston 1999; Walters et al. 2003; Filipe et al. 2010) that influence diversity patterns within major drainage basins. It should be noted that for HUC 10 units, very high goodness of fit values were often produced from small sample sizes. These smaller sample sizes, as seen in Figure 4d, may result in wider confidence intervals, potentially impacting ELF predictive ability. However, due to the phenomenon of ELF variability across spatial scale and improved model performance at finer resolution HUCs, we believe whenever sufficient data are present, modeling ELFs at a HUC 10 resolution or finer can help ensure the accuracy of flow–ecology interpretations.

Only four statistically significant Level III ecoregion ELFs were produced for fish richness with a fixed breakpoint of 530 ft^3/s (15.0 m^3/s) MAF. Both Level III ecoregion and HUC 6 had similar median adjusted-R^2 values, but Level III ecoregion ELFs had slightly worse fit overall (Figure 7a, 7b). Level IV ecoregions were also investigated, but did not contain sufficient data to produce ELFs for the Virginia study area. Based on this initial analysis, ecoregion ELFs did not outperform HUC-based classification approaches, and moreover, are not ideal for applying our ELF framework, which assumes the area units being investigated are hydrologically defined. While we recognize their role in shaping aquatic community structure, and assert that it is possible to generate ELFs with Level III ecoregions, further exploration into their utility was not pursued.

ELFs for Benthic Macroinvertebrates

The ELF modeling framework was assessed for its utility with benthic macroinvertebrate data. ELFs were constructed for benthic macroinvertebrates using high and medium supervision methods, with the Fixed Method proving to be most effective. Sample taxonomic resolution greatly affected distribution of the data, thus altering the location of observed RCC breakpoints. Figure 8 shows how the sample taxonomic resolution varies, limiting the ability to develop ELFs. In the example presented the maximum richness identified using majority family-genus level data was 32 taxa; however, for the same

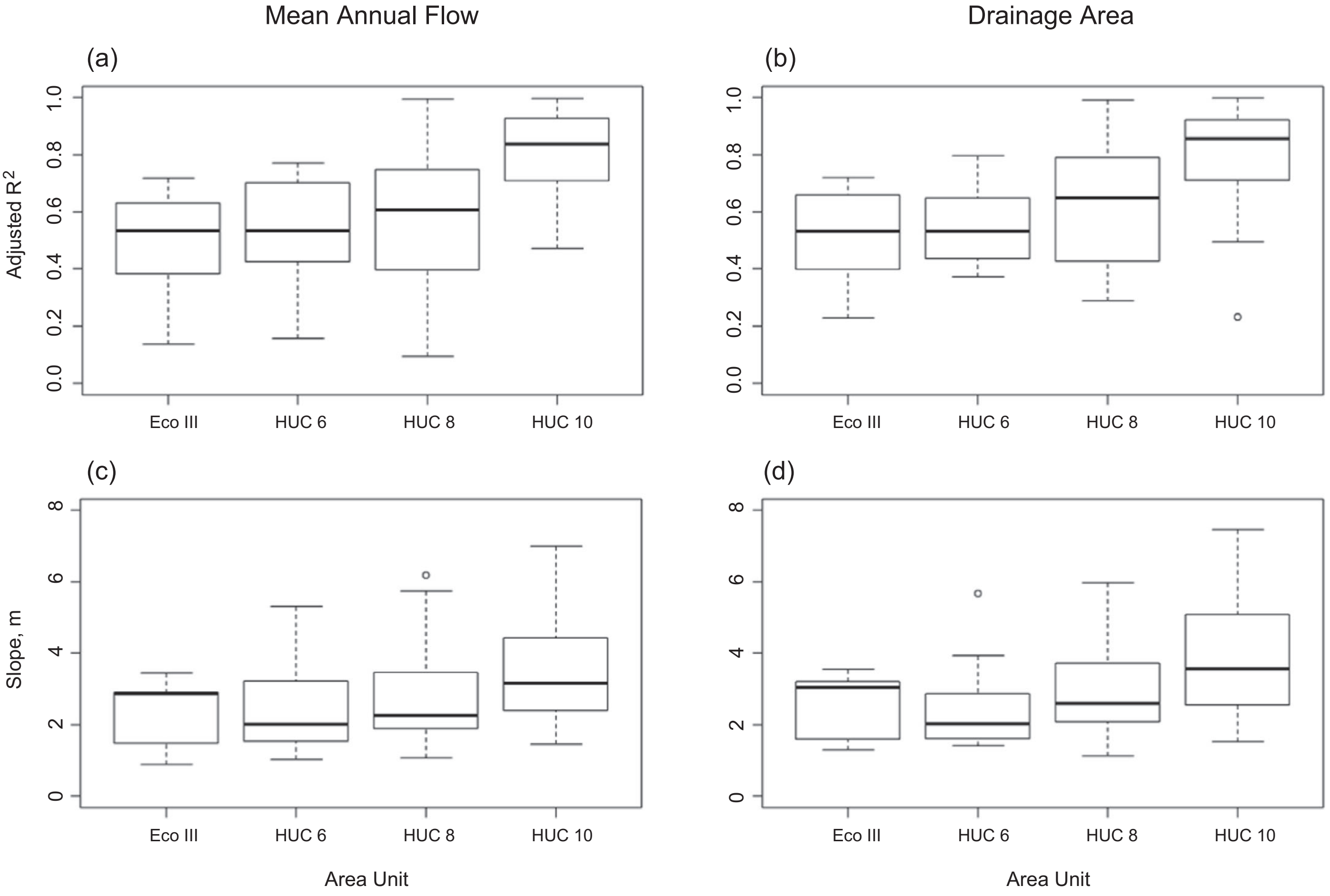

FIGURE 7. ELF model performance by area unit. Comparison of goodness of fit (adjusted-R^2) for (a) MAF and (b) DA across Level III ecoregion, HUC 6, HUC 8 and HUC 10 units. Variation in statistically significant ($p < 0.01$, adjusted-$R^2 > 0$, $m > 0$) ELF slopes for (c) MAF and (d) DA. All ELFs were constructed using fish species richness and a fixed MAF breakpoint of 530 ft^3/s.

hydrologic unit using the majority family-genus-species level data resulted in a maxima of 75 taxa (Figure 8b; Appendix D). Construction of benthic macroinvertebrate ELFs was restricted by species level sample availability, and limited sample data in large streams. However, we believe our ELF framework has utility for characterizing benthic macroinvertebrate flow–ecology relations where sufficient data are available.

