Hydrobiologia
DOI 10.1007/s10750-010-0362-0

PRIMARY RESEARCH PAPER

Determination of biologically significant hydrologic condition metrics in urbanizing watersheds: an empirical analysis over a range of environmental settings

Jeffrey J. Steuer · Krista A. Stensvold · Mark B. Gregory

Received: 4 June 2009 / Revised: 16 June 2010 / Accepted: 19 June 2010
© US Government: United States Geological Survey 2010

Abstract We investigated the relations among 83 hydrologic condition metrics (HCMs) and changes in algal, invertebrate, and fish communities in five metropolitan areas across the continental United States. We used a statistical approach that employed Spearman correlation and regression tree analysis to identify five HCMs that are strongly associated with observed biological variation along a gradient of urbanization. The HCMs related to average flow magnitude, high-flow magnitude, high-flow event frequency, high-flow duration, and rate of change of stream cross-sectional area were most consistently associated with changes in aquatic communities. Although our investigation used an urban gradient design with short hydrologic periods of record ($\leq$1 year) of hourly cross-sectional area time series, these five HCMs were consistent with previous investigations using long-term daily-flow records. The ecological sampling day often was included in the hydrologic period. Regression tree models explained up to 73, 92, and 79% of variance for specific algal, invertebrate, and fish community metrics, respectively. National models generally were not as statistically significant as models for individual metropolitan areas. High-flow event frequency, a hydrologic metric found to be transferable across stream type and useful for classifying habitat by previous research, was found to be the most ecologically relevant HCM; transformation by precipitation increased national-scale applicability. We also investigated the relation between measures of stream flashiness and land-cover indicators of urbanization and found that land-cover characteristic and pattern variables, such as road density, percent wetland, and proximity of developed land, were strongly related to HCMs at both a metropolitan and national scale and, therefore, may be effective land-use management options in addition to wholesale impervious-area reduction.

Keywords Urbanization · Hydrologic condition metrics · Algae · Macroinvertebrates · Fish · Regression tree models · Predictive relations

Electronic supplementary material The online version of this article (doi:10.1007/s10750-010-0362-0) contains supplementary material, which is available to authorized users.

Handling editor: Richard H. Norris

J. J. Steuer (✉) · K. A. Stensvold
U. S. Geological Survey, Wisconsin Water Science Center, 8505 Research Way, Madison, WI 53562, USA
e-mail: jjsteuer@wisc.edu

M. B. Gregory
U. S. Geological Survey, Georgia Water Science Center, 3039 Amwiler Road, Suite 130, Atlanta, GA 30360, USA

Introduction

It has been widely recognized that hydrologic alterations resulting from urbanization affect aquatic ecological conditions. Removal of riparian vegetation, increased impervious surfaces and soil compaction,

channel modifications, and changes in basin drainage characteristics (ditches and piping), which typically occur with urban development, incrementally alter stream hydrology, instream hydraulics, and the aquatic ecosystem (Booth & Jackson, 1997; Paul & Meyer, 2001; Konrad & Booth, 2005; Roy et al., 2005). Streamflow data have provided an important conceptual and empirical link between watershed and instream habitat and the resulting biological communities.

A wide variety of metrics have been proposed to describe the hydrologic regime and ecologically relevant streamflow components that generally include measures of average flow, high-flow magnitude, frequency of flow extremes, duration of flow extremes, and rate of change. Olden & Poff (2003) proposed more than 170 metrics, including the indicators of hydrologic alteration (Richter et al., 1996) that are composed of at least 64 metrics that quantify flow regime change; Clausen & Biggs (2000) reported on 34 hydrologic variables that describe flow conditions, variability, floods, and low-flow components; and McMahon et al. (2003) described 23 stage-based metrics that relate to urban development. Konrad & Booth (2005) identified three types of hydrologic changes of ecological significance that are likely to result from urbanization—increased frequency of high flows, redistribution of water from base flow to flashier stormflow, and increased daily variation in streamflow. The metrics describe overall patterns of the flow regime which, in combination with the physical setting, determined the physical processes that directly affect aquatic biota. Many of these hydrologic metrics have been shown to be useful in relating the effects of urbanization to ecological measures (Poff & Allan, 1995; Clausen & Biggs, 2000; Booth et al., 2004; Cassin et al., 2005; Konrad & Booth, 2005). With the large number of available hydrologic metrics in use today, there remains a need to identify a minimum number of easily computed hydrologic metrics that are biologically relevant and transferable across varying climatic, geologic, and land-use settings.

Extensive study has shown that composition of overall land cover in a watershed can account for much of the variability in stream ecological conditions; however, few investigations have addressed the question of how the configuration of land cover, including fragmentation and the land cover in riparian areas, influences ecological conditions. For example, clustering development to minimize impervious surface cover (Arnold & Gibbons 1996; Booth et al., 2002) and disconnecting drainage systems to allow source infiltration of stormwater (Walsh et al., 2005b) have been hypothesized to result in less impact on the aquatic communities than conventional development. It is important to acknowledge that even small transformations of the natural landscape may result in impacts to streams, and maintenance of healthy biotic assemblages may require limitation of impervious surface cover (Booth et al., 2002). At a time when government is expending substantial resources on storm water management and development on the urban fringe is inevitable, planners must look at designing urban land in ways that minimize the impacts on ecological systems by mimicking natural processes or allowing natural processes to occur (Lloyd et al., 2002). Understanding the effects of urban land use and cover (hereafter referred to as land cover) on the hydrologic condition of streams is an important component of connecting urban development to its ecological effects on stream aquatic communities. Especially important are land-cover variables amenable to management actions.

