Accepted: 9 October 2017
DOI: 10.1111/fwb.13048

SPECIAL ISSUE

WILEY Freshwater Biology

Using regional scale flow–ecology modeling to identify catchments where fish assemblages are most vulnerable to changes in water availability

Ernie F. Hain[1] | Jonathan G. Kennen[2] | Peter V. Caldwell[3] | Stacy A.C. Nelson[1] | Ge Sun[4] | Steven G. McNulty[4]

[1]Department of Forestry and Environmental Resources, Center for Geospatial Analytics, North Carolina State University, Raleigh, NC, U.S.A.

[2]US Geological Survey, Lawrenceville, NJ, U.S.A.

[3]Center for Forest Watershed Research, Coweeta Hydrologic Laboratory, USDA Forest Service, Otto, NC, U.S.A.

[4]Eastern Forest Environmental Threat Assessment Center, USDA Forest Service, Raleigh, NC, U.S.A.

Correspondence
Ernie F. Hain, North Carolina State University, Raleigh, NC, U.S.A.
Email: ernie_hain@ncsu.edu

Funding information
U.S. Department of the Interior Southeast Climate Center; U.S. Department of Agriculture Forest Service Eastern Forest Environmental Threat Assessment Center; U.S. Geological Survey National Water Census

Abstract

1. Streamflow is essential for maintaining healthy aquatic ecosystems and for supporting human water supply needs. Changes in climate, land use and water use practices may alter water availability. Understanding the potential effect of these changes on aquatic ecosystems is critical for long-term water management to maintain a balance between water for human consumption and ecosystem needs.
2. Fish species data and streamflow estimates from a rainfall-runoff and flow routing model were used to develop boosted regression tree models to predict the relationship between streamflow and fish species richness (FSR) under plausible scenarios of (1) water withdrawal, (2) climate change and (3) increases in impervious surfaces in the Piedmont ecoregion of North Carolina, U.S.A. Maximum monthly flow, the fraction of total flow originating from impervious surface runoff, coefficient of monthly streamflow variability, and the specific river basin accounted for 50% of the variability in FSR. This model was used to predict FSR values for all twelve-digit Hydrological Unit Code catchments (HUC-12s) in the North Carolina Piedmont under current flow conditions and under water withdrawal, climate change and impervious surface scenarios.
3. Flow–ecology modeling results indicate that predicted FSR declined significantly with increased water withdrawals. However, the magnitude of decline varied geographically. A "hot-spot" analysis was conducted based on predicted changes in FSR under each scenario to understand which HUC-12s were most likely to be affected by changes in water withdrawals, climate and impervious surfaces. Under the 20% withdrawal increase scenario, 413 of 886 (47%) HUC-12s in the study area were predicted to lose one or more species. HUC-12s in the Broad, Catawba, Yadkin and Cape Fear river basins were most susceptible to species loss.
4. These findings may help decision making efforts by identifying catchments most vulnerable to changing water availability. Additionally, FSR-discharge modeling results can assist resource agencies, water managers and stakeholders in assessing the effect of water withdrawals in catchments to better support the protection and long-term conservation of species.

KEYWORDS
boosted regression trees, environmental water, fish species richness, flow–ecology models, water withdrawal

1 | INTRODUCTION

Environmental water studies over the last two decades have emphasised the inherent linkage and in some cases, the tension, between maintaining water for human use as well as for ecosystem needs (Acreman et al., 2008; Kendy, Apse, & Blann, 2012; Poff et al., 2010; Shenton, Bond, Yen, & Mac Nally, 2012). Substantial emphasis has been placed on the broad discharge patterns that influence the structural, functional and life-history strategies of biotic communities (Bunn & Arthington, 2002; Mims & Olden, 2012; Naiman, Latterell, Pettit, & Olden, 2008). More recently, there has been an emphasis on developing hydrological indices for characterising the flow regime (Henriksen, Heasley, Kennen, & Nieswand, 2006; Monk, Wood, Hannah, & Wilson, 2007; Worrall et al., 2014), systematically arranging streams and rivers into specific stream classes with respect to flow regime characteristics (Archfield et al., 2013; Kennard et al., 2010; Kennen, Henriksen, Heasley, Cade, & Terrell, 2009; Kennen, Henriksen, & Nieswand, 2007; McManamay, Orth, Dolloff, & Frimpong, 2012; Olden & Poff, 2003), and building flow–ecology response models that link changes in streamflow and water availability to changes in assemblage structure and function (e.g. Arthington, Bernardo, & Ilhéu, 2014; Chessman, Jones, Searle, Growns, & Pearson, 2010; Freeman et al., 2013; Kennen, Riskin, & Charles, 2014; Kennen, Riva-Murray, & Beaulieu, 2010; McManamay, Orth, Dolloff, & Mathews, 2013; Stewart-Koster, Olden, & Gido, 2014; Turner & Stewardson, 2014). All of these studies emphasise the identification of the streamflow components needed to help determine ecological and environmental endpoints and the inherent linkages between changes in streamflow processes and ecosystem response.

Changes in riparian and watershed scale land-use and associated alterations in stream habitat and streamflow processes have been linked to declines in native stream fish populations (Olden, 2016) and a general downward trend in aquatic biodiversity across the globe (Dudgeon et al., 2006; Vörösmarty et al., 2010). While minimising impervious surfaces and maximising the conservation of contiguous tracts of forested lands in catchments supports the preservation of stream fish populations (Kennen, Chang, & Tracy, 2005), alterations in water availability including impoundments, streamflow regulation and water-resource development, which are essential to meet the water needs of a growing population, are strongly linked to changes in native fish diversity, abundance and resilience (Conroy, Allen, Peterson, Pritchard, & Moore, 2003; Poff & Zimmerman, 2010; Warren et al., 2000). Confounding the effects of land-use change and streamflow alteration are projected increases in drought frequency and duration associated with climate change (IPCC, 2013; Melillo, Richmond, & Yohe, 2014) which place further stress on water supplies and fish assemblage structure (Keaton, Haney, & Andersen, 2005; Matthews & Marsh-Matthews, 2003; White, Mchugh, & Mcintosh, 2016). Understanding these linkages and the potential effect of changes in water availability on aquatic ecosystems is critical for long-term water management in areas facing significant water stress, especially when the needs of humans and aquatic ecosystems appear to conflict and sometimes result in legal confrontations. This is periodically the case in the southeast United States where water stress is known to occur in conjunction with drought cycles (e.g. Seager, Tzanova, & Nakamura, 2009). It is under these conditions that water managers must identify areas of concern and make informed decisions about water conservation that affect human and ecological use in these areas. Unfortunately, there is a lack of decision support tools that identify areas of concern across broad regions, especially tools with a spatial resolution relevant to management decision making.

The primary objective of this study was to demonstrate the efficacy of using relatively simple, large-scale hydrologic models in conjunction with ecological data to develop empirical flow–ecology response models that predict the effect of changes in water availability on fish species richness (FSR), an easily quantified assemblage metric. Additionally, we sought to use this modeling approach to identify catchments or "hot-spots" of FSR change under a plausible set of future land use, climate and withdrawal change scenarios and test the hypothesis that FSR in the North Carolina Piedmont, U.S.A. will decrease with predicted changes in climate, increases in urban land use (i.e. impervious surfaces) and increases in water withdrawals.

2 | METHODS

2.1 | Study area

We focused on catchments in the Piedmont ecoregion of North Carolina to develop empirical relationships between stream flow and FSR (Figure 1a). Seven major river basins are partially located within the Piedmont region. These include the Broad (BRD), Cape Fear (CPF), Catawba (CTB), Neuse (NEU), Roanoke (ROA), Tar (TAR) and Yadkin (YAD) river basins (Figure 1). The North Carolina Piedmont contains or intersects 886 twelve-digit Hydrological Unit Code catchments (hereafter HUC-12s) identified by the Natural Resources Conservation Service's (NRCS) Watershed Boundary Dataset GIS layer. North Carolina has one of the highest rates of population growth in the U.S., and is now in the top ten most populous states (US Census Bureau, 2016). Much of this growth has occurred in the North Carolina Piedmont, which contains many of the state's most

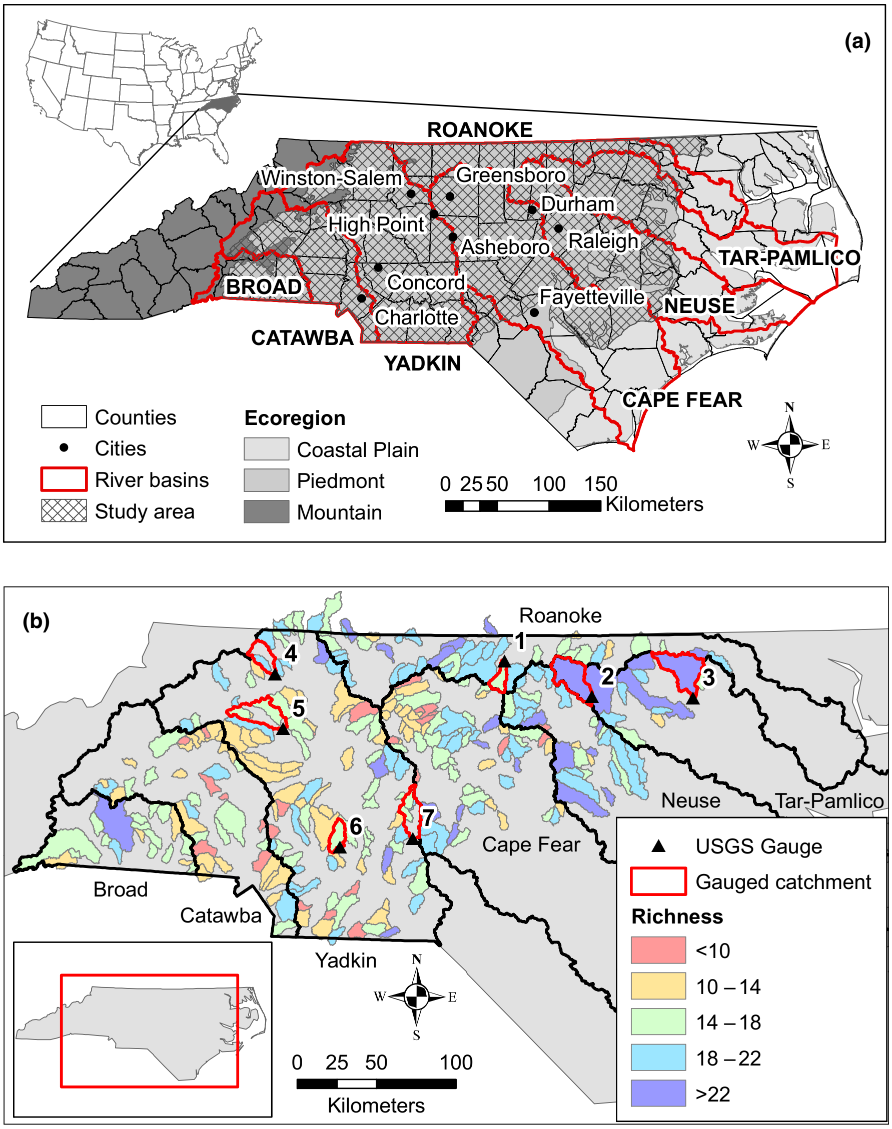

FIGURE 1 Map of North Carolina, U.S.A. showing the study area spanning the Piedmont ecoregion of North Carolina, as well as parts of the Mountain and Coastal Plain ecoregions (a), and the range in fish species richness values for the boosted regression tree model training dataset in the NC Piedmont and locations of seven USGS reference gauges used for Water Supply Stress Index model validation (b). Study area extent was chosen based on best professional judgement and discussions with North Carolina Division of Water Resources personnel (a). Each polygon in (b) represents the delineated contributing watershed for each sample site

rapidly growing cities (i.e. Charlotte, Durham, Raleigh and Winston-Salem) (Figure 1a).

The Piedmont region of North Carolina (Figure 1a) has a humid sub-tropical climate of warm summers and cool moist winters. The region receives, on average, 107–117 cm/year of precipitation (NC Climate Office, 2016), and is made up of gently rolling forested hills, with altitudes ranging from 60 to 470 m. The geology of the Piedmont is dominated by metamorphic (gneisses and schist) and igneous (granite, diorite and gabbro) rocks overlain by “clayey” ultisols (soils with light upper layers and a reddish sub-soil) that were mainly formed through physical weathering and alluvial processes. These soils are rich in aluminium and silicates and contain eroded sediments mixed with organic material. Natural vegetative cover in this region consists mainly of mesic mixed hardwoods (e.g. american beech, tulip poplar, hickory, and red and white oak) though wetlands occur in some lower altitudes and patches of pine forests are found in more xeric regions. Population growth and development in the North Carolina Piedmont have altered the natural landscape and increased water demand. Surface- and groundwater withdrawals have reduced baseflow and estimates of water use show that in 2010 the total gross fresh surface water withdrawals across the 54 counties in the region amounted to 11.7 billion m^3 $year^{-1}$ (Maupin et al., 2014).