SUMMARY AND CONCLUSIONS

Sustainable water resource management relies on a practical understanding of instream flow needs. This paper presents an ELF framework we developed to provide a data-driven approach to flow–ecology analyses. Our focus centers on the characterization of richness data and their relation with stream size characteristics by applying principles of the RCC. The ELF approach utilizes widely available hydrologic and ecological datasets. The analytical tools and best practices we have outlined can be used to characterize the wedge-shaped distribution of increasing richness as a function of stream size characteristics for hydrologic units.

Using the Virginia study area as a case study, we were able to characterize ELFs at the HUC 6, HUC 8, and HUC 10 scales, as well as several ecoregion units. We determined the identification and use of an RCC breakpoint as an organizer of flow–ecology data to be a critical component in flow–ecology analyses. Identification of the RCC breakpoint reduces the risk of mischaracterizing ELF slope by excluding data from larger streams that exhibit shifts in the flow–ecology relation. We found that with only minor user supervision, the identification of an RCC breakpoint improved ELF model performance and resulted in an increased number of statistically significant relations. We also highlight the importance of considering HUC

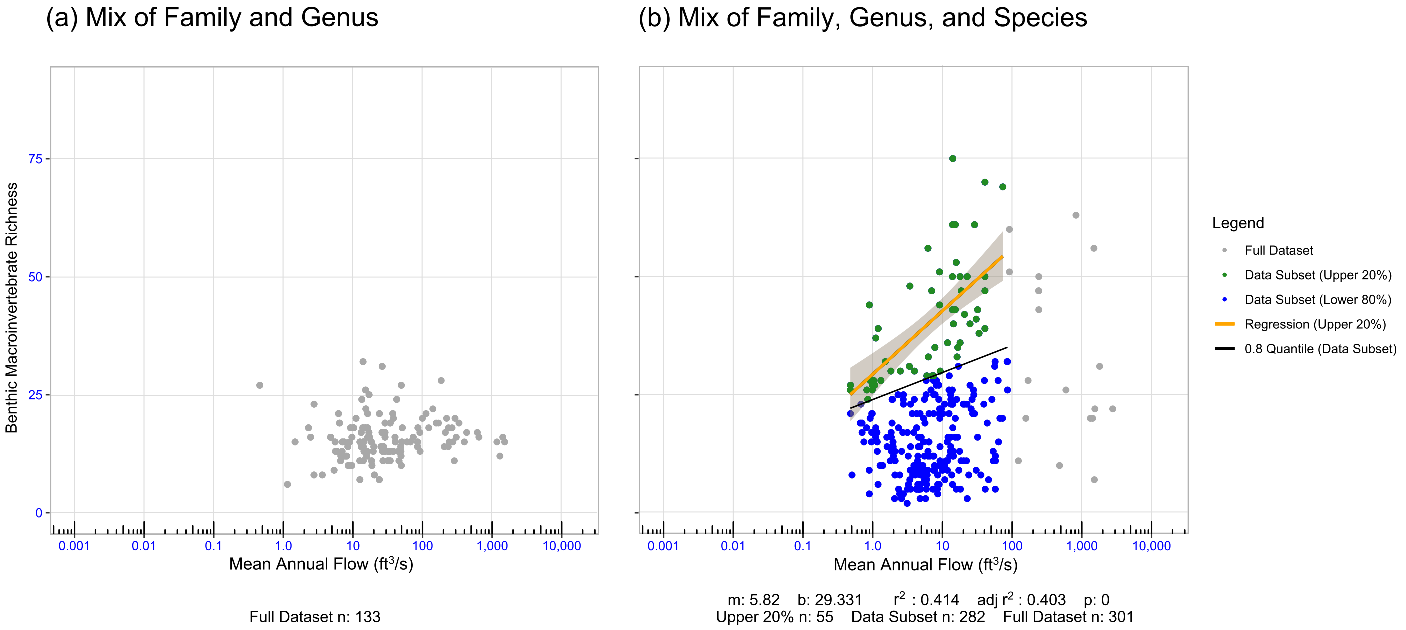

FIGURE 8. Example highlighting the effect of benthic macroinvertebrate sample taxonomic resolution on maximum observed richness values, and the resulting ability to produce ELF models — Relations with MAF for the Potomac River HUC 6 (020700) shown. (a) Data containing a mix of family and genus level samples (b) Data containing a mix of family, genus, and species level samples. The dataset containing only family and genus level data has a flatter distribution with a peak value of 32, compared to the dataset containing species level data which has a peak value of 75 and a discernible flow–ecology relation that can be characterized using the ELF modeling approach.

scale, and whenever sufficient data are present, modeling ELFs at a HUC 10 resolution can improve the goodness of fit of resulting ELF curves. The range of ELF slopes increased at the HUC 10 scale, suggesting that a diverse array of subareas may be nested into larger hydrologic units, with potentially diverse sensitivities to decreases in streamflow. This differing sensitivity, if verified in practice, could form a powerful tool for watershed managers to optimize the location of flow decreases to meet growing off-stream demands.