As part of the U.S. Geological Survey's National Water-Quality Assessment (NAWQA) Program, the effects of urbanization on stream ecosystems (EUSE) have been studied in nine major metropolitan areas across the continental United States: Boston, MA (BOS); Raleigh, NC (RAL); Atlanta, GA (ATL); Birmingham, AL (BIR); Milwaukee-Green Bay, WI (MGB); Denver, CO (DEN); Dallas-Fort Worth, TX (DFW); Salt Lake City, UT (SLC); and Portland, OR (POR; Fig. 1). A rural to urban land-use gradient approach was used in the study areas that differ in potential natural vegetation, temperature, precipitation, basin relief, elevation, and basin slope (Table 1). In contrast to many previous investigations, which used long-term daily streamflow data, our urban gradient design study necessitated the use of short-term (approximately 1 year or less) hourly cross-section hydrologic data. We initially examined all nine metropolitan areas, which resulted in a final focus on the relation between hydrologic condition metrics (HCMs) and response measures of instream biota in five of the nine EUSE metropolitan areas (ATL, RAL, MGB, DFW, and POR) that had comparable measures of HCMs and where the influences of major hydrologic disturbances (e.g., dams, diversions, and transfers)

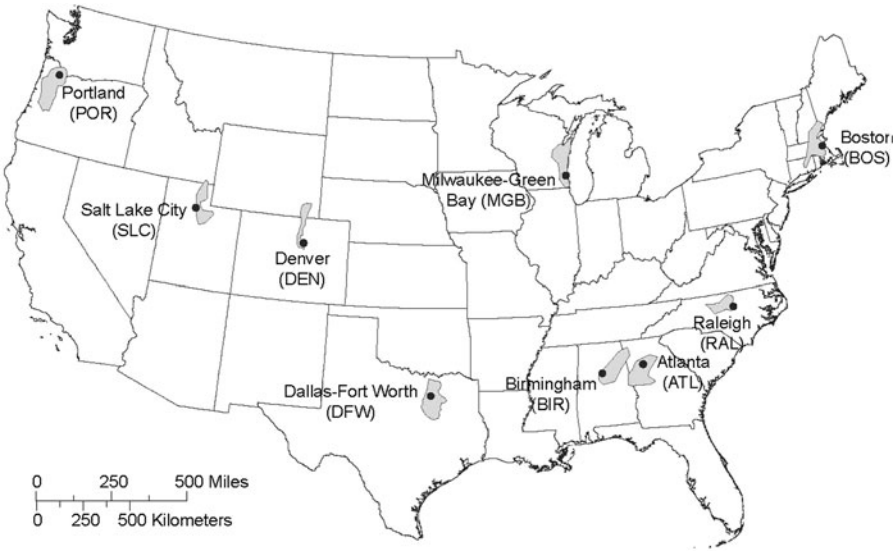


Fig. 1 Location of nine metropolitan areas in the United States sampled as part of the U.S. Geological Survey National Water-Quality Assessment Program study on the effects of urbanization on stream ecosystems, 2000–2004

were relatively small. We also examined these relations in a national-scale data set that combined data from the five metropolitan areas. The purpose of this study was to address two questions. First, is there a subset of HCMs that can predict algal, invertebrate, and fish community richness and abundance variables along a gradient of urbanization in a multiregional setting? Second, what are the urban characteristics (level, composition, and configuration of land cover) that alter the key HCMs?

Methods

Study areas

Nationally available geographic information system (GIS)-derived data were used to define populations of candidate basins within each of the nine metropolitan areas based on 30-m digital elevation models (Falcone et al., 2007). Typically, these basins drained second- to third-order streams and generally were less than 50 km^2. The effects of natural environmental variability were minimized within each metropolitan area by dividing candidate basins into groups with relatively homogeneous natural environmental features (e.g., ecoregion, climate, elevation, and stream size). Urban intensity (MA_NUII) for each candidate basin was defined by combining housing density, percentage of basin area in developed land cover, and road density into an index of urban intensity that ranged from 0 (little or no urban development) to 100 (maximum urban development) within each metropolitan area (Cuffney & Falcone, 2009). The gradient of urbanization in each metropolitan area was represented by using 22–30 study watersheds. This spatially distributed sampling network was intended to represent changes in urbanization through time (i.e., substitute space for time) and allowed the use of HCMs derived from short-term hydrologic data sets.

Fieldwork in the BOS, BIR, and SLC metropolitan areas was conducted in 2000–2001; in ATL, DEN, and RAL in 2002–2003; and in DFW, MGB, and POR in 2003–2004 (Fig. 2). Details about the study designs and landscape characterizations can be found in Coles et al. (2004) and Tate et al. (2005). Basin characteristics were derived for each candidate basin by using a GIS with nationally available spatial data on natural environmental, census, infrastructure, and land-cover characteristics. Falcone et al. (2007) provide a detailed description of the method used to collect the GIS characteristics. A comprehensive presentation of the data used in this study is available in Giddings et al. (2009).

Hydrologic data

Continuous stream-stage data (±3.6 cm) were collected using non-vented, submersible pressure transducers with a 1-h time step. Collection of the hydrologic records began in November 2002 (ATL) and was completed in December 2004 (POR, Fig. 2B).

Table 1 Physiographic region, precipitation occurrence, noteworthy conditions during sampling, and summary statistics for the selected physical variables for the nine studied metropolitan areas in the United States

Metropolitan area (abbreviation)	Physiographic provinces	Precipitation occurrence (Falcone et al., 2007)	Noteworthy conditions during hydrologic data collection	Reach water surface gradient (%) [mean (min–max)]	Mean annual precipitation for period 1980–1997 (cm) [mean (min–max)]	Drainage area (km^2) [mean (min–max)]	Mean elevation (m) [mean (min–max)]	Stage or area calculated HCM	Number of basins in metropolitan area
Portland (POR)	Pacific Mountain, Pacific Border, Cascade-Sierra Mountains	Snowmelt in spring, rainfall October–April	None	0.60 (0.11–1.7)	154 (116–204)	49.2 (17.1–103.8)	222 (53–621)	Area	23
Salt Lake City (SLC)	Rocky Mountain System, Middle Rocky Mountains; Intermontane Plateaus, Basin and Range	Snowmelt in winter/spring major runoff source, scattered light thunderstorms in summer	Flow affected by reservoirs, interbasin transfers, and diversions	1.1 (0.46–2.6)	65 (60–73)	364.5 (8.8–1714.4)	1435 (1370–1504)	Stage	9
Denver (DEN)	Interior Plains, Great Plains; Rocky Mountain System, Southern Rocky Mountains	Snowmelt in spring, rainfall April–September	Flow affected by reservoirs, interbasin transfers, and diversions	Not measured	43 (38–46)	89.8 (4.4–558.6)	1713 (1535–2024)	Area	25
Dallas (DFW)	Atlantic Plain, Coastal Plain	Rainfall spring and summer	None	0.19 (0.008–0.98)	104 (96–111)	85.3 (26.8–291.4)	170 (124–220)	Area	25
Milwaukee (MIL)	Interior Plains, Central Lowland	Snowmelt March–May, rainfall May–September	None	0.31 (0.02–0.63)	86 (81–91)	45.5 (16.3–100.3)	237 (202–270)	Area	22
Birmingham (BIR)	Appalachian Highlands; Valley and Ridge, Plateaus, Piedmont	Rainfall–frontal systems in winter, thunderstorms in summer/fall	Severe drought	0.38 (0.14–0.66)	146 (141–152)	31.8 (4.7–66.1)	233 (162–324)	Stage	23
Atlanta (ATL)	Appalachian Highlands, Piedmont	Rainfall -frontal systems in winter, thunderstorms in summer/fall	Wet year	0.19 (0.04–0.59)	134 (124–140)	89.9 (43.2–146.3)	286 (193–350)	Area	26