2.2 | Modelling approach

For this study, we implemented a multistep modelling approach (Figure 2). First, FSR was calculated for 385 distinct fish sampling sites in the NC Piedmont using data collected by the North Carolina Division of Water Resources (NCDWR). Average monthly streamflow for the sites was then predicted using the well-documented Water Supply Stress Index (WaSSI) model (Caldwell et al., 2015). FSR and streamflow predictions were used to build a boosted regression tree (BRT) flow–ecology model which was used to predict the relationship between a subset of ecologically relevant streamflow metrics and FSR in all HUC-12s in the NC Piedmont. The results of the model were then used to predict FSR under three plausible scenarios of future water withdrawals, climate change and increases in impervious surfaces (Figure 2). Finally, a "hot-spot" analysis was used to identify individual HUC-12s that were most likely to be affected by changes in water availability.

2.3 | Biological data aggregation

The Biological Assessment Unit (BAU) of the NCDWR began sampling each of the state's 17 river basins on a 5-year rotation in 1990 (NCDENR, 2006). Streams wadeable from shoreline to shoreline were sampled by the stream fish community assessment program for an average distance of 183 m (600 ft.). A four-person team collected all fish at each site using a modified two-pass depletion technique with two backpack electrofishing units and two persons netting. All fish were identified to species, enumerated, inspected for disease and deformities, and total length was measured before fish were released back into the stream. Specimens not easily identified in the field were preserved in 10% formalin and transported to the BAU laboratory in Raleigh, North Carolina. Between 1990 and 2012, 967 sampling events were performed by BAU at 385 unique sampling stations in the North Carolina Piedmont region. Where unique stations were sampled multiple times, measures of FSR were averaged across all sampling events, resulting in a final sample size of 385 for developing flow–ecology relationships.

Upstream contributing catchments were delineated for each of the 385 unique NCDWR sampling stations (Figure 1b). Delineations were performed using Arc Hydro tools in ArcGIS 10.1 using digital elevation models procured from the North Carolina Floodplain Mapping Program (NC Floodplain Mapping Program, 2013).

2.4 | Streamflow prediction

Streamflow was predicted for all study catchments using the WaSSI model. WaSSI was developed by the U.S. Forest Service to assess the effect of climate change, land-use change and population growth on water supply stress, river flows and aquatic ecosystems across the contiguous U.S. (Caldwell, Sun, McNulty, Cohen, & Myers, 2012; Caldwell et al., 2015; Sun et al., 2011). WaSSI has been successfully used in climate change assessments in the eastern U.S. (Lockaby et al., 2011; Marion et al., 2013; Sun et al., 2013; Tavernia, Nelson, Caldwell, & Sun, 2013) and examining the nexus of water and energy at the national scale (Averyt et al., 2011, 2013). WaSSI is an integrated monthly water balance and flow routing model that simulates the full hydrologic cycle for each of 10 land cover classes at the HUC-12 scale. The 10 land cover classes are aggregated from the 17 classes of the 2006 National Land Cover Dataset (NLCD) (Fry et al., 2011). Infiltration, surface runoff, soil moisture and baseflow processes for each HUC-12 catchment land cover were computed using algorithms of the Sacramento Soil Moisture Accounting Model (SAC-SMA) (Burnash, 1995; Burnash, Ferral, & Mcguire, 1973). State Soil Geographic (STATSGO) databases (NRCS, 2012) were used to compute the 11 SAC-SMA soil input parameters (Koren, Smith, & Duan, 2003). Monthly evapotranspiration (ET) was modelled with an empirical equation derived from multisite eddy covariance ET measurements (Sun et al., 2011). Required data to estimate ET included monthly mean Moderate Resolution Imaging Spectroradiometer (MODIS) MOD15A2 leaf area index (LAI) (Zhao, Heinsch, Nemani, & Running, 2005), potential ET (PET) calculated as a function of temperature and latitude (Hamon, 1963), and precipitation (PPT). This estimate of ET was then constrained by the soil water content

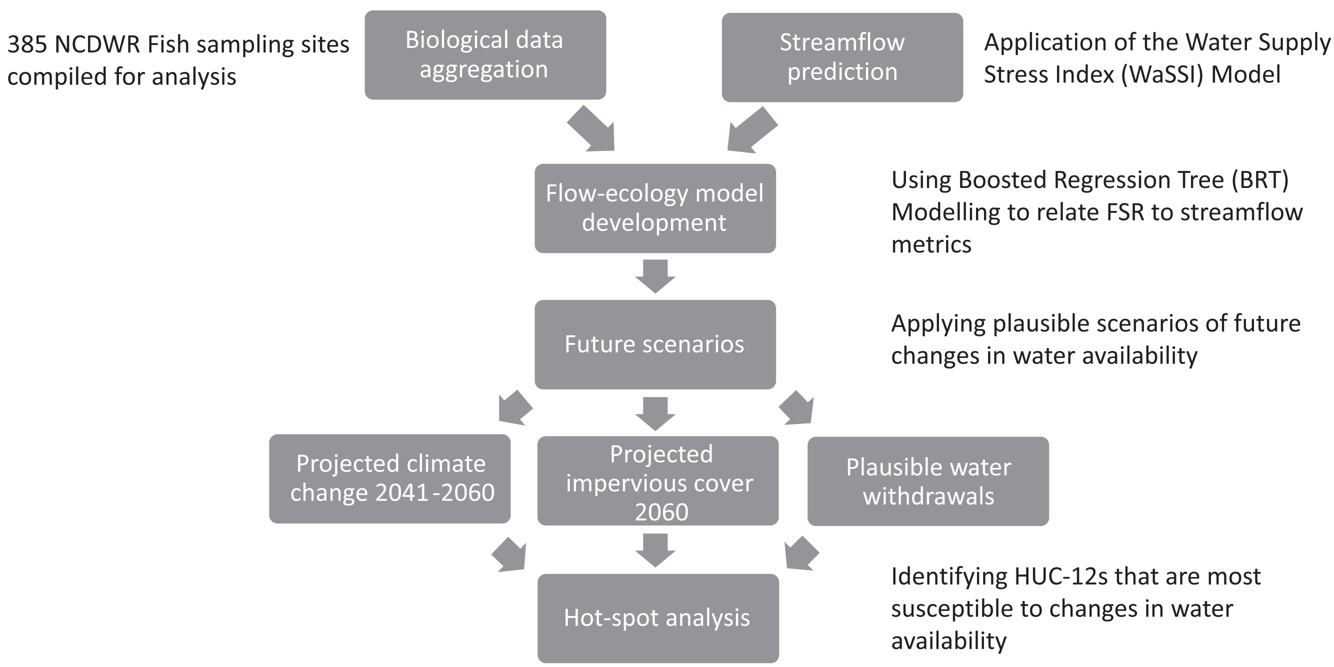

FIGURE 2 Study modelling approach. Biological data were aggregated from North Carolina Division of Water Resources (NCDWR) stream community assessment data. Streamflow was predicted using the Water Supply Stress Index (WaSSI) model

computed by the SAC-SMA algorithm during extreme water-limited conditions. Monthly precipitation and air temperature inputs were based on Precipitation Elevation Regression on Independent Slopes Model (PRISM) estimates (PRISM Climate Group, 2013). All water balance components were computed independently for each land cover class within each catchment and accumulated to estimate the totals for the catchment. For the NLCD-based impervious cover fraction, storage and ET were assumed to be negligible, thus all precipitation falling on the impervious portion of a catchment for a given month was assumed to generate surface runoff in the same month, and was routed directly to the catchment outlet.

Although the WaSSI model can be calibrated, no calibration of model input parameters were performed for this study. WaSSI was developed to include the key ecohydrological processes that affect the water balance with off-the-shelf input datasets while having an acceptable level of predictive performance without calibration. In doing so, the model is not subject to the complexities and uncertainties associated with transferring model parameters from calibrated to ungauged catchments (Sivapalan et al., 2003) and using the model to assess the impact of climate or land cover scenarios outside of the conditions for which the model is calibrated. Despite being uncalibrated, WaSSI has been found to have similar predictive performance at the monthly time-step to other calibrated, process-based models (Caldwell et al., 2015).

Streamflow was estimated using the monthly WaSSI output from the years 1991–2010. This output represented the average streamflow for each month of the year over the 20-year period and resulted in 12 average monthly streamflow values for each site. We calculated a suite of ecologically relevant streamflow statistics from these monthly averages targeting flow metrics that account for the magnitude and seasonality (timing) of streamflow (Poff et al., 1997) in an effort to capture flow signals important to fish life history. This included minimum monthly streamflow in m^3/s (low flow; MinCMS), maximum monthly streamflow (high flow; MaxCMS), average monthly streamflow (average flow; AveCMS), average streamflow for April, May, and June (spring flow; AMJCMS), average streamflow for August (summer flow; AugCMS) and a measure of the coefficient of monthly streamflow variability (FloVar). FloVar was computed by dividing the standard deviation of the decreased monthly flow values by the original monthly mean streamflow. This measure of streamflow variability is equivalent to "MA39," a common measure of monthly streamflow variability presented in Olden and Poff (2003) and Henriksen et al. (2006). Additionally, the proportion of surface flow coming from impervious surfaces was estimated for each catchment (QFRAC_IMP). All statistics were estimated for contributing catchments to each NCDWR sample site, as well as all HUC-12s in the North Carolina Piedmont. The monthly time-step allows us to evaluate the ability of large scale, easily parameterised models to provide flow information that is useful for determining changes in ecosystem integrity.

We validated the WaSSI model flow predictions by computing classical hydrological model fit statistics as well as the prediction of ecologically relevant flow statistics at seven USGS gauges located in the study area (Table 1, Figure 1b). Gauges identified as reference sites in the USGS Gages II database (Falcone, 2011; Falcone, Carlisle, Wolock, & Meador, 2010) were selected for validation procedures, and were either co-located with a fish training site or had one or more fish training sites upstream. Classical fit statistics evaluated included bias in mean streamflow, the Nash–Sutcliffe Efficiency (NSE) statistic (Nash & Sutcliffe, 1970), the root mean squared error and the coefficient of determination (R^2). Flow statistics evaluated included MaxCMS and FloVar that were used in the 4-variable flow–ecology model (described below). Bias in mean streamflow within ±25%, 15%, and 10% were considered indicative of satisfactory, good and very good hydrological model performance, respectively (Moriasi et al., 2007). Similarly, NSE values that are greater than 0.50, 0.65 and 0.75 for prediction of monthly streamflow were considered to be indicative of satisfactory, good and very good model performance, respectively (Moriasi et al., 2007). The NSE can range from negative infinity to 1.0, the closer NSE is to 1.0 the better the model fit. Negative values of NSE indicate that using the mean of the observations provides a better fit than the model. A hydrologic uncertainty of ±30% was used to aid in placing model prediction bias of flow statistics into context with inherent variability in streamflow and flow measurement (Murphy, Knight, Wolfe, & Gain, 2013).

2.5 | Flow–ecology model development

We developed a FSR BRT model using observed biological data and WaSSI streamflow predictions for all 385 NCDWR sample sites in the training dataset. BRT models are only briefly described here as their use and technical details (e.g. Breiman, Friedman, Olshen, & Stone, 1984; De'ath & Fabricius, 2000; Prasad, Iverson, & Liaw, 2006), as well as application (Aertsen, Kint, Van Orshoven, Özkan, & Muys, 2010; Brown et al., 2012; Clapcott, Young, Goodwin, Leathwick, & Kelly, 2011; Elith, Leathwick, & Hastie, 2008; Leclere, Oberdorff, Belliard, & Leprieur, 2011; Waite et al., 2010, 2012) have been widely presented in the literature. BRTs are part of the classification and regression tree (CART) or decision tree family; a family of techniques used to advance single classification or regression trees by averaging the results for each binary split from numerous trees or forests. The objective of BRT models are to reduce the predictive error and improve overall performance (De'ath, 2007; Elith et al., 2008). In BRT, after the initial tree has been developed, successive trees are grown on reweighted versions of the data, giving more weight to cases that are incorrectly classified than those that are correctly classified within each growth sequence (Waite et al., 2012). Thus, as more and more trees are grown in BRT, the large number of trees increases the chance that cases that are difficult to classify initially are correctly classified, thus representing an improvement to the basic averaging algorithm used in random forest (De'ath, 2007). Boosted trees and random forest models retain the positive aspects of single trees seen in CART models, yet have improved predictive performance, nonlinearities and interactions are easily assessed, and they can provide an ordered list of the importance of the explanatory variables (De'ath, 2007; Leclere et al., 2011).