Hydrologic data for our analyses consisted primarily of MAF, though we explored producing ELFs with MMF data as well. With sufficient monthly flow datasets, it should be feasible to characterize ELF relations at the monthly time scale. Ecological data at the species level were shown to be critical for ELF generation. Although our analysis of benthic macroinvertebrate data was limited, we show that a mix of family-genus-species level data can be used to facilitate ELF generation where species level data are lacking. When considering family level benthic macroinvertebrate data only, it was difficult to discern flow–ecology relations with streamflow. These results are similar to a recent study that worked with mixed agency collections (Cuffney and Kennen 2017). While our study relied heavily on the VAHydro-EDAS database, statistically significant ELFs can be generated with the elfgen R package using the IchthyMaps dataset, which contains fish data for the entire conterminous U.S.

Previous investigators have suggested that models describing richness as a function of streamflow can be used to estimate relative or absolute risk to biodiversity resulting from streamflow decreases. This study has shown that flow–ecology relations can be readily characterized using automated methods and commonly available datasets. The elfgen model we developed describing the relation between maximum species richness and stream size characteristic is available on GitHub: https://github.com/HARPgroup/elfgen. The slope of the elfgen equation or “rate of change” may be analogous to ecological sensitivity to consumptive water use. This work suggests that Equation 1 can be used to assess potential richness loss with flow reduction for individual hydrologic units. By developing a framework that makes use of existing datasets, without the need for additional costly sampling, we were able to quickly and efficiently characterize flow–ecology relations that can be used to inform decision making and spur additional scientific investigations across the field of water resources management.

APPENDIX A

ECOLOGICAL DATA PROCESSING

Ecological data were sourced from two datasets: VAHydro-EDAS (Tetra Tech Inc 2012) and IchthyMaps (Frimpong et al. 2016). The VAHydro-EDAS dataset, based originally in Microsoft Access, was built by Tetra Tech, Inc., and made available through the U.S. Environmental Protection Agency and The Virginia Department of Environmental Quality (DEQ) in 2012. The dataset includes raw sample data counts and summary ecological metrics from eight agency and university sources in Virginia, Tennessee, West Virginia, Maryland, North Carolina, and Kentucky. Fish data spanned 1972–2010 and benthic macroinvertebrate data spanned 1992–2010. The elfgen R package utilizes IchthyMaps as the default data source, however, all analyses in this study were conducted using the VAHydro-EDAS dataset.

Richness metrics from the VAHydro-EDAS were generated using Microsoft Access SQL queries. While 44 richness metrics (spanning fish, benthic macroinvertebrate, and periphyton datasets) were available in EDAS, only fish and benthic macroinvertebrate richness were analyzed in this study. These metrics are a sum of the number of unique taxa (species, genus, or family) that are observed at a single location in a single sample event. These metrics were chosen because they are the most general indicators of species biodiversity, address considerations in selecting useful ecological indicators of flow-alteration under the ELOHA framework (Poff et al. 2010), and had an abundance of records. Fish sample identification was at the species level with over 12,000 samples spanning over 9,000 unique sampling locations. The taxonomic level for benthic macroinvertebrate samples varied by agency, project, and year. In an effort to maintain the highest resolution of data possible, benthic macroinvertebrate data were grouped based on sample taxonomic resolution. Three benthic macroinvertebrate groups were developed (Table A1), consisting of: (1) majority family-level samples (around 5,000 samples); (2) majority family-genus-level samples (around 1,700 samples); and (3) majority family-genus-species level samples (around 1,600 samples). More information about the sample resolution groups can be found in Appendix D.

Merging Ecological and Hydrologic Data

Hydrologic data were acquired through the NHDPlus V2 dataset. EDAS stations with geographic coordinates were spatially joined with NHDPlus V2 catchments in ArcGIS 10.3 to obtain an NHDPlus V2 COMID for each station. If multiple stations fell inside the same catchment, all of those stations were assigned the same COMID. The COMID linked stations with the corresponding NHDPlus V2 hydrologic attributes.

NHDPlus V2 modeled streamflow data were sourced from the EROM dataset (McKay et al. 2012). EROM represents the mean annual and monthly streamflows over the 1971–2000 period (McKay et al. 2012). A total cumulative DA value was also assigned to each station based on the value at the NHDPlus V2 segment outlet. Because

TABLE A1. Sample resolution groups based on taxonomic level. Samples were grouped by project and date range to improve resolution uniformity [Project or Agency names from EDAS database (Tetra Tech Inc 2012): INSTAR, Virginia Commonwealth University INteractive STream Assessment Resource; MAHA, USEPA Medical Advocates for Healthy Air; MAIA, USEPA Mid Atlantic Integrated Assessment; NAWQA, U.S. Geological Survey (USGS) National Water Quality Assessment Program data from the USGS BioData Retrieval system database (BioData); PM-Ben-F-Oth, VA Department of Environmental Quality (DEQ) — Probabilistic Monitoring Program — Benthos, Family IDs, non-Probabilistic Sites; PM-Ben-F-Prb, VA DEQ — Probabilistic Monitoring Program — Benthos, Family IDs, Probabilistic Sites; PM-Ben-G-Oth, VA DEQ — Probabilistic Monitoring Program — Benthos, Genus IDs, non-Probabilistic Sites; PM-Ben-G-Prb, VA DEQ — Probabilistic Monitoring Program — Benthos, Genus IDs, Probabilistic Sites].

Sample resolution group	Fish or benthic macroinvertebrate	Project or agency	Date range
Species	Fish	All	All
Majority family, genus, species	Benthic macroinvertebrate	INSTAR	<2009
		MAHA	All
		MAIA	All
		NAWQA	All
Majority family, genus	Benthic macroinvertebrate	PM-Ben-F-Oth	<2009
		PM-Ben-F-Prb	>2007
		PM-Ben-G-Oth	>2008
		PM-Ben-G-Prb	>2007
		INSTAR	=2009
Majority family	Benthic macroinvertebrate	PM-Ben-F-Oth	<2008
		PM-Ben-F-Prb	<2009
		PM-Ben-G-Oth	<2008

streamflow and DA values represent the outlet of NHDPlus V2 segments (McKay et al. 2012), stations on the upstream end of a flowline or on a stream beyond the NHDPlus V2 resolution may have been assigned high values relative to actual site conditions. However, because of the small segment size of the NHDPlus V2 dataset, it was assumed that the impact of these errors would be negligible. Overall, around 11,000 fish stations and just over 2,500 benthic macroinvertebrate stations across Virginia, West Virginia, Maryland, North Carolina, Tennessee, and Kentucky had both DA and flow data available.