Table 1 continued

Metropolitan area (abbreviation)	Physiographic provinces	Precipitation occurrence (Falcone et al., 2007)	Noteworthy conditions during hydrologic data collection	Reach water surface gradient (%) [mean (min–max)]	Mean annual precipitation for period 1980–1997 (cm) [mean (min–max)]	Drainage area (km^2) [mean (min–max)]	Mean elevation (m) [mean (min–max)]	Stage or area calculated HCM	Number of basins in metropolitan area
Raleigh (RAL)	Appalachian Highlands, Piedmont	Rainfall distributed through out the	Wet year	0.63 (0.26–1.3)	120 (115–125)	22.3 (4.9–82.5)	184 (89–284)	Area	25
Boston (BOS)	Appalachian Highlands, New England	Snowmelt March–May, rainfall May–September	None	0.64 (0.12–1.6)	123 (115–136)	75.7 (45.9–124.6)	113 (31–236)	Stage	30

In six of the nine metropolitan areas (POR, DEN, DFW, MGB, ATL, and RAL), a surveyed channel cross-section at the location of the transducer was used to calculate cross-sectional area for each stage value, and this value was converted to a continuous hourly record of cross-sectional area (Table 1). Continuous discharge was not computed because limited discharge measurements during the short data-collection period (<1 year) precluded construction of a stage–discharge rating curve. The POR, DEN, DFW, MGB, and BOS metropolitan areas had periods of record that were less than a full year (Fig. 2B). Study basins with more than 20% missing data were dropped from analysis. Cross-sectional area-based (POR, DEN, DFW, MGB, ATL, and RAL) or stage-based (BIR, BOS, and SLC) HCMs were calculated to summarize overall hydrograph central tendency and variability and the magnitude, frequency, and duration of high- and low-flow periods (McMahon et al., 2003; Giddings et al., 2009). Three additional metrics were calculated: (1) the Richards-Baker flashiness index (rb_flash; Baker et al., 2004), which was based on daily values; (2) the percent daily change in mean daily values (day_pctchg); and (3) the duration of time that streamflow exceeded mean streamflow (Tqmean; Konrad & Booth, 2005). The 83 candidate metrics used in the initial HCM reduction are listed in Appendix Table 1 (Supplementary material) and are arranged according to flow component as defined in Olden & Poff (2003).

By definition, many of the HCMs (mean, median, maximum change, etc.) computed from stage-based time series data (in BIR, BOS, and SLC) were not comparable with the more commonly computed area-based HCM (five metropolitan areas). In addition, numerous streams in the SLC and DEN metropolitan areas were altered by interbasin transfer and reservoir storage, which likely resulted in poorer relations ($r^2 = 0.03, 0.20$) between watershed imperviousness and the frequency of high flow events (Fig. 3). Therefore, the metropolitan areas of BIR, BOS, DEN, and SLC were invalid for comparative purposes and not used in further analysis.

Previous investigations have found biology to be related to both the absolute size of flow (Clausen & Biggs, 1997) and drainage area (DA) normalized flow. To include the possibility that some HCMs may not have relations without accounting for drainage area, our initial set of candidate metrics contained 30 HCMs transformed by DA. HCMs with positive

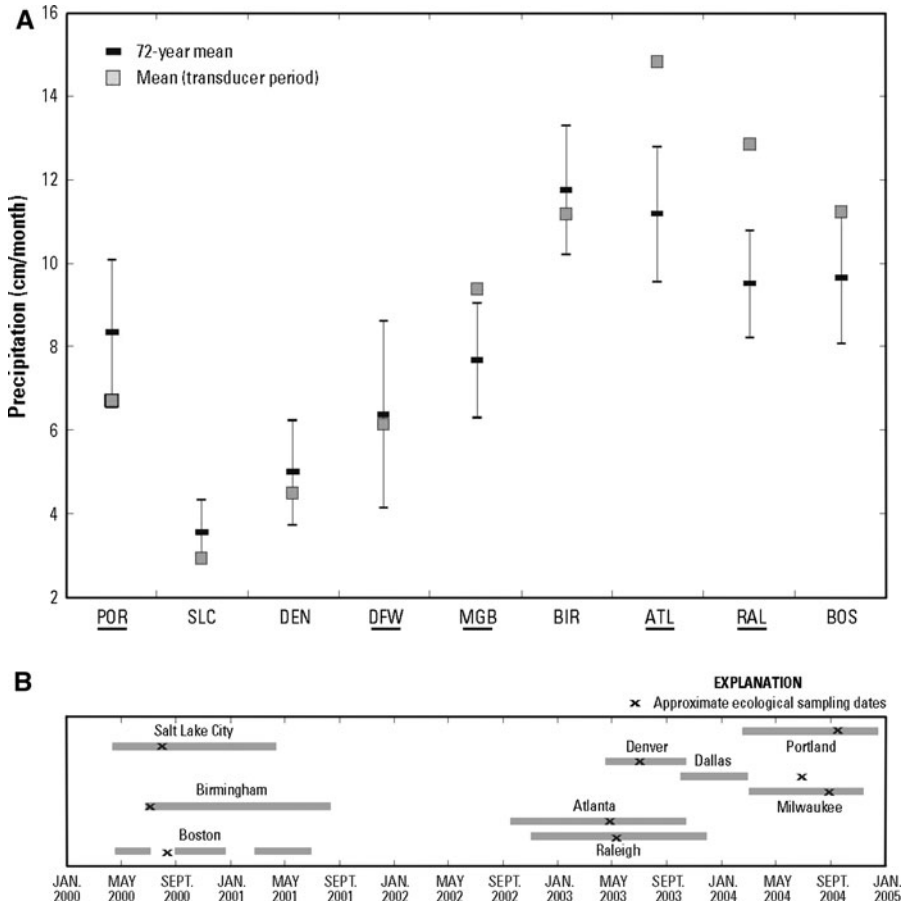


Fig. 2 Long-term (72-year mean) precipitation (±1 SD) and precipitation during the transducer collection period at nine metropolitan areas in the continental United States (**A**). *Underlined* metropolitan areas were used in the hydrologic condition metric analysis. Hydrologic data-collection periods and approximate ecological sampling dates are displayed (**B**)