TABLE 1 Summary of classical hydrological model fit statistics and bias in prediction of the flow statistics used in the 4-variable boosted regression tree model across the seven USGS gauges used for Water Supply Stress Index model validation. Bias in mean streamflow within ±25%, 15% and 10% are considered indicative of satisfactory, good and very good hydrological model performance, respectively, while Nash–Sutcliffe efficiency (NSE) values that are greater than 0.50, 0.65 and 0.75 for prediction of monthly streamflow are considered to be indicative of satisfactory, good and very good model performance, respectively. A hydrologic uncertainty of ±30% was used to aid in placing model prediction bias of flow statistics into context with inherent variability in streamflow and flow measurement (Murphy et al., 2013). RMSE, root mean square error; CMS, cubic metres per second. Definitions of ecologically relevant streamflow statistics can be found in the streamflow prediction section

				Classical model fit statistics				Flow statistics	
Site	Gauge	Description	Drainage area, km^2	Bias in mean (%)	NSE	RMSE, cms	R^2	Bias in MaxCMS (%)	Bias in FloVar (%)
1	02077200	Hyco Cr. Near Leasburg, NC	121.7	16%	0.60	0.85	0.63	−12%	−26%
2	02081500	Tar R. near Tar River, NC	428.4	16%	0.74	2.05	0.76	−17%	−34%
3	02082950	Little Fishing Cr. near White Oak, NC	460.9	10%	0.78	2.01	0.79	−17%	−44%
4	02112360	Mitchell R. near State Road, NC	205.3	−10%	0.50	1.17	0.70	2.4%	73%
5	02118500	Hunting Cr. near Harmony, NC	400.5	13%	0.60	2.30	0.69	3.4%	3.7%
6	02125000	Big Bear Cr. near Richfield, NC	144.5	6.9%	0.81	0.76	0.81	−11%	−18%
7	02128000	Little R. near Star, NC	273.5	−8.3%	0.71	1.45	0.72	−16%	0.8%
Mean (standard deviation)			290.7 (140.0)	6.2% (11.0%)	0.68 (0.11)	1.51 (0.62)	0.73 (0.06)	−9.5% (8.9%)	−6.4% (39.1%)

Although BRT offers improved modeling performance over CART, the simple single tree obtained from CART is lost, making it more difficult to visualise the results. Partial dependency plots (PDPs) provide a way to visualise the effect of a specific explanatory variable on the response variable after accounting for the average effects of all other explanatory variables (De'ath, 2007; Elith et al., 2008); PDPs for selected variables important in models appear as examples in the results. BRT models were run using the gbm library in R and specific code from Elith et al. (2008).

Boosted regression tree models were developed using FSR as the response variable and streamflow statistics and river basin as explanatory variables. BRT models were developed using the training dataset (Figure 1b). We used a bag fraction of 0.5, a learning rate of 0.004 and a tree complexity of 3. A bag fraction of 0.5 indicates that each tree is developed using a random selection of 50% of the data. The learning rate influences the total number of trees evaluated for a model, while tree complexity controls whether interactions are fitted, a value of 3 allows the assessment of up to 3-way interactions. Variable relative importance (VRI) was calculated using formulae developed by Friedman (2001) and implemented in the R gbm library to estimate the relative importance of predictor variables (Waite et al., 2012). Calculations of VRI are based on the number of times a variable is selected for splitting, weighted by the squared improvement to the models as a result of each split, averaged over all trees. The relative importance of each variable is scaled so that the sum adds to 100, with higher numbers indicating stronger influence on the modelled response. Due to the size of the training dataset, we implemented a *k*-fold cross-validation technique using the R function gbm.step. The *k*-fold cross-validation splits the dataset into *k* partitions, keeping one partition for testing and the remaining partitions for fitting the model (Hastie, Tibshirani, & Friedman, 2009). This technique generally has low bias and the predictive performance of a *k*-fold cross-validation and validation using an independent dataset are highly similar (Elith et al., 2008). Additionally, *k*-fold cross-validation is known to provide a computational advantage over leave-one-out techniques and provides a more accurate estimate of the test error rate (James, Witten, Hastie, & Tibshirani, 2013). Goodness of fit was measured using the equivalent R^2, estimated as (TD-RD)/TD where TD = total deviance and RD = residual deviance.

Initially, we developed an 8-variable BRT model using the primary subset of ecologically relevant streamflow statistics outlined above. BRT approaches have been shown to overfit models (Aertsen et al., 2010; Elith et al., 2008). Therefore, we developed a reduced-variable model using only those variables identified as having a relative importance greater than 10% (Figure 3). The final model variables were selected after evaluating a Spearman rank correlation matrix of explanatory variables (Table 2), the effects on model fit (i.e. equivalent R^2), and by examining the PDPs of all eight explanatory variables (data not shown). This secondary evaluation allowed us to reduce the number of explanatory variables from eight to four

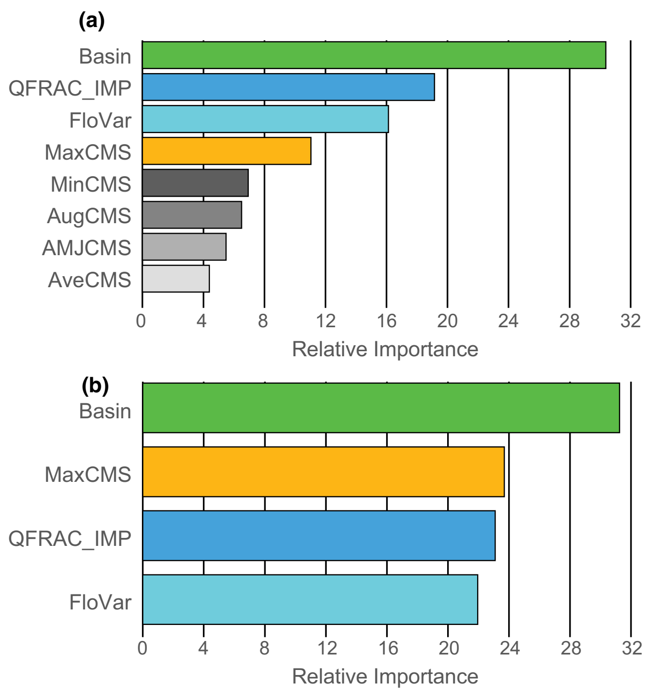

FIGURE 3 Summary of the relative importance of the predictor variables included in (a) the full 8-variable (equivalent R^2 = 0.47) and (b) 4-variable (equivalent R^2 = 0.50) boosted regression tree models

without a loss of variability accounted for by the BRT response model. The final reduced 4-variable model identified those variables most critical for assessing the effects of climate, streamflow and land-use changes on FSR and insured a high level of parsimony for use in future management scenarios. We used the final reduced 4-variable BRT model to predict FSR for all 385 NCDWR training sites and regressed observed FSR against predicted FSR with a 95% confidence interval (Piñeiro, Perelman, Guerschman, & Paruelo, 2008).

2.6 | Scenarios

The final reduced 4-variable model developed for the training data was then used to predict FSR in each HUC-12 in the North Carolina Piedmont under current conditions as well as three future scenarios. These scenarios included (1) projected average climate for the years 2041–60, (2) impervious cover projections for the year 2060, and (3) plausible water withdrawals from each HUC-12. For climate projections, statistically downscaled 1/8 × 1/8° (*c.* 12 × 12 km) 1961–2099 monthly precipitation and temperature predicted by NOAA's Geophysical Fluid Dynamics Laboratory coupled climate model CM2.0 for the A2 growth and emission scenario was obtained from the World Climate Research Programme Coupled Model Intercomparison Project Phase 3 (CMIP3) dataset (Meehl et al., 2007). The Intergovernmental Panel on Climate Change (IPCC) Special Report on Emissions Scenarios (SRES) (Nakicenovic et al., 2000) characterised the A2 storyline as a very heterogeneous world with continuously increasing global population and regionally oriented economic growth with relatively slow technological change. The A2 (High) SRES scenario was selected because it represents potentially the most likely emission scenario as post-2000 global carbon emissions estimates indicate that current emissions are tracking the higher of the SRES emission projections (Raupach et al., 2007). The CM2.0 climate model was selected because it represents a "mid-range" scenario among the 16 climate models evaluated in CMIP3 for the U.S. (Treasure, McNulty, Myers, & Jennings, 2014). Average monthly precipitation and air temperature predictions were estimated for each HUC-12 using area-weighted means. Predicted FSR for the 2040–60 time period was compared to the 1991–2010 time period to evaluate potential climate change effects.

Impervious surface projections for the year 2060 were derived from the U.S. Environmental Protection Agency's Integrated Climate and Land-Use Scenarios (ICLUS) project for A2 growth and emission scenario (Bierwagen et al., 2010; USEPA, 2009) to match the climate change scenario. The ICLUS project develops future impervious surface scenarios that are "broadly consistent with global-scale, peer-reviewed storylines of population growth and economic development" (USEPA, 2009). Projections are based on regression models that relate the 2001 NLCD impervious surface dataset with housing density estimates (a derivative of demographic projections), which enables forecasting likely changes under SRES growth scenarios (scenario A2 was used for this project) (USEPA, 2009). Impervious cover effects were assessed by comparing FSR under projected 2060 impervious cover to that of 2006 using the baseline climate data from 1991 to 2010.

The North Carolina Ecological Flow Science Advisory Board (NCEFSAB) recommended a "flow-by" criteria where ecological flow should be 80%–90% of the instantaneous modelled baseline flow (NCEFSAB, 2013). Consistent with this recommendation, we modelled the effect of reduced flows on FSR by systematically decreasing the amount of total streamflow predicted for each HUC-12 in the North Carolina Piedmont from 5% to 25% at 5% intervals (i.e. a 95%–75% flow-by), thereby bracketing the range recommended by the NCEFSAB. We then used these decreased flow values as explanatory variables in our BRT prediction models. Welch's *t* tests were performed to compare richness predictions for future withdrawal scenarios to the 2010 predictions using the R function *t* test. Welch's t tests were performed to test for differences between mean expected values for current conditions and values predicted under future scenarios (water withdrawals, impervious surface projections and climate change). The Welch's modification adjusts the degrees of freedom for predictions whose variances are not equal (Welch, 1947). Significant ($\alpha < .05$) *p*-values indicate that predictions are not equal.

3 | RESULTS

The WaSSI model reasonably captured the magnitude and variability in observed flows at the seven validation sites within the study region (Figure 4, Table 1). Model performance was satisfactory or better at all sites evaluated for classical model fit statistics. Absolute bias in mean flow was satisfactory (15%–25%) at two sites, good (10%–15%) at three sites and very good (<10%) at two sites. The

TABLE 2 Spearman correlations and *p*-values (in parentheses) of predictor variables for the boosted regression tree (BRT) models across the 385 North Carolina Division of Water Resources sample sites. Bolded *p*-values are significant (α < .05). Bolded variables represent those retained in the 4-variable BRT model. Definitions of ecologically relevant streamflow statistics can be found in the streamflow prediction section. FSR, fish species richness

	Basin	AveCMS	MinCMS	MaxCMS	AMJCMS	AugCMS	FloVar	QFRAC_IMP	FSR
Basin		0.07 (.1802)	0.01 (.8556)	0.11 (**.0258**)	0.04 (.4061)	0.02 (.6536)	0.11 (**.0310**)	−0.15 (**.0025**)	0.1 (.0508)
AveCMS			0.96 (**<.0001**)	0.99 (**<.0001**)	0.99 (**<.0001**)	0.94 (**<.0001**)	−0.22 (**<.0001**)	0.05 (.2893)	0.2 (**<.0001**)
MinCMS				0.9 (**<.0001**)	0.96 (**<.0001**)	0.99 (**<.0001**)	−0.42 (**<.0001**)	0.11 (**.0359**)	0.13 (**.0092**)
MaxCMS					0.96 (**<.0001**)	0.88 (**<.0001**)	−0.09 (.0869)	0.01 (.8612)	0.22 (**<.0001**)
AMJCMS						0.94 (**<.0001**)	−0.29 (**<.0001**)	0.03 (.6049)	0.18 (**.0004**)
AugCMS							−0.42 (**<.0001**)	0.13 (**.0128**)	0.12 (**.0147**)
FloVar								−0.21 (**<.0001**)	0.2 (**<.0001**)
QFRAC_IMP									−0.21 (**<.0001**)
FSR									

NSE was satisfactory at three sites (0.50–0.65), good (0.65–0.75) at two sites, and very good (>0.75) at two sites. The mean bias in mean flow across all sites was +6.2% (absolute bias 11.5%) while the mean NSE was 0.68, both statistics reflecting good performance overall. Similarly, bias in the MaxCMS and FloVar statistics used in the 4-variable flow–ecology model was generally within the range of hydrologic uncertainty (<30%). Bias for four of the seven sites was <30% for FloVar and was <30% at all sites for MaxCMS. Sites 2, 3 and 4 FloVar bias was >30%, however site 2 was only marginally outside the range of hydrologic uncertainty at 34%. FloVar bias was greatest for site 4 at 73%. Comments on flow modification in the USGS Gages II database indicate that there are small reservoirs in the headwaters of this catchment (Falcone, 2011; Falcone et al., 2010), possibly reducing FloVar in the observed time series relative to natural conditions simulated by WaSSI by supplementing flows during dry periods (Figure 4). Indeed, FloVar was lowest at site 4 among all sites, thus small absolute differences in FloVar result in large relative differences. Overall, the validation results indicate that our hydrologic modeling approach provides reasonable approximations of flow statistics for flow–ecology modeling that fall within commonly applied bounds of uncertainty and bias.