Data quality was assessed by plotting richness with streamflow or DA, revealing two primary issues: (1) gaps in data during winter months (January, February, and March) in very small streams (<1 ft^3/s, 0.02 m^3/s), and (2) stations with unrealistically high richness values for very low flow values (richness values an order of magnitude greater than these small streams can realistically support). Comparison of area-normalized EROM streamflow and U.S. Geological Survey (USGS) gage historic streamflow records showed that gaps in data were likely the result of overestimates of winter streamflow (Appendix B). These overestimates were observed in a minority of hydrologic units and were likely related to reference gage regression flow adjustments during the EROM dataset compilation (personal correspondence with Timothy Bondelid, consulting engineer for USGS NHDPlus V2 project, April 2018). Winter streamflow estimates were retained in the dataset to support further analyses and were considered a limitation of the EROM dataset. NHDPlus V2 divergent flowpaths were likely the cause of stations with unrealistically high richness values when plotted against MAF, but NOT when plotted with DA. These flows were assumed to be unrepresentative of the true conditions at the stations and were excluded from subsequent analysis (see Appendix C).

APPENDIX B

QA OF EROM STREAMFLOW DATA

Initial plots of species richness with MMFs revealed gaps in the data during winter months

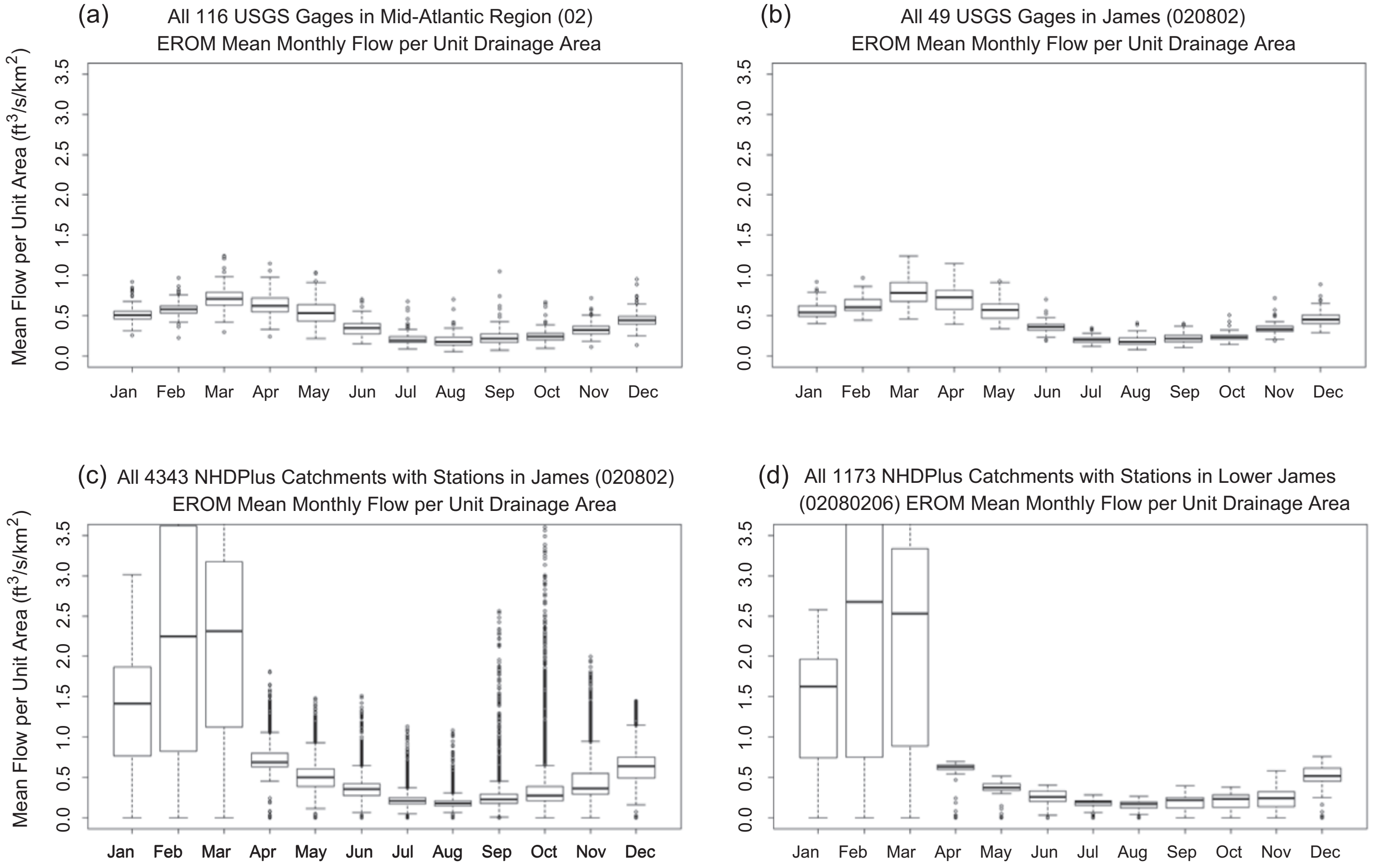

FIGURE B1. Boxplots of historic DA-normalized MMF ($\text{ft}^3/\text{s}/\text{km}^2$) for all USGS gages in (a) the Mid-Atlantic Region (02) and (b) the James (020802). Boxplots highlighting exaggerated DA-normalized winter Enhanced Unit Runoff Method (EROM) MMF ($\text{ft}^3/\text{s}/\text{km}^2$) for NHDPlus V2 catchments containing EDAS stations in (c) the James (020802) and (d) the Lower James (02080206).