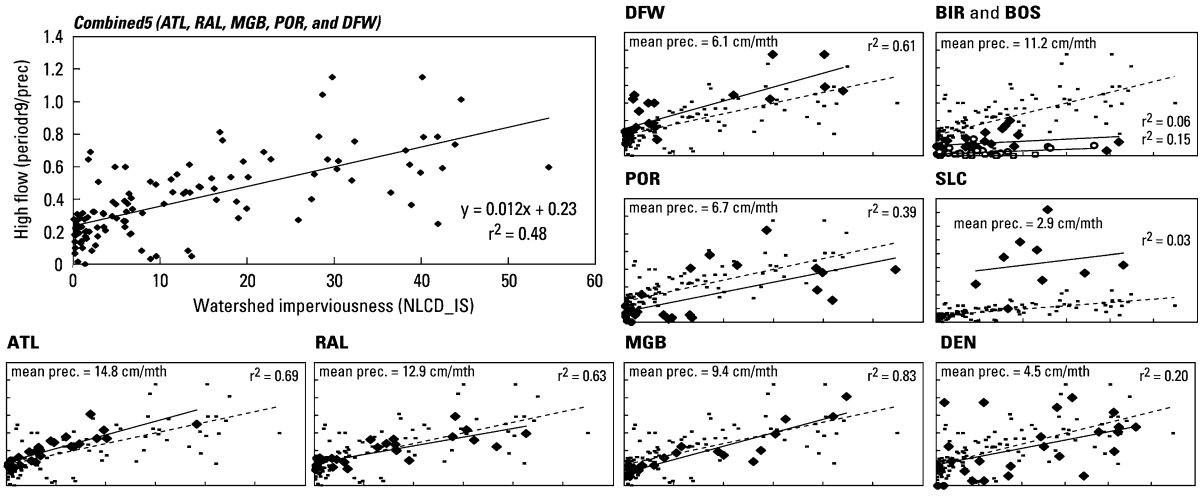


Fig. 3 Scatter plots of watershed imperviousness (NLCD_IS) against precipitation (prec) transformed hydrologic condition metric periodr9 for nine metropolitan areas and the national-scale (combined5) data set (ATL, RAL, MGB, POR, and DFW) with linear regression lines. *Light symbols* and *dotted line* in individual metropolitan plots are to enhance comparison to the combined5 results. Small scatter plots have same scales as the combined5 scatter plot. Mean precipitation as determined at one site in metropolitan area during the transducer collection period

correlation to DA were divided by DA; negative correlations were multiplied by DA (Appendix Table 1 in Supplementary material). The six stream cross-sectional area percentile metrics (low flow pct_25a, pct_10a, pct_5a, and high flow pct_99a, pct_95a, pct_75a) were normalized by the median cross-sectional area value, rather than DA, to reduce the effect of variable drainage sizes.

Precipitation varied widely across the five metropolitan areas, both in terms of long-term mean (77–134 cm/year) and during the actual period of data collection (Fig. 2A). The time series data were collected in the DFW area when precipitation was near the 72-year mean; POR flow data were collected during drier than normal conditions; and RAL, ATL, and MGB hydrologic periods of record were much wetter than their respective normal conditions. To compensate for the regional variations in climate in the combined5 data set, a limited number of HCMs were divided by the mean monthly precipitation from one site representative of the entire metropolitan area (value shown in Fig. 3; Appendix Table 1 in Supplementary material).

Biological data

Biological endpoints were derived from algal, invertebrate, and fish community data collected within each of the nine metropolitan areas in up to 30 stream reaches. Sampling reaches in each metropolitan area were selected from watersheds along gradients of increasing urbanization and were inclusive of or near the location of the hydrologic pressure transducer. Biological sampling occurred during an index period appropriate for each geographic area between 1999 and 2004. Index periods were chosen to represent the optimal time to collect biological data and were based on multiple factors, such as life history attributes of the organisms to be sampled and predictable hydrologic conditions that allow safe and effective sampling. The use of index periods ensures year-to-year data comparability within the study area and may be designated in different times of the year for different community types in some study areas (Cuffney et al., 1993). Community data were collected from these reaches using standard NAWQA data-collection methods (Moulton et al., 2002), which are outlined briefly here. Algal community data were derived from diatom communities sampled from hard substrates in each stream. Invertebrate communities were sampled twice at each site—one semiquantitative sample was collected from either rock or woody debris, and a separate qualitative sample was collected from all available habitats at each site. Fish communities were sampled using two-pass electrofishing techniques from reaches that were between 150 and 300 m in length. Algal metrics were selected from Lange-Bertalot (1979), Bahls (1993), and van Dam et al. (1994) and were summarized in Porter (2008) and were based on taxonomic composition, salinity indicators, trophic status, pollution tolerance, oxygen tolerance, motility, and biomass (Coles et al., 2009). Metrics calculated from the invertebrate community data were based on measures of abundance, richness, functional groups, tolerance, and indices of diversity. Invertebrate traits used in the calculation of metrics were from Barbour et al. (1999) and Cummins (1973). Metrics calculated from the fish data were based on richness, abundance, biomass, total length, tolerance, trophic guild, and traits (Goldstein & Meador, 2005); in addition, locally derived fish Index of Biotic Integrity (IBI) scores were calculated and used when available. Redundancy of the 26 biological metrics was examined using Spearman rank correlation analysis (Data Desk; Data Description, Inc., ver. 6.1, 1996); no metrics were intercorrelated ($|r| > \sim 0.90$) in more than two metropolitan areas. Two invertebrate metrics (EPTr and richTOL) were intercorrelated in ATL and POR, and two algal metrics (SiltIdx and MT_YS_DP) were intercorrelated in RAL and MGB. For use in predictive modeling, we selected a subset of biological metrics from the larger list of metrics that characterized broad features of community health, including measures of both abundance and richness. Selections were based on current metric use by various State and national assessment programs and the professional judgment of the authors (Table 2).