Observed FSR values in the North Carolina Piedmont ranged from 5.0 to 28.5 across the 385 sites (Figure 1b, Table 3), with an average of 16.4. Among river basins, the TAR, ROA and NEU had the highest average FSR values, whereas the BRD, YAD, CPF and CTB had slightly lower averages (Table 3). By comparison, the highest MaxCMS values were found in the TAR, BRD and NEU while the CTB, ROA, YAD and CPF were slightly lower (Table 3). In addition to having the lowest MaxCMS values, the CPF had the highest QFRAC_IMP values, followed by the NEU, CTB, YAD, BRD, TAR and ROA (Table 3).

The explanatory variables in the reduced 4-variable model (all variables with relative importance <10% removed) consisted of Basin, QFRAC_IMP, MaxCMS and FloVar, with an estimated equivalent $R^2 = 0.50$ (Table 4). Basin was the most influential variable in the model (31.3% relative importance), followed by MaxCMS (23.7%), QFRAC_IMP (23.1%) and FloVar (21.9%) (Figure 3b). All flow variables show distinct relations with the fitted values (Figure 5). Although each variable exhibits some variability, the overall response pattern indicates a negative response between FSR and QFRAC_IMP, and a positive response between FSR with both MaxCMS and FloVar, although this response may not be linear (Figure 5). Further, the PDP for QFRAC_IMP shows a fairly rapid linear decline in FSR at relatively low levels of surface flow coming from impervious surfaces (Figure 5). Conversely, there is a strong increase in FSR between 3 and 7 m^3/s in the PDP for MaxCMS. The interactions between MaxCMS, QFRAC_IMP and FloVar indicate that FSR is highest when MaxCMS and FloVar are high, but QFRAC_IMP is low (Figure 6). Across the ranges of MaxCMS and FloVar values, FSR remains quite low when QFRAC_IMP is high (Figure 6). The slope of the regression for observed FSR values versus those predicted by the BRT model was 0.41 (predicted FSR = 9.73 + 0.41 × observed FSR) with an adjusted R^2 value of 0.48, indicating a relatively good predictive fit with only slight bias across the range of values.

The northeastern part of the Piedmont including the ROA, TAR and NEU river basins had higher FSR values than the rest of the region under the baseline scenario as illustrated in Figures 2 and 10a. Projected climate change by 2041–60 increased FSR by 0.35 species (Table 4) on average ($p = .0042$), ranging from a decrease of 2.2 to an increase of 3.1 (Figure 7b). Projected changes in impervious cover resulted in an insignificant decrease ($p = .1817$) in FSR of 0.16 species across the region on average (Table 4). FSR decreased significantly across the region as water withdrawals increased from 5% to 25% of baseline flows (Table 4; Figures 7d and 8). Under the water withdrawal scenarios, a significant loss in FSR of 0.49 species was predicted with a 15% reduction in flow ($p = .0001$) while a reduction in flow of 25% was predicted to have an average loss of one species ($p < .0001$) (Table 4).

Under each future scenario, some HUC-12s were more likely to experience changes in FSR than the average HUC-12 (Figures 7 and 8). These results indicated that some HUC-12s will lose species

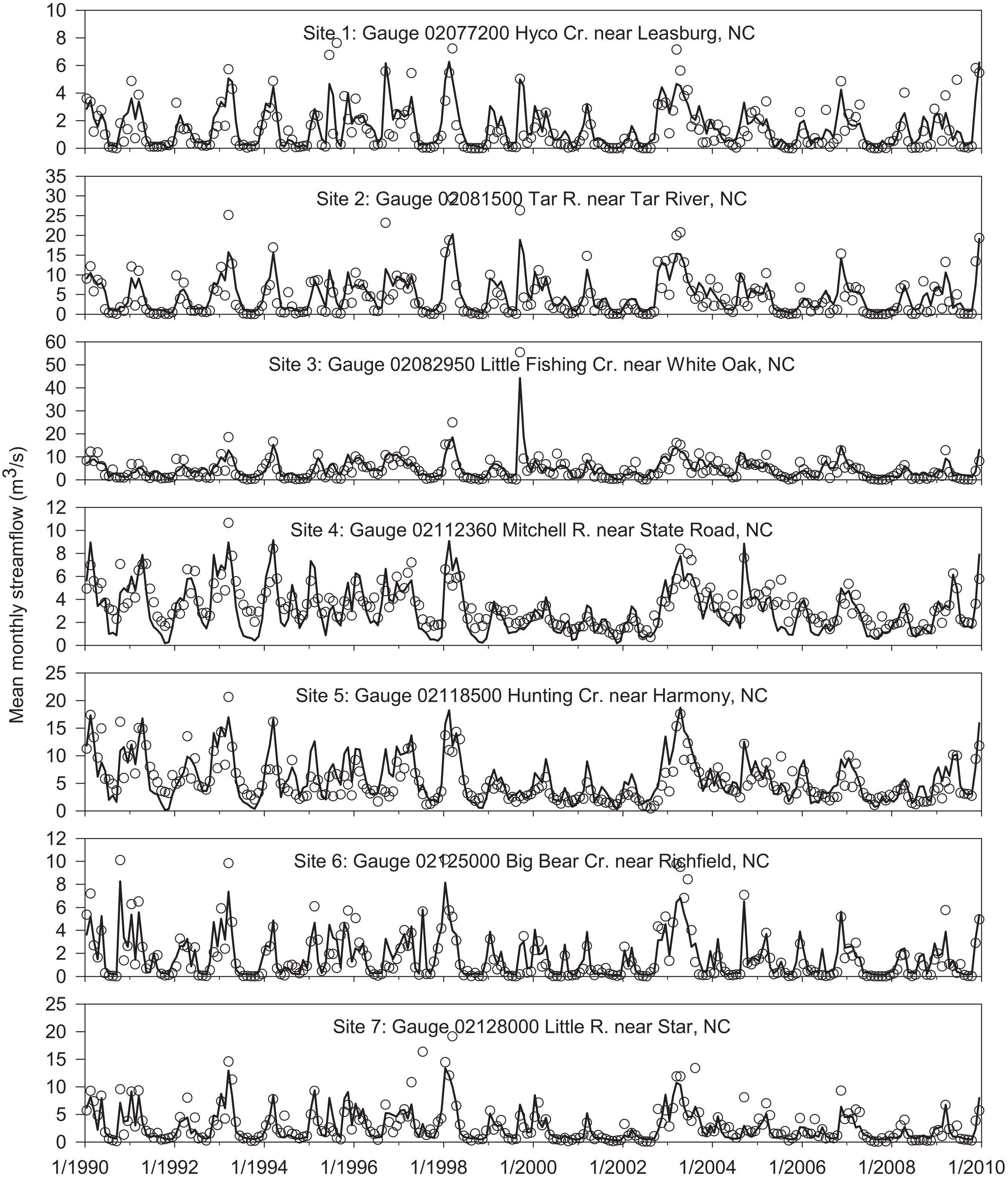

FIGURE 4 Observed (circles) and predicted (lines) mean monthly streamflow hydrographs for the seven USGS reference gauges used for Water Supply Stress Index model validation

more quickly than the region average. The hot-spots shown in Figure 8 (HUC-12s that could potentially lose more than one fish species under the 5%, 10%, 15%, and 20% withdrawal scenarios) identify the catchments that are most vulnerable to a loss in FSR and provide managers with a mechanism for prioritising catchments that are most susceptible to changing water availability. The percent of HUC-12s predicted to decrease in FSR varied across river basins, with some basins (CPF, CTB, TAR and YAD) appearing to

TABLE 3 Predicted mean (standard deviation) flow statistics and fish species richness (FSR) across the 385 North Carolina Division of Water Resources sample sites for the Broad (BRD), Cape Fear (CPF) Catawba (CTB), Neuse (NEU), Roanoke (ROA), Tar (TAR) and Yadkin (YAD) river basins. Definition of ecologically relevant streamflow statistics can be found in the streamflow prediction section

	BRD	CPF	CTB	NEU	ROA	TAR	YAD	All Basins
FSR	15.63 (2.8)	15.48 (4.36)	13.59 (3.63)	18.95 (3.82)	19.36 (3.37)	19.91 (3.65)	15.56 (3.95)	16.35 (4.32)
MaxCMS	2.8 (3.89)	1.43 (1.68)	1.76 (1.15)	2.46 (2.06)	1.79 (1.39)	3.07 (3.68)	1.75 (1.17)	1.95 (2.05)
QFRAC_IMP	0.08 (0.07)	0.21 (0.18)	0.2 (0.17)	0.21 (0.14)	0.05 (0.04)	0.06 (0.05)	0.12 (0.12)	0.14 (0.15)
FloVar	0.75 (0.04)	0.81 (0.12)	0.74 (0.09)	0.78 (0.08)	0.8 (0.1)	0.85 (0.06)	0.83 (0.16)	0.8 (0.12)

TABLE 4 Boosted regression tree predictions for the 4-variable model across the North Carolina Piedmont ecoregion. Welch's *t* tests were performed to test for differences between mean expected values for current conditions and values predicted under future scenarios (water withdrawals, impervious surface projections and climate change). Bolded values represent a significant difference ($p < .05$) between current conditions and predicted values. 2041–60 and IMP 2041–60 represent the time periods for the climate change and impervious surface scenarios, respectively. Percentage values from −5 to −25 represent water withdrawal scenarios

4-Variable model (equivalent R^2 = 0.50)	2041–60	IMP 2041–60	−5%	−10%	−15%	−20%	−25%
t Test *p*-value	**.0042**	.1817	.7909	.1047	**.0001**	**<.0001**	**<.0001**
Mean expected	18.0441	18.0441	18.0441	18.0441	18.0441	18.0441	18.0441
Mean predicted	18.3896	17.8802	18.0119	17.8424	17.5503	17.2323	17.0478
Mean change (predicted − expected)	0.3455	−0.1639	−0.0322	−0.2017	−0.4938	−0.8118	−0.9963

be particularly susceptible to changes in flow (Table 5). Predicted FSR decreased in the majority of HUC-12s under all flow reduction scenarios (Table 5). Even under the climate change scenario, where average FSR across all HUC-12s was predicted to increase, a large percentage (33%) of HUC-12s were predicted to decrease in FSR (Table 5). Further, the climate scenario showed a large increase in mean MaxCMS (3.32–15.28 m^3/s) across all major river basins (Table 6), resulting in a mean increase in FSR. MaxCMS increased in some catchments under the impervious scenario, however QFRAC_IMP also increased (Table 6) resulting in a net decrease in FSR (Table 4).

4 | DISCUSSION

In this study, we evaluated whether relatively simple hydrologic models can be used in conjunction with ecological data to develop empirical flow–ecology response models that predict the effect of changes in water availability on FSR at a spatial scale relevant to management. We also sought to use the empirical flow–ecology models to identify "hot-spots" of fish richness change under plausible scenarios representing changes in water withdrawals (e.g. 5%–25%), land use (derived from known build out scenarios), and climate (the A2 high emission scenario), at the HUC-12 level. We postulated that a decline in predicted FSR would be attributable to changes in climate and increases in impervious surfaces and water withdrawals. Our findings indicate that changes in streamflow associated with plausible future water withdrawals may result in a significant loss in fish species richness, and for the withdrawals scenarios across the region as a whole, losses appear to be directly linked to the quantity of water withdrawn. Although the future impervious scenario was not found to be significant across the region as a whole, decreases in FSR of one or more species were predicted in many HUC12s proximal to the highly urban regions of North Carolina including Raleigh, Durham, Chapel Hill and Charlotte (Figure 7c). Under the climate change scenario, FSR was actually predicted to increase significantly across all HUC-12s. While this was contrary to our hypothesis, there were many individual HUC-12s where FSR was predicted to decrease (Table 5).