(January, February, and March) in very small streams (<1 ft^3/s, 0.02 m^3/s). EROM streamflow data for all EDAS stations were compared to historic USGS streamgage records (USGS, National Water Information System. Accessed August 2016, https://waterdata.usgs.gov/nwis) to assess the limitations of the EROM model. MMFs (ft^3/s) were normalized based on the DA in km^2 of their NHDPlus V2 catchment or their USGS gage.

EROM streamflow data for a HUC 8 (Lower James) and historic USGS streamgage records for the HUC 6 (James River) and HUC 2 (Mid-Atlantic) containing this HUC 8 were examined. Figure B1a, B1b shows monthly box plots of the flow record normalized by DA for all the gages contained within the James River and Mid-Atlantic Region. Figure B1c, B1d shows similar monthly box plots using the EROM modeled streamflow for the Lower James and James River HUCs. The EROM data closely match the historic flow data, except for the months of January, February, and March where EROM appears to overestimate mean streamflow values.

Since the EROM model appeared to exaggerate mean streamflow in the winter, boxplots of MMFs normalized by DA were produced at the HUC 6 level. Three of the six HUC 6 units examined showed exaggerated EROM estimates in January, February, and March. Five of those HUCs, including the James River, also showed a high number of outliers in October. The other normalized MMFs and MAFs for all the HUCs examined appeared to be in line with those from the USGS gages. The exaggerated MMFs and outliers were accepted as a limitation of using EROM data.

APPENDIX C

DIVERGENT FLOWPATHS IN NHDPLUS V2 NETWORK

High taxa count outliers appeared in plots of MAF but not DA for the same HUC. We examined the locations of stations with high taxa counts and low MAFs (<1 ft^3/s, 0.02 m^3/s) and found that these stations fell in catchments containing divergent flowpaths. The NHDPlus V2 handles divergent flowpaths by assigning 100°1 of cumulative flow to the main path and 0°1 of cumulative flow to minor paths (McKay et al. 2012). Therefore, if a station was closest to a minor flowpath, it ended up with a very low gage-adjusted streamflow value. DA values for these stations were unaffected because those values came from the stations' catchment outlet. DA:MAF ratios were examined to identify stations that may have been located on divergent flowpaths. A DA:MAF ratio greater than 1,000 was considered unrealistic, and these stations were excluded from subsequent analysis.

APPENDIX D

BENTHIC MACROINVERTEBRATE DATA PROCESSING

Ambiguous taxonomic identifications appeared in both benthic macroinvertebrate and fish data. These identifications were likely the result of damaged or immature specimens that could not be identified at the lowest taxonomic unit of the rest of their sample (Cuffney et al. 2007). Tetra Tech, Inc., dealt with inconsistent taxonomic levels among data sources and ambiguous taxa identifications through the use of family and genus level Operational Taxonomic Units (Tetra Tech Inc 2012). However, differences in taxonomic levels between data sources persisted. Figure D1 shows the distribution of taxa rank by project or collecting agency and sampling date.

Benthic macroinvertebrate data were collected and identified at a variety of taxonomic resolutions: family, genus and species, with multiple taxonomic resolutions present in a single sample. Preliminary analysis showed that relations between streamflow and benthic macroinvertebrate richness were difficult to discern when a majority of data was at the genus and family level; however, statistically significant trends could be observed using data that included abundant species level benthic macroinvertebrate data (Figure 8 shows example plots of two sample resolution groups for the same HUC). Therefore, the benthic macroinvertebrate data were analyzed in terms of distributions of taxonomic resolution, and three benthic macroinvertebrate sample resolution groups were categorized based on these distributions (see Table A1). Each individual sample was assigned to only one sample resolution group. Statistically significant ELFs were generated using only data from the group that included species level identification, the "majority family, genus,