Data analysis

Identifying a subset of HCMs

Regression tree analysis (CART software; Salford Systems, San Diego, California) was used to identify a subset of HCMs strongly related to algal, invertebrate, and fish community metrics. This approach was

Table 2 Algal, invertebrate, and fish metric abbreviations and descriptions used in exploratory regression tree modeling (eRT), cross-validated predictive regression tree modeling (xpRT), and Spearman correlation analysis (SC)

Analysis type	Metric abbreviation	Metric description
Algal community[a]		
xpRT, SC	Diat_rich	Number of diatom taxa
eRT	SP_OL_DP	Saprobien index: oligosaprobous (zone of near complete organic load oxidation)
eRT	SP_BM_DP	Saprobien index: beta-mesosaprobous (zone of organic load oxidation, N as NH_4)
eRT	MT_YS_DP	Motile algae
eRT	Siltldx	Silt motile diatoms
xpRT, SC	CHLa[b]	Chlorophyll *a* concentration on hard substrate
xpRT, SC	Eutrophic	Percent eutrophic diatoms
Invertebrate community[a]		
eRT	SC_rich	Richness composed of scrapers (multihabitat)
eRT	DIPRp	Percent of total richness composed of Diptera (multihabitat)
eRT, xpRT, SC	RichTOL	Richness-based average USEPA tolerance value for sample (multihabitat)
xpRT, SC	EPTr	Richness of Ephemeroptera, Plecoptera, and Trichoptera
eRT, xpRT, SC	ABUND[b]	Total number of organisms in the sample
eRT	PR_abund	Total abundance composed of predators
eRT	FC_abund	Total abundance composed of filtering-collectors
eRT	SC_abund	Total abundance composed of scrapers
eRT	pSH_abund	Percent of total abundance composed of shredders
eRT, xpRT, SC	AbundTOL	Abundance weighted USEPA tolerance value for sample
eRT	Intol_abund	Total abundance in the tolerant class
eRT	Modtol_abund	Total abundance in the moderately tolerant class
Fish community		
eRT, xpRT, SC	Rich	Number of fish species
eRT, xpRT, SC	CenR	Richness of fishes in the Family Centrarchidae (sunfish)
eRT	IctR	Richness of fishes in the Family Ictalueridae (catfish)
eRT	aMud	Percent of community whose substrate preference is mud
xpRT, SC	TotAbund[b]	Total abundance of fish collected in sampling reaches
xpRT, SC	aGravel[b]	Total abundance of gravel-spawning fishes
xpRT	Fish TBI[c]	Locally derived Index of Biotic Integrity Score

[a] Unless otherwise noted in metrics description, metrics were calculated from samples collected from rock substrate except in ATL and DFW metro areas where woody debris were sampled

[b] Data transformed (4th root) prior to analysis

[c] Fish IBI scores were used only for the DFW, MGB, ATL, and RAL metropolitan areas in correlation table (Appendix Table 2 in Supplementary material) and were not used in computation of values in Fig. 5

successfully used to reduce habitat and land-use stressor variables in a fish distribution investigation (Steen et al., 2008). Regression trees are a powerful analytical method for descriptive and predictive modeling of complex data (Efron & Tibshirani, 1991; De'ath & Fabricius, 2000) and have been used to explore non-linear or step function relations between a single dependent variable and several independent variables (Breiman et al., 1984). Regression tree modeling has gained popularity as a screening method for variables, for summarizing relations in large multivariate data sets and for providing a flexible non-parametric alternative to generalized linear modeling for regression problems (Kennard et al., 2006). This method uses recursive binary partitioning to divide response data into homogenous groups with

combinations of explanatory variables. Explanatory variables and value thresholds are selected to minimize the total sum of squared residuals (deviance) for each specific subgroup. Regression tree models result in output resembling a tree, with binary-decision nodes that split data into branches and terminal-end nodes, or leaves that contain groups of similar data. For complex data, such as ours, regression tree models offer several important advantages in that they are distribution-free, are robust to outliers and missing data, do not require a priori selection of variables, can accommodate large numbers of exploratory variables, and can accommodate threshold responses and context-dependent interactions. Further, regression trees can effectively model variable relations despite significant spatial autocorrelation in data (Cablk et al., 2002). Relative importance scores of model variables were calculated by summing the improvements in model error for each variable across all nodes for the primary split and first five surrogate splits. The relative importance score was scaled (range = 0–100) to the variable with highest improvement score (Breiman et al., 1984; Steinberg & Colla, 1997).

Exploratory regression tree models were run for the full set of HCM variables with the biological metrics listed in Table 2 as the target endpoints. The initial set of HCMs was reduced to the most relevant metrics based on score magnitude and frequency of appearance in the variable relative importance tables. Retained HCMs were examined using Spearman rank correlation analysis (Data Desk; Data Description, Inc., ver. 6.1, 1996) with intercorrelated variables ($|r| > \sim 0.90$) removed. The 37 HCMs retained from the data-reduction phase (referred to as reduced set in Appendix Table 1 in Supplementary material) were organized according to flow component (Olden & Poff, 2003). Principal component analysis and factor loading plots were used to examine structure of the reduced set of HCMs between metropolitan areas and enable comparison to previous investigations. Five primary HCMs were selected from the reduced set based on regression tree analysis and Spearman rank correlation.

Predictive relations between HCMs and biological metrics

We considered the relation between HCMs and biological metrics on two spatial scales—individual metropolitan areas and a national scale using the combined5 data set. Regression tree models constructed on the reduced data set of HCMs and 11 biological endpoints (Table 2) incorporated v-fold cross-validation ($v = 10$) to aid in the selection of the most parsimonious tree and assess error rates (Breiman et al., 1984). Briefly, the procedure behind v-fold cross-validation is to randomly divide the data into equal subsets ($v = 10$) and omit each subset in turn (test data), building a series of tree models on the remaining subsets (learn data). The omitted subsets (test data) are then used to estimate the true error rates for the combined tree model. Our aim was to identify predictive biological endpoint tree models with minimum relative cross-validated error (Breiman et al., 1984; Steinberg & Colla, 1997). A measure of overall model R^2 (r^2_{eq}) is 1—relative resubstituted error (Breiman et al., 1984). We discarded candidate trees with >7 terminal nodes for being too complex (for combined5 data set). Additional comparisons were made with regression tree models formulated on the reduced HCM data set (excluding DA-transformed variables) and (1) drainage area as an independent variable (metropolitan-area scale) and (2) three contextual variables (water–surface gradient, precipitation, and drainage area) on the national scale (combined5 data set).

Relation between HCMs and urbanization

To examine the relation of HCMs to urbanization, we primarily used Spearman rank correlations between HCMs and land-cover characteristics commonly represented in deterministic models used for hydrograph computation, such as Hydrologic Simulation Program-Fortran (HSPF) or Precipitation-Runoff Modeling System (PRMS; Cassin et al., 2005). We identified management options beyond wholesale impervious-surface reduction, which were associated with biologically relevant HCMs. Specifically, we sought to identify land-cover and pattern variables that (1) had stronger relations with HCMs (periodr9 and mdh_95) than total percent imperviousness or developed land in the basin (NLCD_2) or (2) had significant correlation with HCMs but did not operate in concert with percent developed land in the basin (not highly correlated with NLCD_2).