The key variable driving the average increase in FSR for the climate scenario appears to be MaxCMS (maximum monthly streamflow for the 20-year period of record). Under this scenario, MaxCMS was the only variable from the 4-variable BRT model to change substantially from the 1981–2010 average. MaxCMS was highly correlated with FSR in the training data, so these results should not be surprising. However, these findings could indicate a link between predicted changes in climate, maximum monthly or seasonal flows in river systems, and increasing FSR. In contrast to the climate scenario, the flow variables for the impervious scenario changed very little from the 1981–2010 average over the entire study area (Table 6), which may help explain why there was no significant changes in average FSR across all HUC12s for that scenario. Increases in impervious surfaces are predicted to occur in and around urban areas (USEPA, 2009), and likely would not impact all HUC12s within a region the way that climate change could. For example, even though increases in MaxCMS are positively correlated with FSR, when QFRAC_IMP is high, FSR tends to be low (Figure 6). Conversely, when QFRAC_IMP is low, FSR tends to be high, especially in larger streams with higher MaxCMS. Some level of interaction is expected among flow attributes that

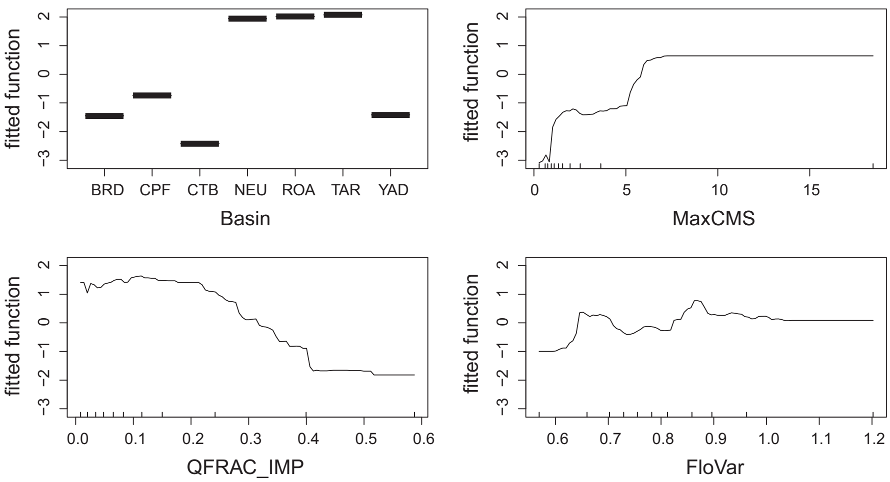

FIGURE 5 Partial dependency plots for variables in the final 4-variable boosted regression tree model for fish species richness (FSR). Boosted regression tree partial dependency plots show the response form of FSR (*y*-axis = fitted function of FSR) based on the effect of individual explanatory variables with the response of all other variables removed

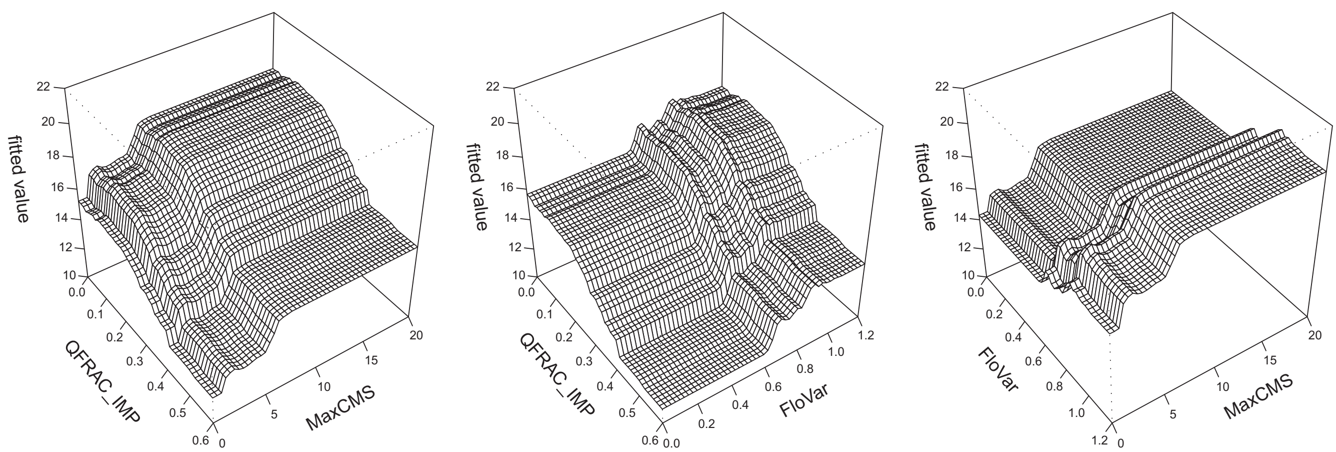

FIGURE 6 Interaction plots of QFRAC_IMP, MaxCMS and FloVar, the three continuous variables in the 4-variable model. The *y*-axis fitted function represents the effect of the interaction on the response variable fish species richness

summarise information across broad spatial scales; however, such findings are essential for supporting decision makers by giving them the tools and information needed to manage water resources when faced with multiple sources of change. For example, limiting water withdrawals in an undeveloped catchment to maintain or enhance FSR may not result in the desired endpoint if land-use change results in increases in impervious cover and thus increases in QFRAC_IMP. Therefore, focusing on a management action that addresses only one streamflow component or that does not take into consideration non-stationarity principles (Milly et al., 2008) could cause an over-estimation of water availability, and result in a significant over-allocation of the resource.

The value of a strong hydrologic foundation cannot be understated for supporting a broader understanding of the connection between changes in water availability and sustaining the long-term viability of fish assemblages. The modeling approach presented in this paper was vital for systematically assessing regional scale effects and identifying areas of concern (i.e. “hot-spots”) where the combined effects of land cover change, climate change and/or streamflow alteration may threaten water resources. Once hot-spots are identified, fine-scale, physically based models of higher temporal resolution could potentially be applied to those areas of concern to provide more quantitative estimates of changes in water availability and support sub-monthly ecologically relevant flow statistics using more site specific inputs.

4.1 | Response of FSR to hydrologic change

Maintenance of hydrologic variability is critical to protecting biodiversity and maintaining the integrity of aquatic, riparian and wetland ecosystems, and is the foundation of the Natural Flow Regime Paradigm (NFRP) presented by Poff et al. (1997). Decades of observation of the effects of human alteration of natural flow regimes have established that streamflow variability is critical for maintaining the ecological integrity of river systems because many aquatic species have developed life-history strategies in response to these flow attributes (Hill, Platts, & Beschta, 1991; Lytle & Poff, 2004; Mims & Olden, 2012, 2013; Poff & Ward, 1989; Postel & Richter, 2012; Richter, Braun, Mendelson, & Master, 1997; Stalnaker, 1990). The coefficient of monthly streamflow variability (FloVar) was one of the important predictors in our model. The PDP plot (Figure 5), before it flattens out, generally indicates a strong positive response between

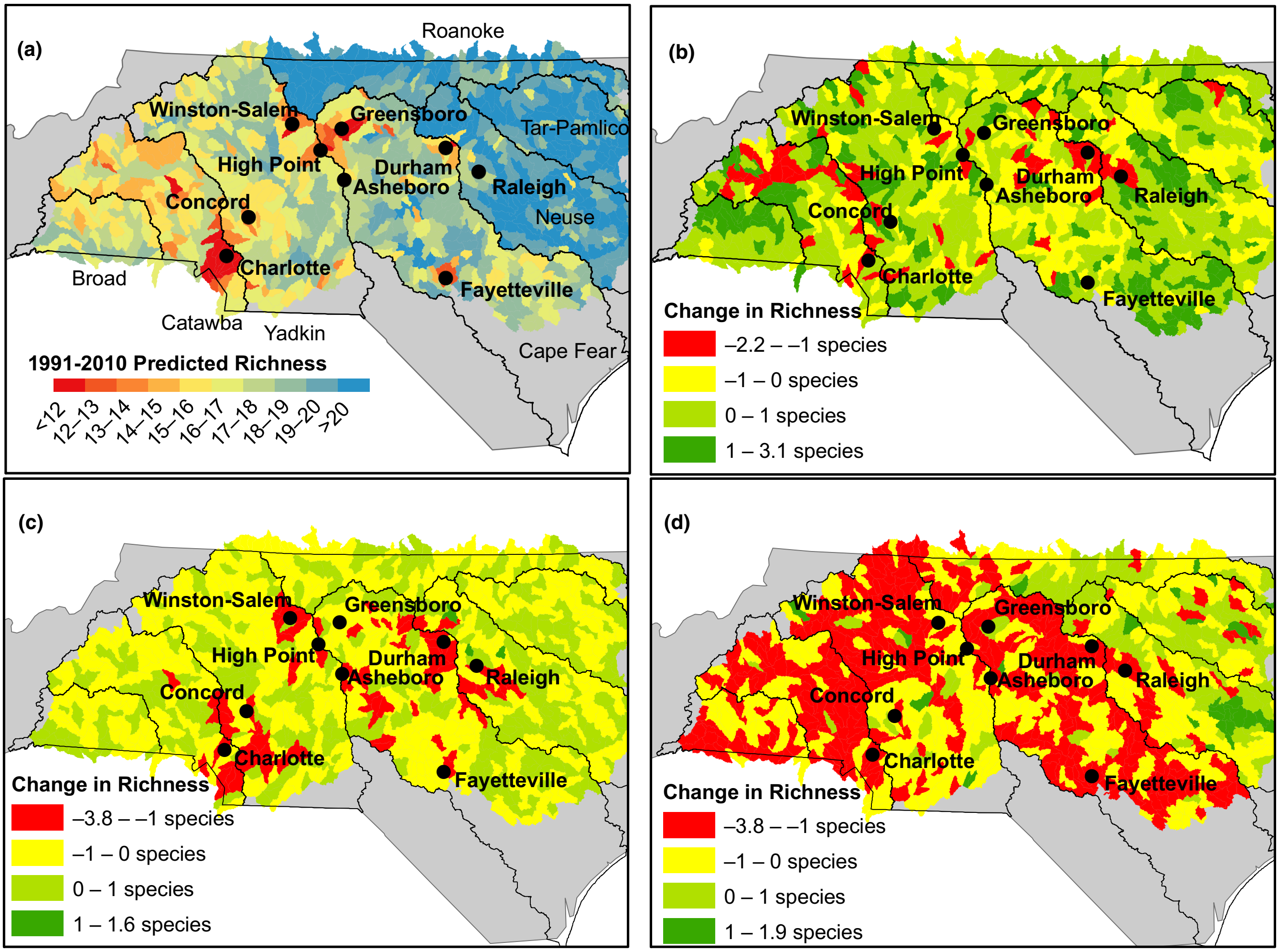

FIGURE 7 Predicted 1991–2010 baseline fish species richness (FSR) values across the study area (a), and change in FSR under climate projections by 2060 (b), 2060 impervious projections (c), and 20% withdrawals (d) using the 4-variable boosted regression tree model

FSR and increasing streamflow variability which is in keeping with the principles of the NFRP. Therefore, the strength of the response for FlowVar in the model underscores the importance of maintaining streamflow variability in support of a thriving fish assemblage (Bunn & Arthington, 2002; Carlisle, Falcone, Wolock, Meador, & Norris, 2010; Poff & Zimmerman, 2010).

Although the goal of restoring streamflow to its "natural" condition may be unachievable in moderately to highly degraded urban systems with high human demand for water or in systems with numerous reservoirs designed for water supply purposes, it still may be possible to offset future alterations in water availability resulting from climate or land-use change by implementing proactive strategies that maintain variable passing flows or flow-by standards that are consistent with NFRP principles. For example, the North Carolina Ecological Flow Science Advisory Board (NCEFSAB), which was tasked with developing a scientifically defensible approach to establishing flows that protect the ecological integrity of streams and rivers in North Carolina as required under Session Law 2010-143, suggested an 80%–90% flow-by (i.e. 80%–90% of ambient modelled flow remains in the stream; NCEFSAB, 2013) in combination with a critical low-flow component. Results of our plausible withdrawal scenarios are highly consistent with the NCEFSAB's recommendations.

4.2 | Modelling limitations

Streamflow metrics predicted by the WaSSI model were subject to similar uncertainties associated with other hydrologic models (Caldwell et al., 2015), including uncertainty in climate, land cover, soil and leaf area index input data, as well as uncertainty in the representation of the physical processes that govern streamflow magnitude and timing. Unlike calibrated models, the WaSSI model will be less sensitive to errors associated with expanding the model domain to catchments not included in the model calibration process, and using the model to assess the effect of climate or land cover scenarios outside of the conditions for which it was calibrated. The overall accuracy of the model was considered satisfactory given the many uncertainties in model inputs, model representation of the physical system, and observed stream flow data (see Caldwell et al., 2015).