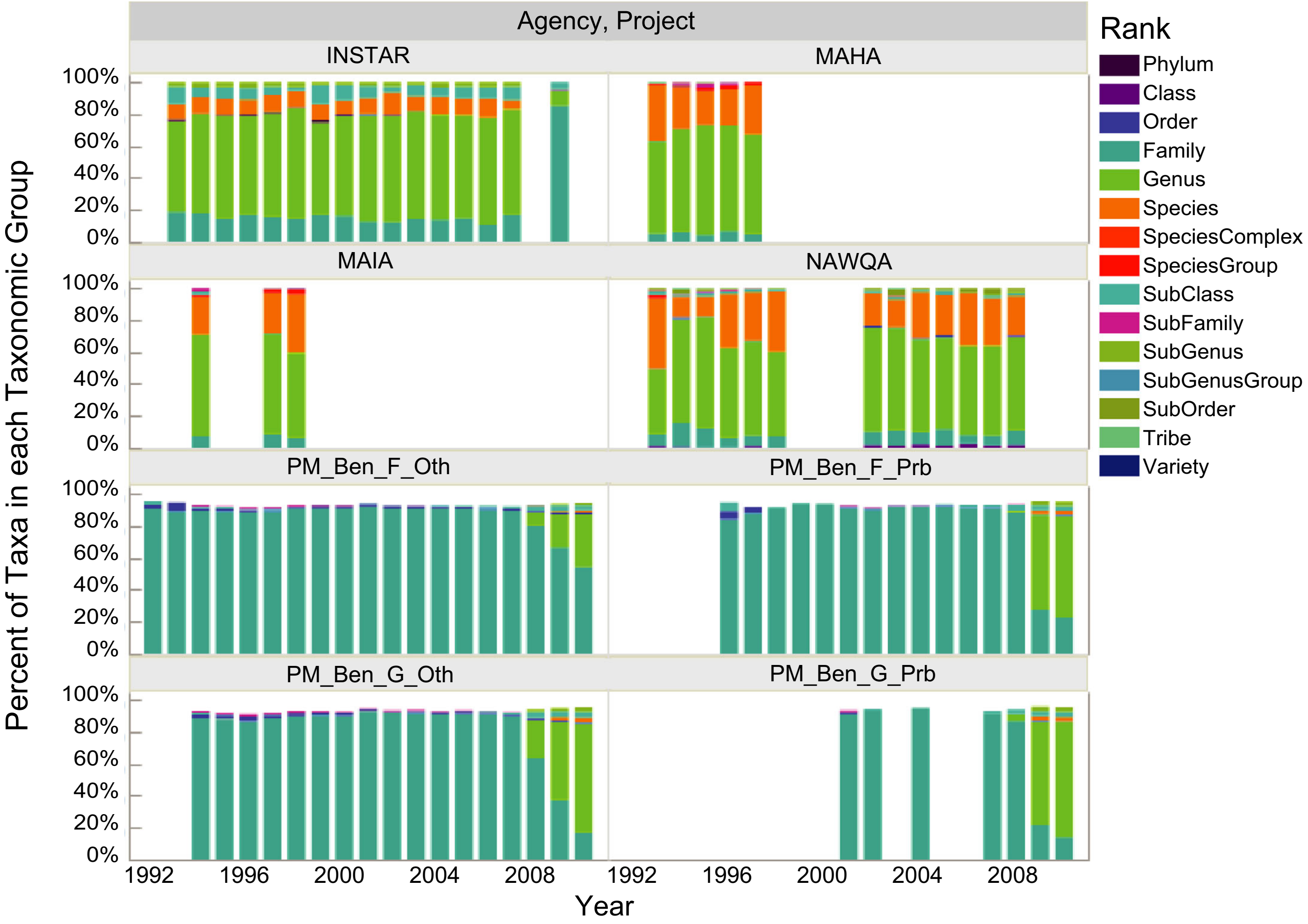

FIGURE D1. Distribution of benthic macroinvertebrate sample taxa rank by agency or project and sampling date. [Project or Agency names from EDAS database (Tetra Tech Inc 2012): INSTAR, Virginia Commonwealth University INteractive STream Assessment Resource; MAHA, USEPA Medical Advocates for Healthy Air; MAIA, USEPA Mid Atlantic Integrated Assessment; NAWQA, U.S. Geological Survey (USGS) National Water Quality Assessment Program data from the USGS BioData Retrieval system database (BioData); PM-Ben-F-Oth, VA DEQ — Probabilistic Monitoring Program — Benthos, Family IDs, non-Probabilistic Sites; PM-Ben-F-Prb, VA DEQ — Probabilistic Monitoring Program — Benthos, Family IDs, Probabilistic Sites; PM-Ben-G-Oth, VA DEQ — Probabilistic Monitoring Program — Benthos, Genus IDs, non-Probabilistic Sites; PM-Ben-G-Prb, VA DEQ — Probabilistic Monitoring Program — Benthos, Genus IDs, Probabilistic Sites].

species" grouping. Table A1 outlines projects and date ranges that define each sample resolution group.

SUPPORTING INFORMATION

Additional supporting information may be found online under the Supporting Information tab for this article: See Kleiner, J. 2020. "elfgen." GitHub repository. (available online: https://github.com/HARPgroup/elfgen) for additional notes, along with the modeling tools we developed in the "elfgen" R package, which can be used to generate ELFs using built-in datasets for the conterminous U.S., or applied anywhere else stream size based characteristics and species richness data inputs are available.

ACKNOWLEDGMENTS

This study benefited from the development of a comprehensive database, Ecological Data Application System (EDAS) compiled by Tetra Tech Inc (2012), who checked locations, removed duplicate records, and compiled multiple-source data in EDAS. The authors are grateful for access to eight sources for ecological data from: DEQ Probabilistic Monitoring (ProbMon) database; Virginia Department of Game and Inland Fisheries Virginia Fish and Wildlife Information Service (VaFWIS) database (provided through cooperative agreement with VDGIF); two Environmental Protection Agency sources: Medical Advocates for Healthy Air (MAHA) and Mid Atlantic Integrated Assessment (MAIA) databases; USGS

National Water Quality Assessment (NAWQA) data sourced from the BioData Retrieval system database (BioData); Virginia Commonwealth University INteractive STream Assessment Resource (INSTAR) database; and Multistate Aquatic Resources Information System (MARIS) including Tennessee, and Virginia Department of Conservation and Recreation Natural Heritage database. Each of these data providers shared information with DEQ and Tetra Tech in 2010 for the historic period of record. The authors acknowledge the various reviewers at The Virginia DEQ, USGS, and Virginia Tech who provided critical input throughout the project development process. Any use of trade, firm, or product names is for descriptive purposes only, and does not imply endorsement by the U.S. Government.

AUTHORS' CONTRIBUTIONS

Joseph Kleiner: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing-original draft; writing-review & editing. **Elaina Passero:** Data curation; formal analysis; investigation; methodology; resources; software; validation; visualization; writing-original draft. **Robert Burgholzer:** Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing-original draft; writing-review & editing. **Jennifer Rapp:** Conceptualization; data curation; formal analysis; investigation; methodology; resources; software; supervision; validation; visualization; writing-original draft; writing-review & editing. **Durelle Scott:** Conceptualization; data curation; formal analysis; investigation; methodology; resources; software; supervision; validation; visualization; writing-original draft; writing-review & editing.

LITERATURE CITED

Angermeier, P.L., and I.J. Schlosser. 1989. "Species-Area Relationship for Stream Fishes." *Ecology* 70 (5): 1450–62.

Angermeier, P.L., and M.R. Winston. 1999. "Characterizing Fish Community Diversity across Virginia Landscapes: Prerequisite for Conservation." *Ecological Applications* 9 (1): 335–49.

Armstrong, D.T. Richards, and S. Levin. 2011. "Factors Influencing Riverine Fish Assemblages in Massachusetts." *Technical Report, U.S. Geological Survey Scientific-Investigations Report 2011–5193*, 58 pp.