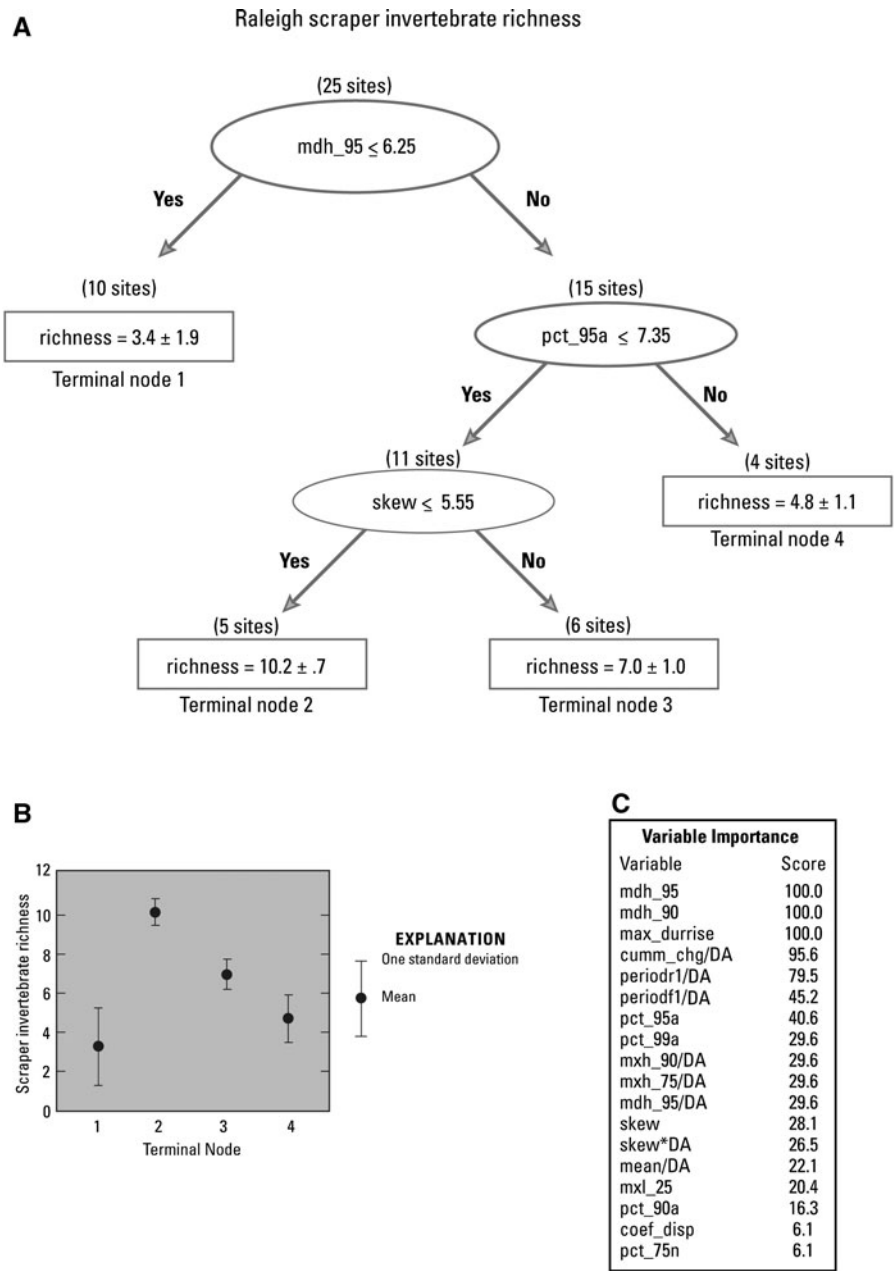


Fig. 4 Regression tree model of the influence of hydrologic condition metrics on scraper invertebrate richness in the RAL metropolitan area (**A**). The parentheses indicate the number of sites at a node. Terminal nodes are denoted by a rectangle with invertebrate richness (mean ±1 SD). To read the model, start from the top and work down. For example, if mdh_95 is >6.25 and pct_95a is >7.35, four sites meet these criteria with a mean richness of 4.8. Terminal node means and standard deviations are displayed (**B**) as are importance of variables used in model formation (**C**)

Results

Reduction of HCMs

The RAL invertebrate scraper richness regression tree model is an example of the ~100 exploratory regression tree models constructed and analyzed in the initial variable-reduction process (Fig. 4). Duration of high flow (mdh_95) was the primary split variable (Fig. 4A) and one of the most important variables in model formation (Fig. 4C). Sites lowest in scraper richness had shorter duration of high flow (terminal node 1, mean = 3.4), a finding that is consistent in invertebrate scraper–streamflow relations determined from 111 sites in Western United States (Konrad et al., 2008). The highest scraper richness was found at the sites with increased high-flow duration, decreased high-flow magnitude