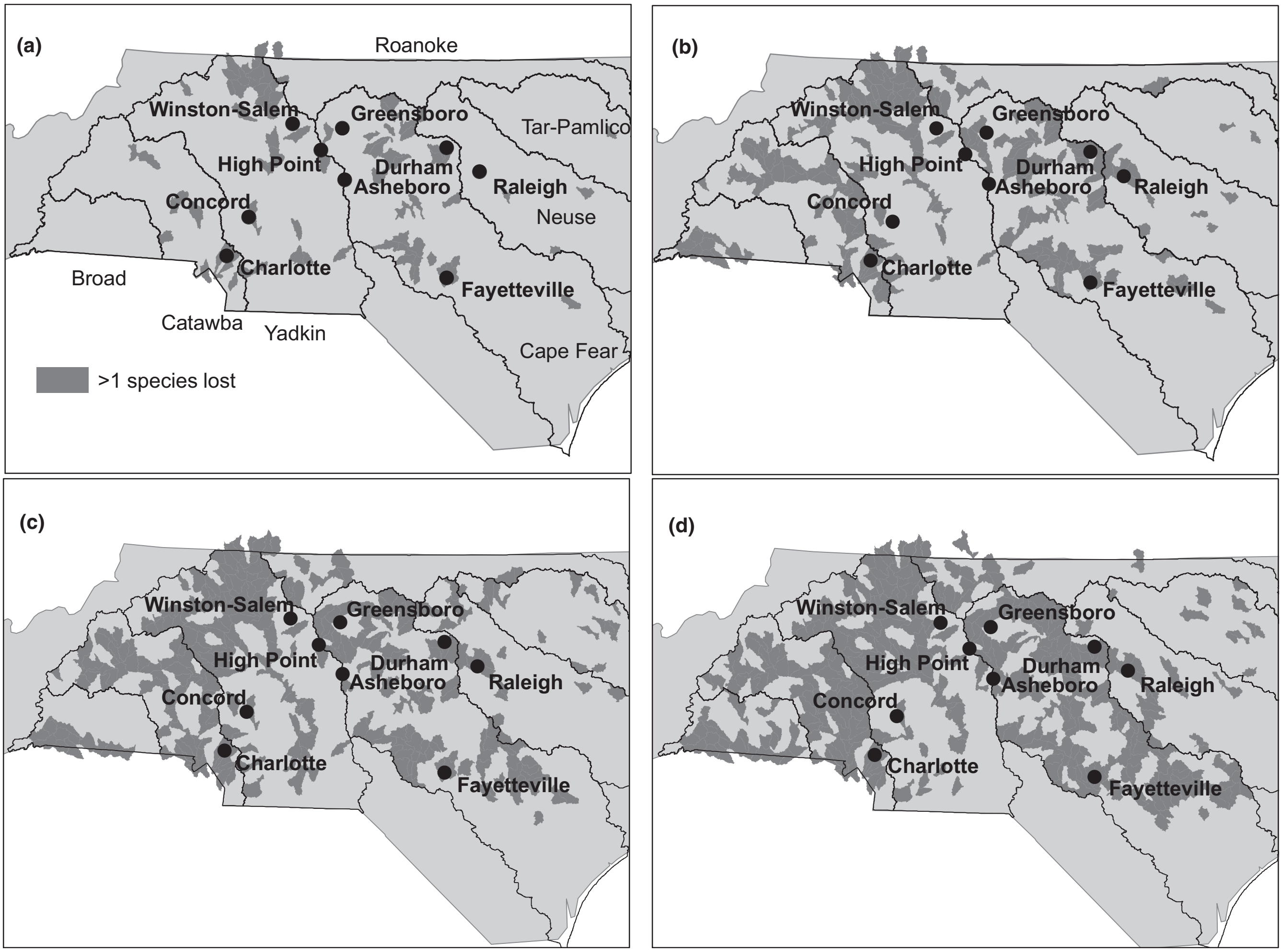

FIGURE 8 HUC-12s in the North Carolina Piedmont predicted to lose greater than 1 fish species based on the 5% (a), 10% (b), 15% (c) and 20% (d) withdrawal scenarios

TABLE 5 Number (percentage) of HUC-12s, by river basin, with a predicted decrease in fish species richness under each scenario. Bolded values represent a significant difference ($p < .05$) between current conditions and predicted values (note that significance was tested for the region as a whole, not per basin). 2041–60 and IMP 2041–60 represent the time periods for the climate change and impervious surface scenarios, respectively. Percentage values from −5 to −25 represent water withdrawal scenarios. Broad (BRD), Cape Fear (CPF) Catawba (CTB), Neuse (NEU), Roanoke (ROA), Tar (TAR) and Yadkin (YAD) river basins

Basin	2041–60	IMP 2041–60	−5%	−10%	−15%	−20%	−25%
BRD (n = 49)	3 (6%)	15 (31%)	13 (27%)	13 (27%)	29 (59%)	49 (100%)	48 (98%)
CPF (n = 215)	87 (40%)	128 (60%)	148 (69%)	151 (70%)	177 (82%)	207 (96%)	211 (98%)
CTB (n = 100)	45 (45%)	50 (50%)	48 (48%)	66 (66%)	87 (87%)	98 (98%)	98 (98%)
NEU (n = 134)	39 (29%)	62 (46%)	50 (37%)	53 (40%)	54 (40%)	81 60%)	109 (81%)
ROA (n = 97)	25 (26%)	36 (37%)	45 (46%)	31 (32%)	40 (41%)	61 63%)	61 (63%)
TAR (n = 85)	33 (39%)	23 (27%)	57 (67%)	48 (56%)	43 (51%)	44 (52%)	59 (69%)
YAD (n = 206)	62 (30%)	95 (46%)	98 (48%)	123 (60%)	148 (72%)	189 (92%)	191 (93%)
All basins (n = 886)	**294 (33%)**	409 (46%)	459 (52%)	485 (55%)	**578 (65%)**	**729 (82%)**	**777 (88%)**

We acknowledge that there is considerable uncertainty in the prediction of future climate and land cover, however the projections we used provided a reasonable scenario of how they may change and this was supported by our model validation results (Table 1).

We were not able to capture some of the more specific submonthly streamflow attributes that may be important for fish migration and reproduction (e.g. annual daily minimum and maximum streamflow, daily streamflow exceedances and recession rates; see

TABLE 6 Average difference between scenarios and 2010 predictions (scenario-2010) for changes in climate and imperviousness in major river basins across the North Carolina Piedmont. Positive values represent a predicted increase under scenarios. Definitions of ecologically relevant streamflow statistics can be found in the streamflow prediction section. Broad (BRD), Cape Fear (CPF) Catawba (CTB), Neuse (NEU), Roanoke (ROA), Tar (TAR) and Yadkin (YAD) river basins

Scenario	Variable	BRD	CPF	CTB	NEU	ROA	TAR	YAD	All basins
Climate	MaxCMS	3.32	6.51	7.82	5.98	15.28	5.18	7.88	7.55
	QFRAC_IMP	−0.01	0.00	−0.01	0.00	0.00	0.00	0.00	0.00
	FloVar	−0.07	0.00	−0.07	−0.03	−0.02	−0.09	−0.05	−0.04
Impervious	MaxCMS	0.01	0.24	0.21	0.51	0.04	0.03	0.18	0.21
	QFRAC_IMP	0.00	0.03	0.03	0.05	0.00	0.00	0.03	0.03
	FloVar	0.00	−0.01	−0.01	−0.02	0.00	0.00	−0.01	−0.01

Kennen et al., 2007; Konrad et al., 2008; Olden & Poff, 2003) because the WaSSI model functions at a monthly time-step. However, even with this limitation we were able to develop a significant 4-variable BRT model that had good predictive power and helped to better understand the potential effects of increasing water withdrawal on FSR in the North Carolina Piedmont region. Hydrologic models vary in their levels of complexity, temporal and spatial resolution, and required level of calibration. Detailed and highly parameterised fine-resolution models such as distributed physically based watershed and rainfall–runoff models are well suited for smaller domains but can be computationally expensive and difficult to parameterise at larger scales. In contrast, simple, easily parameterised models such as monthly water balance models (e.g. WaSSI, the U.S. Geological Survey Monthly Water Balance Model; Hay & McCabe, 2002) are useful for assessing broad implications of streamflow alteration at a large scale and identifying potential water-limited areas, but may have difficulty resolving unique sub-watershed scale physical and ecological processes and associated anthropogenic effects. Leveraging the benefits of both simple, large-scale models with more complex, high resolution models has the potential to allow more robust evaluations of the effects of water withdrawal on aquatic ecosystems. WaSSI, as demonstrated in this paper, can be used in conjunction with biological data to develop flow–ecology models that assess broad-scale regional impacts and identify specific catchments of concern ("hot-spots") where the combined effects of land cover change, climate change and/or flow alteration may threaten water resources.

There are also limitations implicit in flow–ecology models constructed using machine learning techniques such as the BRT model presented in this paper. The strength of BRT models is that they improve on the basic averaging algorithm used in random forest (De'ath, 2007), however, the improvements in prediction accuracy may come at the expense of some loss of interpretation. For example, many of the advanced machine learning techniques, such as BRT, have a tendency to over fit the data (Aertsen et al., 2010). Goodness-of-fit measures and *k*-fold cross-validation techniques, as applied in this study, have been implemented to help practitioners understand and offset this limitation (Elith et al., 2008). However, care should be taken to make sure results are not influenced by spatial sorting bias or spatial autocorrelation (Hijmans, 2012; Randin et al., 2006). A general weakness of BRT models is that they are not as familiar to scientists and managers as modeling methods such as multiple linear regression. Thus, explaining how BRT models work and how to interpret the results in a manner that positively impacts management decisions can be a challenge. The general robustness and greater predictive power of machine learning techniques greatly outweighs their limitations and, as their application becomes more commonplace in ecology, especially for modeling nonlinear relationships, their level of acceptance in the management arena will also increase.

4.3 | Improvements/future work

The use of FSR as the primary measure of fish assemblage integrity as part of this study provides a level of simplicity and parsimony that supports scientific reproducibility and management application at the state and regional level. However, richness is only one measure of assemblage integrity, and alone, it may limit broader interpretation of the hydrologic effects on fish reproduction, life-history processes and species of special concern. Moreover, there is a need to better understand underlying mechanisms (sensu Poff, 2018, this Special Issue) that explain local abundance and regional distributions of fish species. Examining fish species traits is one such method that has been shown to be a powerful tool in ecology for identifying trends within and among species assemblages (Statzner, Hildrew, & Resh, 2001) and represent measurable characteristics based on morphological, physiological or life-history attributes (Violle et al., 2007). Additionally, there is a need to more broadly implement trait-based research in fish ecology, which to date, has largely been focused on terrestrial plants and aquatic invertebrates (Verberk, Van Noordwijk, & Hildrew, 2013). Therefore, results of this study could be enhanced through the applications of functional traits as a means to better understand the effects of hydrologic alteration on fish assemblages and support the conservation of fish species of special concern in the North Carolina Piedmont.

5 | CONCLUSION

In this study, streamflow indices including the maximum monthly streamflow and the coefficient of streamflow variability were shown, in part, to be particularly important for supporting the richness of fish assemblages in the North Carolina Piedmont. The results strongly support other studies that have shown that as the magnitude of high flows and natural variability in annual streamflow is altered, the richness of species with life-history and behavioural constraints that rely on annual high flow patterns or fluctuations in flow for reproduction may be reduced. Implementing water management measures that meet the constraints of the NFRP has been a major challenge for management agencies. Developing practical flow-protection standards that limit ground- and surface-water withdrawals, interbasin transfers, or the implementation of designed flow releases that protect essential streamflow variability, have been difficult to achieve or have been met with strong resistance or legal actions. Therefore, it is essential that management strategies developed in collaboration with stakeholders that minimise flow alteration strive to conserve FSR. Improved water management incentives need to be established within the constraints of existing water law and government statutes that support designated uses, meet existing regulatory requirements and promote a balance between water supply to support human needs and conservation of biological integrity.