Arthington, A.H., R.J. Naiman, M.E. McClain, and C. Nilsson. 2010. "Preserving the Biodiversity and Ecological Services of Rivers: New Challenges and Research Opportunities." *Freshwater Biology* 55 (1): 1–16.

Beecher, H.A., E.R. Dott, and R.F. Fernau. 1988. "Fish Species Richness and Stream Order in Washington State Streams." *Environmental Biology of Fishes* 22 (3): 193–209.

Brooks, A.J., and T. Haeusler. 2016. "Invertebrate Responses to Flow: Trait-Velocity Relationships during Low and Moderate Flows." *Hydrobiologia* 773 (1): 23–34.

Cade, B.S., J.W. Terrell, and R.L. Schroeder. 1999. "Estimating Effects of Limiting Factors with Regression Quantiles." *Ecology* 80 (1): 311–23.

Carlisle, D.M., S.M. Nelson, and K. Eng. 2012. "Macroinvertebrate Community Condition Associated with the Severity of Streamflow Alteration." *River Research and Applications* 30 (1): 29–39.

Cuffney, T.F., and J.G. Kennen. 2017. "Potential Pitfalls of Aggregating Aquatic Invertebrate Data from Multiple agency sources: Implications for Detecting Aquatic Assemblage Change across Alteration Gradients." *Freshwater Biology* 63 (8): 738–51.

Cuffney, T.F., M.D. Bilger, and A.M. Haigler. 2007. "Ambiguous Taxa: Effects on the Characterization and Interpretation of Invertebrate Assemblages." *Journal of the North American Benthological Society* 26 (2): 286–307.

Dewson, Z.S., A.B.W. James, and R.G. Death. 2007. "Stream Ecosystem Functioning under Reduced Flow Conditions." *Ecological Applications* 17 (6): 1797–808.

Döll, P., and J. Zhang. 2010. "Impact of Climate Change on Freshwater Ecosystems: A Global-Scale Analysis of Ecologically Relevant River Flow Alterations." *Hydrology and Earth System Sciences* 14 (5): 783–99.

Filipe, A.F., M. Filomena Magalhães, and M.J. Collares-Pereira. 2010. "Biodiversity Research: Native and Introduced Fish Species Richness in Mediterranean Streams: The Role of Multiple Landscape Influences." *Diversity and Distributions* 16 (5): 773–85.

Frimpong, E.A., J. Huang, and Y. Liang. 2016. "IchthyMaps: A Database of Historical Distributions of Freshwater Fishes of the United States." *Fisheries* 41 (10): 590–99.

Gido, K.B., D.L. Propst, J.D. Olden, and K.R. Bestgen. 2013. "Multidecadal Responses of Native and Introduced Fishes to Natural and Altered Flow Regimes in the American Southwest." *Canadian Journal of Fisheries and Aquatic Sciences* 70 (4): 554–64.

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)." *Technical Report, U.S. Geological Survey, Biological Resources Discipline, Open File Report 2006–1093*.

Iwasaki, Y., M. Ryo, P. Sui, and C. Yoshimura. 2012. "Evaluating the Relationship between Basin-Scale Fish Species Richness and Ecologically Relevant Flow Characteristics in Rivers Worldwide." *Freshwater Biology* 57 (10): 2173–80.

Jenkins, R., and N. Burkhead. 1994. *Freshwater Fishes of Virginia*. Bethesda, MD: American Fisheries Society.

Jowett, I.G., J. Richardson, and M.L. Bonnett. 2005. "Relationship between Flow Regime and Fish Abundances in a Gravel-Bed River, New Zealand." *Journal of Fish Biology* 66 (5): 1419–36.

Kennen, J.G., M.L. Riskin, and E.G. Charles. 2014. "Effects of Streamflow Reductions on Aquatic Macroinvertebrates: Linking Groundwater Withdrawals and Assemblage Response in Southern New Jersey Streams, USA." *Hydrological Sciences Journal* 59 (3–4): 545–61.

Knight, R.R., J.C. Murphy, W.J. Wolfe, C.F. Saylor, and A.K. Wales. 2014. "Ecological Limit Functions Relating Fish Community Response to Hydrologic Departures of the Ecological Flow Regime in the Tennessee River Basin, United States." *Ecohydrology* 7 (5): 1262–80.

Koenker, R., and G. Bassett, Jr. 1978. "Regression Quantiles." *Econometrica* 46 (1): 33–50.

Lemoine, N.2012. "R for Ecologists: Putting Together a Piecewise Regression." https://www.r-bloggers.com/r-for-ecologists-putting-together-a-piecewise-regression/.

McGarvey, D., and G. Ward. 2008. "Scale Dependence in the Species-Discharge Relationship for Fishes of the Southeastern U.S.A." *Freshwater Biology* 53: 2206–19.

McKay, L., T. Bondelid, T. Dewald, J. Johnston, R. Moore, and A. Rea. 2012. “Nhdplus Version 2: User Guide.” http://ftp.horizon-systems.com/NHDplus/NHDPlusV21/Documentation/NHDPlusV2_User_Guide.pdf.

McManamay, R.A., D.J. Orth, C.A. Dolloff, and E.A. Frimpong. 2012. “A Regional Classification of Unregulated Stream Flows: Spatial Resolution and Hierarchical Frameworks.” *River Research and Applications* 28 (7): 1019–33.

Minshall, G.W., R.C. Petersen, and C.F. Nimz. 1985. “Species Richness in Streams of Different Size from the Same Drainage Basin.” *The American Naturalist* 125 (1): 16–38.

Novak, R., J.G. Kennen, R.W. Abele, C.F. Baschon, D.M. Carlisle, L. Dlugolecki, D.M. Eignoret al.2016. “Final EPA-USGS Technical Report: Protecting Aquatic Life from Effects of Hydrologic Alteration.” *Technical Report, U.S. Geological Survey Scientific Investigations Report*.