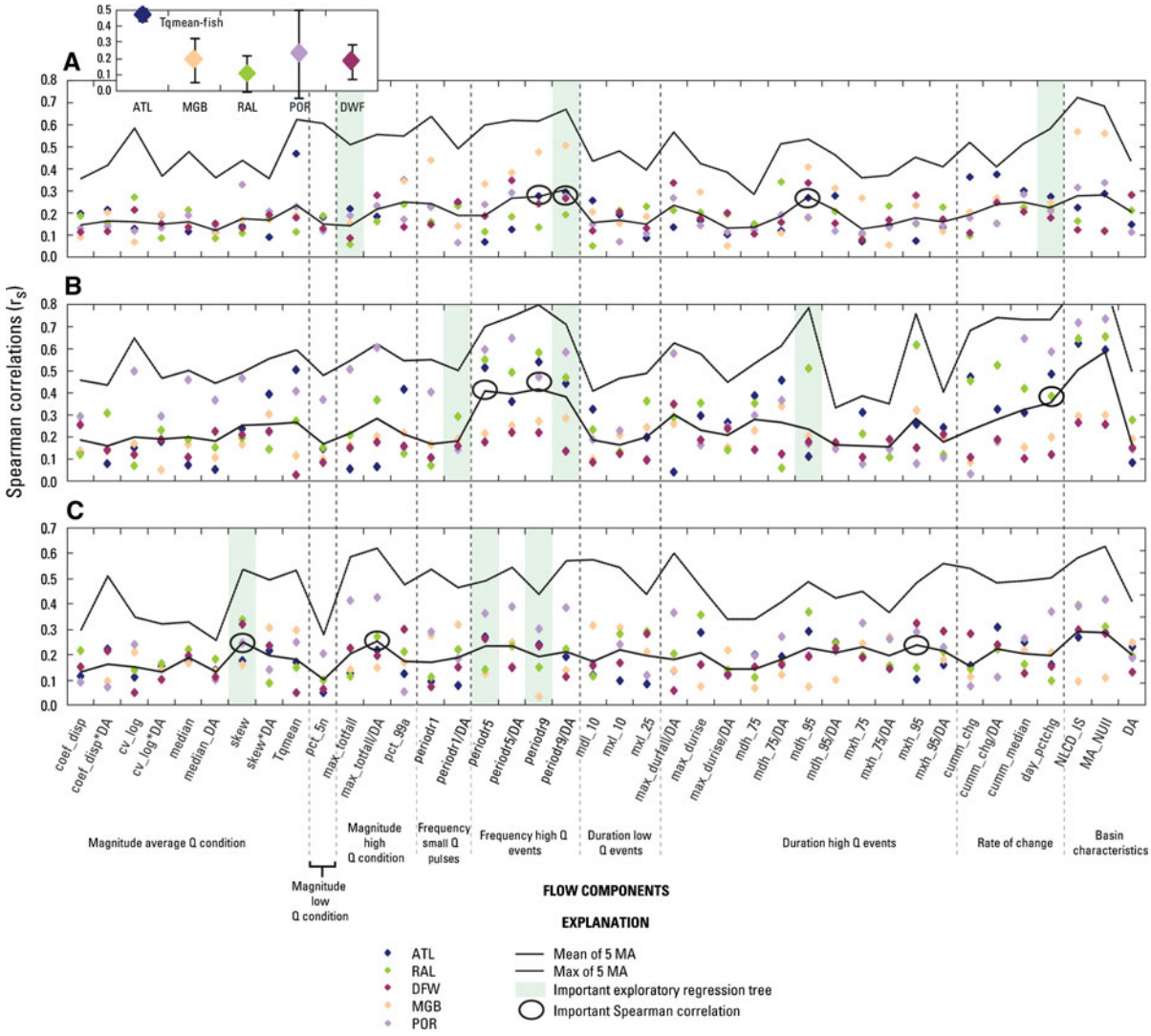


Fig. 5 Scatter plots of Spearman correlations (r_s) between reduced set of hydrologic condition metrics (HCM) and **A** fish, **B** invertebrate, and **C** algal metrics. *Symbols* represent mean HCM correlation for three to four biological metrics in a metropolitan area (see inset for Tqmean-fish example). *Solid line* indicates mean correlation for five metropolitan areas (MA) with largest values circled. *Upper dashed line* indicates maximum correlation value for an HCM and a biological metric. HCM arranged according to flow component (Olden & Poff 2003). Most important HCMs from exploratory regression tree analysis are denoted. Biological metric definitions are listed in Table 2. Correlation values (r_s) are listed in Appendix Table 2 (Supplementary material)

(pct_95a), and decreased skew (terminal node 2; mean = 10.2). Two other high-flow duration metrics, max_durrise and mdh_90, had equal importance in model formation to mdh_95 (Fig. 4C); however, mdh_90 was not retained due to high correlation with mdh_95. Skew was an important variable in a subset of the data (11 sites, split nodes for terminal nodes 2 and 3; Fig. 4) and illustrates regression tree model capability for context-dependent variable interaction. (Skew was only important in the 11 sites with reduced pct_95a and increased mdh_95.) The 15 drainage area-transformed HCMs and 22 non-transformed HCMs most important in construction of the individual metropolitan area regression tree models (reduced set) are denoted in Appendix Table 1 (supplementary material) ("reduced data set"). The

most important HCMs (frequent and high values in regression tree variable importance tables) in each of the three taxa groups are denoted in Fig. 5.

Reduced HCM data structure

The factor loadings [correlation with individual variables and principal components axis 1 (PCA1) and principal components axis 2 (PCA2)] indicated the reduced set of HCMs were relatively independent of each other and accounted for the majority of variation in the combined5 and individual metropolitan area HCM data sets (Appendix Table 1 in Supplementary material; Fig. 6). Factor plots for the five individual metropolitan areas and the combined5 data set were similar, with the rate-of-change metric (day_pctchg) consistently dominant on factor axis 1 ($r > 0.84$ except ATL where $r = 0.76$) and the frequency of high-flow metric (periodr9) located often in the quadrant position between the two primary factor axes (combined5, POR, DFW, RAL). In some instances HCMs were nearby in the two-dimensional factor plot but differed on the third-factor axis or in correlation with biological endpoints. For example, the high-flow duration metric (max_durrise/DA) and magnitude average flow (cv_log*DA) were proximate in the lower right quadrant; however, third axis values were 0.43 and −0.80, respectively. The duration of high-flow metrics mxh_75/DA and mxh_95/DA was similar in all the three factor axes; however, mxh_95/DA had stronger maximum correlation with algae metrics (DFW; CHLa) than mxh_75/DA (Fig. 5C; Appendix Table 2 in Supplementary material). Similarly, two precipitation-transformed high-flow frequency metrics (periodf5, periodr9) were proximate in the factor plot (Fig. 6); however, periodr9 had stronger correlation with fish metrics in MGB (Fig. 5A). In general, HCM data structure was similar with or without inclusion of the drainage area-transformed variables.