ACKNOWLEDGMENTS

We are indebted to Bryn Tracy who worked tirelessly during his long career with the North Carolina Division of Water Resources to collect and curate the fish species dataset which was integral to this project and many others. We thank Jerry McMahon, Director of the Southeast Climate Science Center, for his guidance and support throughout the duration of this project. We also thank David Dudgeon, Eric Stein, and the reviewers for helping to improve this manuscript. Funding for the WaSSI model was provided by the USDA Forest Service Eastern Forest Environmental Threat Assessment Center. Funding for this study was provided by the U.S. Department of Interior, Southeast Climate Science Center and the U.S. Geological Survey National Water Census. The use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

ORCID

Ernie F. Hain http://orcid.org/0000-0003-0492-5900
Jonathan G. Kennen http://orcid.org/0000-0002-5426-4445
Peter V. Caldwell http://orcid.org/0000-0003-0537-3546
Ge Sun http://orcid.org/0000-0002-0159-1370
Steven G. McNulty http://orcid.org/0000-0003-4518-5646

REFERENCES

Acreman, M., Dunbar, M., Hannaford, J., Mountford, O., Wood, P., Holmes, N., . . . Aldrick, J. (2008). Developing environmental standards for abstractions from UK rivers to implement the EU Water Framework Directive. *Hydrological Sciences Journal*, *53*, 1105–1120. https://doi.org/10.1623/hysj.53.6.1105

Aertsen, W., Kint, V., Van Orshoven, J., Özkan, K., & Muys, B. (2010). Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. *Ecological Modelling*, *221*, 1119–1130. https://doi.org/10.1016/j.ecolmodel.2010.01.007

Archfield, S. A., Kennen, J. G., Carlisle, D. M., Wolock, D. M., Kiang, J. E., & Eng, K. (2013). An objective and parsimonious approach for classifying natural flow regimes at a continental scale. *River Research and Applications*, *30*, 1166–1183. https://doi.org/10.1002/rra.2710

Arthington, A. H., Bernardo, J. M., & Ilhéu, M. (2014). Temporary rivers: Linking ecohydrology, ecological quality and reconciliation ecology. *River Research and Applications*, *30*, 1209–1215. https://doi.org/10.1002/rra.2831

Averyt, K., Fisher, J., Huber-Lee, A., Lewis, A., Macknick, J., Madden, N., . . . Tellinghuisen, S. (2011). *Freshwater use by US power plants: Electricity's thirst for a precious resource*. Cambridge, MA: Union of Concerned Scientists, 52 pp.

Averyt, K., Meldrum, J., Caldwell, P., Sun, G., McNulty, S., Huber-Lee, A., & Madden, N. (2013). Sectoral contributions to surface water stress in the coterminous United States. *Environmental Research Letters*, *8*, 035046. https://doi.org/10.1088/1748-9326/8/3/035046

Bierwagen, B. G., Theobald, D. M., Pyke, C. R., Choate, A., Groth, P., Thomas, J. V., & Morefield, P. (2010). National housing and impervious surface scenarios for integrated climate impact assessments. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 20887–20892. https://doi.org/10.1073/pnas.1002096107

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. New York, NY: Chapman & Hall, 358 pp.

Brown, L. R., May, J. T., Rehn, A. C., Ode, P. R., Waite, I. R., & Kennen, J. G. (2012). Predicting biological condition in southern California streams. *Landscape and Urban Planning*, *108*, 17–27. https://doi.org/10.1016/j.landurbplan.2012.07.009

Bunn, S. E., & Arthington, A. H. (2002). Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management*, *30*, 492–507. https://doi.org/10.1007/s00267-002-2737-0

Burnash, R. J. C. (1995). The NWS river forecast system—Catchment modeling. In V. J. Singh (Ed.), *Computer models of watershed hydrology* (pp. 311–366). Littleton, CO: Water Resources Publications.

Burnash, R. J. C., Ferral, R. L., & Mcguire, R. A. (1973). *A generalized streamflow simulation system, conceptual modeling for digital computers*. Sacramento, CA: U.S. Department of Commerce, National Weather Service, and State of California, Department of Water Resources, 204 pp.

Caldwell, P. V., Kennen, J. G., Sun, G., Kiang, J. E., Butcher, J. B., Eddy, M. C., . . . Nelson, S. A. (2015). A comparison of hydrologic models for ecological flows and water availability. *Ecohydrology*, *8*, 1525–1546. https://doi.org/10.1002/eco.1602

Caldwell, P., Sun, G., McNulty, S., Cohen, E., & Myers, J. M. (2012). Impacts of impervious cover, water withdrawals, and climate change on river flows in the conterminous US. *Hydrology and Earth System Sciences*, *16*, 2839–2857. https://doi.org/10.5194/hess-16-2839-2012

Carlisle, D. M., Falcone, J., Wolock, D. M., Meador, M. R., & Norris, R. H. (2010). Predicting the natural flow regime: Models for assessing hydrological alteration in streams. *River Research and Applications*, *26*, 118–136. https://doi.org/10.1002/rra.1247

Chessman, B. C., Jones, H. A., Searle, N. K., Growns, I. O., & Pearson, M. R. (2010). Assessing effects of flow alteration on macroinvertebrate assemblages in Australian dryland rivers. *Freshwater Biology*, *55*, 1780–1800. https://doi.org/10.1111/j.1365-2427.2010.02403.x

Clapcott, J., Young, R., Goodwin, E., Leathwick, J., & Kelly, D. (2011). Relationships between multiple land-use pressures and individual and combined indicators of stream ecological integrity. DOC Research and Development Series 326. Department of Conservation, Wellington. 57 p.

Conroy, M. J., Allen, C. R., Peterson, J. T., Pritchard, Jr, L., & Moore, C. T. (2003). Landscape change in the southern Piedmont: Challenges, solutions, and uncertainty across scales. *Conservation Ecology*, *8*, 3. https://doi.org/10.5751/ES-00598-080203

De'ath, G. (2007). Boosted trees for ecological modeling and prediction. *Ecology*, *88*, 243–251. https://doi.org/10.1890/0012-9658(2007)88[243:BTFEMA]2.0.CO;2

De'ath, G., & Fabricius, K. E. (2000). Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology*, *81*, 3178–3192. https://doi.org/10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., … Sullivan, C. A. (2006). Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Review*, *81*, 163–182. https://doi.org/10.1017/S1464793105006950

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*, 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Falcone, J. A. (2011). *Geospatial attributes of gages for evaluating streamflow. Digital spatial data set*. Retrieved from http://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml

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. *Ecology*, *91*(2), 621. https://doi.org/10.1890/09-0889.1

Freeman, M. C., Buell, G. R., Hay, L. E., Hughes, W. B., Jacobson, R. B., Jones, J. W., … Peterson, J. T. (2013). Linking river management to species conservation using dynamic landscape-scale models. *River Research and Applications*, *29*, 906–918. https://doi.org/10.1002/rra.2575

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Fry, J. A., Xian, G., Jin, S., Dewitz, J. A., Homer, C. G., Limin, Y., … Wickham, J. D. (2011). Completion of the 2006 national land cover database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, *77*, 858–864.

Hamon, W. R. (1963). Computation of direct runoff amounts from storm rainfall. *International Association of Scientific Hydrology*, *Publication*, *63*, 52–62.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer Series in Statistics (2nd ed.). New York, NY: Springer-Verlag, 533 pp.

Hay, L. E., & McCabe, G. J. (2002). Spatial variability in water-balance model performance in the conterminous United States. *Journal of the American Water Resources Association*, *38*, 847–860. https://doi.org/10.1111/j.1752-1688.2002.tb01001.x

Henriksen, J. A., Heasley, J., Kennen, J. G., & Nieswand, S. (2006). *Users' manual for the Hydroecological Integrity Assessment Process software (including the New Jersey Assessment Tools)*. U.S. Geological Survey Open-File Report 2006-1093, 71 pp.

Hijmans, R. J. (2012). Cross-validation of species distribution models: Removing spatial sorting bias and calibration with a null model. *Ecology*, *93*(3), 679–688. https://doi.org/10.1890/11-0826.1

Hill, M. T., Platts, W. S., & Beschta, R. L. (1991). Ecological and geomorphological concepts for instream and out-of-channel flow requirements. *Rivers*, *2*, 198–210.

Intergovernmental Panel on Climate Change (IPCC). (2013). *Climate change 2013: The physical science basis*. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. M. Midgley (Eds.). Cambridge, U.K. and New York, NY: Cambridge University Press, 1535 pp.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. New York, NY: Springer.

Keaton, M., Haney, D., & Andersen, C. B. (2005). Impact of drought upon fish assemblage structure in two South Carolina Piedmont streams. *Hydrobiologia*, *54*, 209–223. https://doi.org/10.1007/s10750-005-2674-z

Kendy, E., Apse, C., & Blann, K. (2012). *A practical guide to environmental flows for policy and planning*. Arlington, VA: The Nature Conservancy, 72 pp. Retrieved from https://www.conservationgateway.org/ConservationPractices/Freshwater/EnvironmentalFlows/MethodsandTools/ELOHA/Documents/Practical%20Guide%20Eflows%20for%20Policy-low%20res.pdf

Kennard, M. J., Pusey, B. J., Olden, J. D., Mackay, S. J., Stein, J. L., & Marsh, N. (2010). Classification of natural flow regimes in Australia to support environmental flow management. *Freshwater Biology*, *55*, 171–193. https://doi.org/10.1111/j.1365-2427.2009.02307.x

Kennen, J. G., Chang, M., & Tracy, B. H. (2005). Effects of landscape change on fish assemblage structure in a rapidly growing metropolitan area in North Carolina, USA. *American Fisheries Society Symposium*, *47*, 39–52.

Kennen, J. G., Henriksen, J. A., Heasley, J., Cade, B. S., & Terrell, J. W. (2009). *Application of the hydroecological integrity assessment process for Missouri streams*. U.S. Geological Survey Open-File Report 2006-1093, 71 pp.

Kennen, J. G., Henriksen, J. A., & Nieswand, S. P. (2007). *Development of the hydroecological integrity assessment process for determining environmental flows for New Jersey streams*. U.S. Geological Survey Scientific Investigations Report 2007-5206, 55 pp.

Kennen, J. G., Riskin, M. L., & Charles, E. G. (2014). Effects of streamflow reductions on aquatic macroinvertebrates: Linking groundwater withdrawals and assemblage response in southern New Jersey streams, USA. *Hydrological Sciences Journal*, *59*, 545–561. https://doi.org/10.1080/02626667.2013.877139

Kennen, J. G., Riva-Murray, K., & Beaulieu, K. M. (2010). Determining hydrologic factors that influence stream macroinvertebrate assemblages in the northeastern US. *Ecohydrology*, *3*, 88–106. https://doi.org/10.1002/eco.99

Konrad, C. P., Brasher, A. M. D., & May, J. T. (2008). Assessing streamflow characteristics as limiting factors on benthic invertebrate assemblages in streams across the western United States. *Freshwater Biol*, *53*, 1983–1998. https://doi.org/10.1111/j.1365-2427.2008.02024.x

Koren, V., Smith, M., & Duan, Q. (2003). Use of a priori parameter estimates in the derivation of spatially consistent parameter sets of rainfall-runoff models. In Q. Duan, H. V. Gupta, S. Sorooshian, A. N. Rosseau, & H. Turcotte (Eds.), *Calibration of watershed models*, Vol. *6* (pp. 239–254). Washington, D. C.: American Geophysical Union. http://onlinelibrary.wiley.com/doi/10.1002/9781118665671.ch18/summary

Leclere, J., Oberdorff, T., Belliard, J., & Leprieur, F. (2011). A comparison of modeling techniques to predict juvenile 0+ fish species occurrences in a large river system. *Ecological Informatics*, *6*, 276–285. https://doi.org/10.1016/j.ecoinf.2011.05.001

Lockaby, G., Nagy, C., Vose, J. M., Ford, C. R., Sun, G., McNulty, S., … Meyers, J. M. (2011). Chapter 13: Water and forests. In: D. N. Wear & J. G. Greis (Eds.), *Southern forest futures project*. Asheville, NC: USDA Forest Service, Southern Research Station, General Technical Report, 85 pp.

Lytle, D. A., & Poff, N. L. (2004). Adaptation to natural flow regimes. *Trends in Ecology & Evolution*, *19*, 94–100. https://doi.org/10.1016/j.tree.2003.10.002

Marion, D. A., Sun, G., Caldwell, P. V., Ford, C. R., Ouyang, Y., Amatya, D. M., … Trettin, C. (2013). Managing forest water quantity and quality under climate change in the Southern US. In J. Vose (Ed.), *Climate change adaptation and mitigation management options*. Boca Raton, FL: CRC Press.

Matthews, W. J., & Marsh-Matthews, E. (2003). Effects of drought on fish across axes of space, time and ecological complexity. *Freshwater Biology*, *48*, 1232–1253. https://doi.org/10.1046/j.1365-2427.2003.01087.x

Maupin, M. A., Kenny, J. F., Hutson, S. S., Lovelace, J. K., Barber, N. L., & Linsey, K. S. (2014). *Estimated use of water in the United States in 2010*. U.S. Geological Survey Circular 1405, 56 pp.