Omernik, J.M., and G.E. Griffith. 2014. “Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework.” *Environmental Management* 54 (6): 1249–66.

Ostby, B.J., J.L. Krstolic, and G.C. Johnson. 2014. “Reach-Scale Comparison of Habitat and Mollusk Assemblages for Select Sites in the Clinch River with Regional Context.” *Journal of the American Water Resources Association* 50 (4): 859–77.

Paller, M.H., W.L. Specht, and S.A. Dyer. 2006. “Effects of Stream Size on Taxa Richness and Other Commonly Used Benthic Bioassessment Metrics.” *Hydrobiologia* 568 (1): 309–16.

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 (11): 769–84.

Poff, N., B.D. Richter, A.H. Arthington, S.E. Bunn, R.J. Naiman, E. Kendy, M. Acreman et al. 2010. “The Ecological Limits of Hydrologic Alteration (ELOHA): A New Framework for Developing Regional Environmental Flow Standards.” *Freshwater Biology* 55 (1): 147–70.

Poff, N.L., and J.K.H. Zimmerman. 2010. “Ecological Responses to Altered Flow Regimes: A Literature Review to Inform the Science and Management of Environmental Flows.” *Freshwater Biology* 55 (1): 194–205.

Pracheil, B.M., P.B. McIntyre, and J.D. Lyons. 2013. “Enhancing Conservation of Large-River Biodiversity by Accounting for Tributaries.” *Frontiers in Ecology and the Environment* 11 (3): 124–28.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rapp, J.L., and P.A. Reilly. 2017. “Virginia Flow-Ecology Modeling Results: An Initial Assessment of Flow Reduction Effects on Aquatic Biota.” *Technical Report, U. S: Geological Survey Open-File Report*.

Richter, B.D., M.M. Davis, C. Apse, and C. Konrad. 2012. “A Presumptive Standard for Environmental Flow Protection.” *River Research and Applications* 28 (8): 1312–21.

Rolls, R.J., C. Leigh, and F. Sheldon. 2012. “Mechanistic Effects of Low-Flow Hydrology on Riverine Ecosystems: Ecological Principles and Consequences of Alteration.” *Freshwater Science* 31 (4): 1163–86.

Rosenfeld, J.S. 2017. “Developing Flow–Ecology Relationships: Implications of Nonlinear Biological Responses for Water Management.” *Freshwater Biology* 62 (8): 1305–24.

Schlosser, I.J. 1987. “A Conceptual Framework for Fish Communities in Small Warmwater Streams.” In *Community and Evolutionary Ecology of North American Stream Fishes*, edited by W. Matthews, and D. Heins, 17–24. Norman, OK: University of Oklahoma Press.

Schmidt, T.S., W.H. Clements, and B.S. Cade. 2012. “Estimating Risks to Aquatic Life Using Quantile Regression.” *Freshwater Science* 31 (3): 709–23.

Terrell, J.W., B.S. Cade, J. Carpenter, and J.M. Thompson. 1996. “Modeling Stream Fish Habitat Limitations from Wedge-Shaped Patterns of Variation In Standing Stock.” *Transactions of the American Fisheries Society* 125 (1): 104–17.

Tetra Tech Inc. 2012. “Virginia Ecological Limits of Hydrologic Alteration (ELOHA): Development of Metrics of Hydrologic Alteration.” *Technical Report, Draft Report Prepared for the U.S. Environmental Protection Agency and Virginia Department of Environmental Quality*.

Vander Vorste, R., P. McElmurray, S. Bell, K.M. Eliason, and B.L. Brown. 2017. “Does Stream Size Really Explain Biodiversity Patterns in Lotic Systems? A Call for Mechanistic Explanations.” *Diversity* 9 (3): 26.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell, and C.E. Cushing. 1980. “The River Continuum Concept.” *Canadian Journal of Fisheries and Aquatic Sciences* 37 (1): 130–37.

Vaz, S., C.S. Martin, P.D. Eastwood, B. Ernande, A. Carpentier, G.J. Meaden, and F. Coppin. 2008. “Modelling Species Distributions Using Regression Quantiles.” *Journal of Applied Ecology* 45 (1): 204–17.

Walters, D.M., D.S. Leigh, M.C. Freeman, B.J. Freeman, and C. Pringle. 2003. “Geomorphology and Fish Assemblages in a Piedmont River Basin, U.S.A.” *Freshwater Biology* 48 (11): 1950–70.

Ward, J. 1998. “Riverine Landscapes: Biodiversity Patterns, Disturbance Regimes, and Aquatic Conservation.” *Biological Conservation* 83 (3): 269–78.

Xenopoulos, M.A., and D.M. Lodge. 2006. “Going with the Flow: Using Species-Discharge Relationships to Forecast Losses in Fish Biodiversity.” *Ecology* 87 (8): 1907–14.

Xenopoulos, M.A., D.M. Lodge, J. Alcamo, M. Märker, K. Schulze, and D.P. Van Vuuren. 2005. “Scenarios of Freshwater Fish Extinctions from Climate Change and Water Withdrawal.” *Global Change Biology* 11 (10): 1557–64.

Xu, Z.-H., X.-A. Yin, C. Zhang, and Z.-F. Yang. 2016. “Piecewise Model for Species–Discharge Relationships in Rivers.” *Ecological Engineering* 96: 208–13.

Zipper, C.E., B. Beaty, G.C. Johnson, J.W. Jones, J.L. Krstolic, B.J. Ostby, W.J. Wolfe, and P. Donovan. 2014. “Freshwater Mussel Population Status and Habitat Quality in the Clinch River, Virginia and Tennessee, USA: A Featured Collection.” *Journal of the American Water Resources Association* 50 (4): 807–19.

Zorn, T.G., P.W. Seelbach, and E.S. Rutherford. 2012. “A Regional-Scale Habitat Suitability Model to Assess the Effects of Flow Reduction on Fish Assemblages in Michigan Streams.” *Journal of the American Water Resources Association* 48 (5): 871–95.