McManamay, R. A., Orth, D. J., Dolloff, C. A., & Frimpong, E. A. (2012). A regional classification of unregulated streamflows: Spatial resolution and hierarchical frameworks. *River Research and Applications*, *28*, 1019–1033. https://doi.org/10.1002/rra.v28.7

McManamay, R. A., Orth, D. J., Dolloff, C. A., & Mathews, D. C. (2013). Application of the ELOHA framework to regulated rivers in the Upper Tennessee River Basin: A case study. *Environmental Management*, *51*, 1210–1235. https://doi.org/10.1007/s00267-013-0055-3

Meehl, G., Covey, C., Delworth, T., Latif, M., Mcavaney, B., Mitchell, J., ... Taylor, K. (2007). The WCRP CMIP3 multi-model dataset: A new era in climate change research. *Bulletin of the American Meteorological Society*, *88*, 1383–1394. https://doi.org/10.1175/BAMS-88-9-1383

Melillo, J. M., Richmond, T. C., & Yohe, G. W. (Eds.). (2014). *Climate change impacts in the United States: The third national climate assessment*. Washington, D.C.: U.S. Global Change Research Program, 841 pp. https://doi.org/10.7930/j0z31wj2

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Climate change—Stationarity is dead: Whither water management? *Science*, *319*, 573–574. https://doi.org/10.1126/science.1151915

Mims, M. C., & Olden, J. D. (2012). Life history theory predicts fish assemblage response to hydrologic regimes. *Ecology*, *93*, 35–45. https://doi.org/10.1890/11-0370.1

Mims, M. C., & Olden, J. D. (2013). Fish assemblages respond to altered flow regimes via ecological filtering of life history strategies. *Freshwater Biology*, *58*, 50–62. https://doi.org/10.1111/fwb.12037

Monk, W. A., Wood, P. J., Hannah, D. M., & Wilson, D. A. (2007). Selection of river flow indices for the assessment of hydroecological change. *River Research and Applications*, *23*, 113–122. https://doi.org/10.1002/rra.964

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the American Society of Agricultural and Biological Engineers*, *50*(3), 885–900. https://doi.org/10.13031/2013.23153

Murphy, J. C., Knight, R. R., Wolfe, W. J., & Gain, W. S. (2013). Predicting ecological flow regime at ungaged sites: A comparison of methods. *River Research and Applications*, *29*, 660–669. https://doi.org/10.1002/rra.2570

Naiman, R. J., Latterell, J. J., Pettit, N. E., & Olden, J. D. (2008). Flow variability and the biophysical vitality of river systems. *Comptes Rendus Geoscience*, *340*, 629–643. https://doi.org/10.1016/j.crte.2008.01.002

Nakicenovic, N., Alcamo, J., Davis, G., De Vries, B., Fenhann, J., Gaffin, S., ... Kram, T. (2000). *Special report on emissions scenarios: A special report of Working Group III of the Intergovernmental Panel on Climate Change*. Richland, WA: Pacific Northwest National Laboratory, Environmental Molecular Sciences Laboratory.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models, I, A discussion of principles. *Journal of Hydrology*, *10*, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

NC Climate Office. (2016). *Normal monthly precipitation totals based on 1971–2000 normals*. Retrieved from http://climate.ncsu.edu/climate/monthlyprecip.html

NC Floodplain Mapping Program. (2013). *Digital elevation model of North Carolina*. Retrieved from http://www.ncfloodmaps.com/

NCDENR. (2006). *Standard operating procedure biological monitoring, stream fish community assessment program*. North Carolina Department of Environment and Natural Resources, Division of Water Quality, Version 4, 51 pp, Published in Raleigh, NC.

NCEFSAB. (2013). *Recommendations for estimating flows to maintain ecological integrity in streams and rivers in North Carolina*. Final report submitted to the North Carolina Department of Environment and Natural Resources by the North Carolina Ecological Flows Science Advisory Board, November 2013, 52 pp. Retrieved from http://www.ncwater.org/files/eflows/sab/EFSAB_Final_Report_to_NCDENR.pdf

NRCS, United States Department of Agriculture. (2012). *U.S. General Soil Map (STATSGO2)*. Retrieved from https://catalog.data.gov/dataset/u-s-general-soil-map-statsgo2-for-the-united-states-of-america

Olden, J. D. (2016). Challenges and opportunities for fish conservation in dam-impacted waters. In G. P. Closs, M. Krkosek, & J. D. Olden (Eds.), *Conservation of freshwater fishes* (pp. 107–148). Cambridge, U.K.: Cambridge University Press.

Olden, J. D., & Poff, N. L. (2003). Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. *River Research and Applications*, *19*, 101–121. https://doi.org/10.1002/rra.700

Piñeiro, G., Perelman, S., Guerschman, J. P., & Paruelo, J. M. (2008). How to evaluate models: Observed vs. predicted or predicted vs. observed? *Ecological Modelling*, *216*, 316–322. https://doi.org/10.1016/j.ecolmodel.2008.05.006

Poff, N. L. (2018). Beyond the natural flow regime? Broadening the hydro-ecological foundation to meet environmental flows challenges in a non-stationary world. *Freshwater Biology*, https://doi.org/10.1111/fwb.13038.

Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., ... Stromberg, J. C. (1997). The natural flow regime—A paradigm for conservation and restoration of river ecosystems. *BioScience*, *47*, 769–784. https://doi.org/10.2307/1313099

Poff, N. L., Richter, B. D., Arthington, A. H., Bunn, S. E., Naiman, R. J., Kendy, E., ... Warner, A. T. (2010). The ecological limits of hydrologic alteration (ELOHA): A new framework for developing regional environmental flow standards. *Freshwater Biology*, *55*, 147–170. https://doi.org/10.1111/j.1365-2427.2009.02204.x

Poff, N. L., & Ward, J. V. (1989). Implications of streamflow variability and predictability for lotic community structure: A regional analysis of streamflow patterns. *Canadian Journal of Fisheries and Aquatic Sciences*, *46*, 1805–1818. https://doi.org/10.1139/f89-228

Poff, N. L. (2018). Beyond the natural flow regime? Broadening the hydro-ecological foundation to meet environmental flows challenges in a non-stationary world. *Freshwater Biology*. https://doi.org/10.1111/fwb.13038

Postel, S., & Richter, B. (2012). *Rivers for life: Managing water for people and nature*. Island Press, Washington, D.C., 253 pp.

Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems*, *9*, 181–199. https://doi.org/10.1007/s10021-005-0054-1

PRISM Climate Group. 2013. PRISM Climate Data. Available at: http://www.prism.oregonstate.edu/ [Accessed 01 July 2013].

Randin, C. F., Dirnbock, T., Dullinger, S., Zimmermann, N. E., Zappa, M., & Guisan, A. (2006). Are niche-based species distribution models transferable in space? *Journal of Biogeography*, *33*, 1689–1703. https://doi.org/10.1111/jbi.2006.33.issue-10

Raupach, M. R., Marland, G., Ciais, P., Le Quéré, C., Canadell, J. G., Klepper, G., & Field, C. B. (2007). Global and regional drivers of accelerating CO_2 emissions. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 10288–10293. https://doi.org/10.1073/pnas.0700609104

Richter, B. D., Braun, D. P., Mendelson, M. A., & Master, L. L. (1997). Threats to imperiled freshwater fauna. *Conservation Biology*, *11*, 1081–1093. https://doi.org/10.1046/j.1523-1739.1997.96236.x

Seager, D., Tzanova, A., & Nakamura, J. (2009). Drought in the southeastern United States, variability over the last millennium, and the potential for future hydroclimate change. *Journal of Climate*, *22*, 5021–5045. https://doi.org/10.1175/2009JCLI2683.1

Shenton, W., Bond, N. R., Yen, J. D., & Mac Nally, R. (2012). Putting the "ecology" into environmental flows: Ecological dynamics and demographic modelling. *Environmental Management*, *50*, 1–10. https://doi.org/10.1007/s00267-012-9864-z

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., . . . Zehe, E. (2003). IAHS decade on predictions in ungauged basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal*, *48*(6), 857–880. https://doi.org/10.1623/hysj.48.6.857.51421

Stalnaker, C. B. (1990). Minimum flow is a myth. In Bain, M.B. (ed.) *Ecology and assessment of warmwater streams: Workshop synopsis* (pp. 31–33). Washington, D.C.: US Fish & Wildlife Service. Biological Report.

Statzner, B., Hildrew, A. G., & Resh, V. H. (2001). Species traits and environmental constraints: Entomological research and the history of ecological theory. *Annual Review of Entomology*, *46*, 291–316. https://doi.org/10.1146/annurev.ento.46.1.291

Stewart-Koster, B., Olden, J. D., & Gido, K. B. (2014). Quantifying flow–ecology relationships with functional linear models. *Hydrological Sciences Journal*, *59*, 629–644. https://doi.org/10.1080/02626667.2013.860231

Sun, G., Caldwell, P. V., Georgakakos, A. P., Arumugam, S., Cruise, J., McNider, R. T., . . . Marion, D. A. (2013). Impacts of climate change and variability on water resources in the Southeastern US. In K. T. Ingram, K. Dow, & L. Carter (Eds.), *Southeast regional technical report to the national climate change assessment, water resources*. Washington, D.C.: Island Press.

Sun, G., Caldwell, P., Noormets, A., McNulty, S. G., Cohen, E., Moore Myers, J., . . . Xiao, J. (2011). Upscaling key ecosystem functions across the conterminous United States by a water-centric ecosystem model. *Journal of Geophysical Research: Biogeosciences*, *116*, G00J05. https://doi.org/10.1029/2010JG001573

Tavernia, B. G., Nelson, M. D., Caldwell, P., & Sun, G. (2013). Water stress projections for the northeastern and Midwestern United States in 2060: Anthropogenic and ecological consequences. *Journal of the American Water Resources Association*, *49*, 938–952. https://doi.org/10.1111/jawr.12075

Treasure, E., McNulty, S., Moore Myers, J., & Jennings, L. N. (2014). *Template for assessing climate change impacts and management options: TACCIMO user guide version 2.2. Gen. Tech. Rep. SRS-GTR-186* (p. 33). Asheville, NC: USDA-Forest Service, Southern Research Station.

Turner, M., & Stewardson, M. (2014). Hydrologic indicators of hydraulic conditions that drive flow–biota relationships. *Hydrological Sciences Journal*, *59*, 659–672. https://doi.org/10.1080/02626667.2014.896997

US Census Bureau. (2016). *United States census 2010*. Retrieved from http://www.census.gov/2010census/

US Environmental Protection Agency (USEPA) (2009). *Land-use scenarios: National-scale housing-density scenarios consistent with climate change storylines, global change research program, EPA/600/R-08/076F*. Washington, D.C.: National Center for Environmental Assessment.

Verberk, W., Van Noordwijk, C., & Hildrew, A. (2013). Delivering on a promise: Integrating species traits to transform descriptive community ecology into a predictive science. *Freshwater Science*, *32*(2), 531–547. https://doi.org/10.1899/12-092.1

Violle, C., Navas, M. L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. (2007). Let the concept of trait be functional!. *Oikos*, *116*, 882–892. https://doi.org/10.1111/j.0030-1299.2007.15559.x

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., . . . Davies, P. M. (2010). Global threats to human water security and river biodiversity. *Nature*, *467*, 555–561. https://doi.org/10.1038/nature09440

Waite, I. R., Brown, L. R., Kennen, J. G., May, J. T., Cuffney, T. F., Orlando, J. L., & Jones, K. A. (2010). Comparison of watershed disturbance predictive models for stream benthic macroinvertebrates for three distinct ecoregions in western US. *Ecological Indicators*, *10*, 1125–1136. https://doi.org/10.1016/j.ecolind.2010.03.011

Waite, I. R., Kennen, J. G., May, J. T., Brown, L. R., Cuffney, T. F., Jones, K. A., & Orlando, J. L. (2012). Comparison of stream invertebrate response models for bioassessment metrics. *PLoS ONE*, *9*(3), e90944. https://doi.org/10.1371/journal.pone.0090944

Warren, Jr, M. L., Burr, B. M., Walsh, S. J., Bart, Jr, H. L., Cashner, R. C., Etnier, D. A., . . . Starnes, W. C. (2000). Diversity, distribution, and conservation status of the native freshwater fishes of the southern United States. *Fisheries*, *25*, 7–29. https://doi.org/10.1577/1548-8446(2000)025<0007:DDACSO>2.0.CO;2

Welch, B. L. (1947). The generalization of student's' problem when several different population variances are involved. *Biometrika*, *34*, 28–35. https://doi.org/10.2307/2332510

White, R. S. A., Mchugh, P. A., & Mcintosh, A. R. (2016). Drought survival is a threshold function of habitat size and population density in a fish metapopulation. *Global Change Biology*, *22*, 3341–3348. https://doi.org/10.1111/gcb.13265

Worrall, T. P., Dunbar, M. J., Extence, C. A., Laize, C. L., Monk, W. A., & Wood, P. J. (2014). The identification of hydrological indices for the characterization of macroinvertebrate community response to flow regime variability. *Hydrological Sciences Journal*, *59*, 645–658. https://doi.org/10.1080/02626667.2013.825722

Zhao, M., Heinsch, F. A., Nemani, R. R., & Running, S. W. (2005). Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment*, *95*, 164–176. https://doi.org/10.1016/j.rse.2004.12.011

How to cite this article: Hain EF, Kennen JG, Caldwell PV, Nelson SAC, Sun G, McNulty SG. Using regional scale flow–ecology modeling to identify catchments where fish assemblages are most vulnerable to changes in water availability. *Freshwater Biol*. 2018;63:928–945. https://doi.org/10.1111/fwb.13048