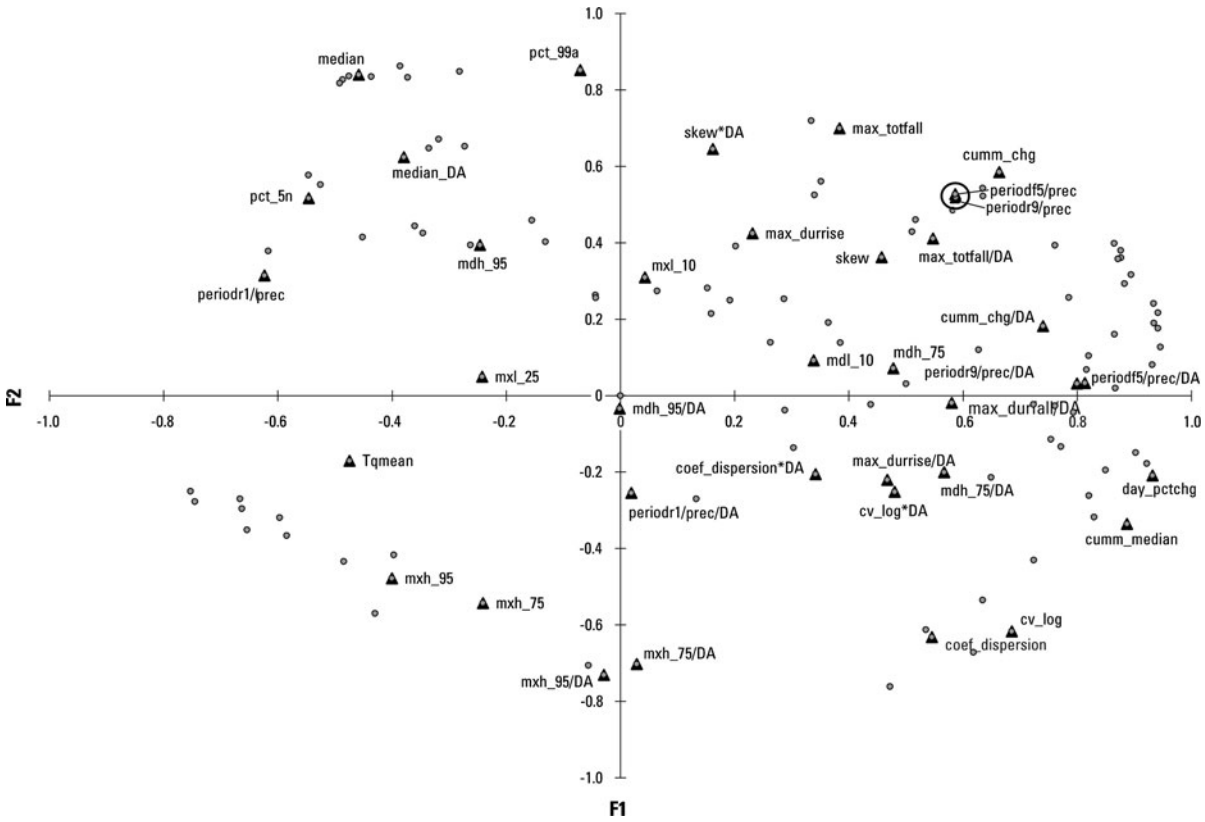


Fig. 6 Factor loading plot of the first two principal components of the 128 candidate hydrologic condition metrics (HCM) for the national scale combined5 data set. *Labeled triangles* indicate the reduced set of HCM. The two circled HCMs indicate precipitation-transformed variables periodr9 and periodf5

Selection of five HCMs based on relations to measures of aquatic biota

Results of Spearman rank correlation (r_s) analysis between the reduced set of HCMs and biological variables (three algal, four invertebrates, and four fish; Table 2) are presented in Fig. 5 for individual metropolitan areas. The HCMs are arranged according to flow component (Olden & Poff, 2003) with individual metropolitan area mean correlation between HCM and community type represented (illustrated in Fig. 5 inset for Tqmean). Specific Spearman rank correlation values are listed in Appendix Table 2 (Supplementary material). The HCM–biological correlations were used as a screening tool to highlight the HCMs substantially associated with biological metrics, to contrast the biological response among community types, and to assess flow-component relations with biological communities. Metropolitan area correlations were not directly comparable because of varied sample size; however, Appendix Table 2 (Supplementary material) provides comparison based on P value.

The strongest individual HCM relations with fish metrics were between the high-flow frequency metric (periodr9/DA) and total abundance (Fig. 5A, Appendix Table 2 in Supplementary material; $r_s = 0.67$, MGB). Milwaukee-Green Bay (MGB) had some of the strongest relations with fish metrics; this was especially true for the frequency of high-flow HCM (periodf5, periodr9). Two rate-of-change metrics (day_pctchg and cumm_median) had similar fish-relation strength for all five metropolitan areas (clustered symbols), although correlations were not as strong as the high-flow frequency (periodr9) or duration of high-flow (mdh_95) metrics. The rate-of-change (day_pctchg), high-flow magnitude (max_totfall), and frequency of high-flow (periodr9/DA) metrics were most important in the fish regression tree models (Fig. 5A).

The frequency of high-flow metrics (periodr9 and periodf5) had the strongest mean correlations ($|r_s| = 0.42$ and 0.41, respectively) of all HCMs with invertebrate metrics (Fig. 5B). The rate-of-change metric day_pctchg also had correlation with invertebrates (mean $|r_s| = 0.35$), albeit with large differences between metropolitan areas. The strongest HCM–invertebrate relation was between periodr9 and RichTol ($r_s = 0.80$, ATL), although the two high-flow duration metrics (mdh_95 and mxh_95) in RAL had similar strong relations with EPTr ($r_s = 0.79$ and 0.76, respectively; Appendix Table 2 in Supplementary material). The duration and frequency of high-flow events (mdh_95, periodr9/DA) and the frequency of small flow pulse (periodr1) metrics were most important in the exploratory regression tree models (Fig. 5B).

The magnitude and duration of high flow (max_totfall/DA and mxh_95) and average flow magnitude (skew) metrics had the strongest mean correlations with algae (Fig. 5C). The strongest correlation was with a high-flow metric (max_totfall/DA) and eutrophic algae in POR ($|r_s| = 0.62$). Skew, periodf5, and periodr9 were the three most important HCMs in the algal regression tree models.

Primary HCMs, based on Spearman rank correlation and regression tree analysis in the individual metropolitan areas, were selected from each of five flow components. Metrics representing magnitude of average flow (skew), high-flow magnitude (max_totfall), frequency of high-flow events (periodr9), duration of high flow (mdh_95), and rate of change (day_pctchg) were selected as most useful in biological relations.

We found that frequency of high-flow events (periodr9) had significant relations ($P < 0.05$) in each of the five metropolitan areas with all the three taxa groups except algae in DFW and MGB (Appendix Table 2 in Supplementary material; Fig. 7). Increased periodr9 was associated with decreased fish IBI and abundance, decreased invertebrate EPTr, increased number of tolerant invertebrate species, and increased percent of eutrophic diatoms. Transformation of periodr9 with DA improved 7 of the 15 relations.

Comparison of HCM–biological relations for individual metropolitan areas

In general, stronger HCM–biological correlations were observed in metropolitan areas with forested non-urban sites (ATL, RAL, and POR) compared with metropolitan areas with agricultural non-urban sites (MGB, DFW; Appendix Table 2 in Supplementary material; Fig. 5). The highest numbers of significantly ($P < 0.05$) associated algal, invertebrate, and fish biological metrics with HCMs were observed in the POR, ATL, and RAL metropolitan areas with 82, 66, and 61 correlations, respectively. Within all the three of these metropolitan areas, the

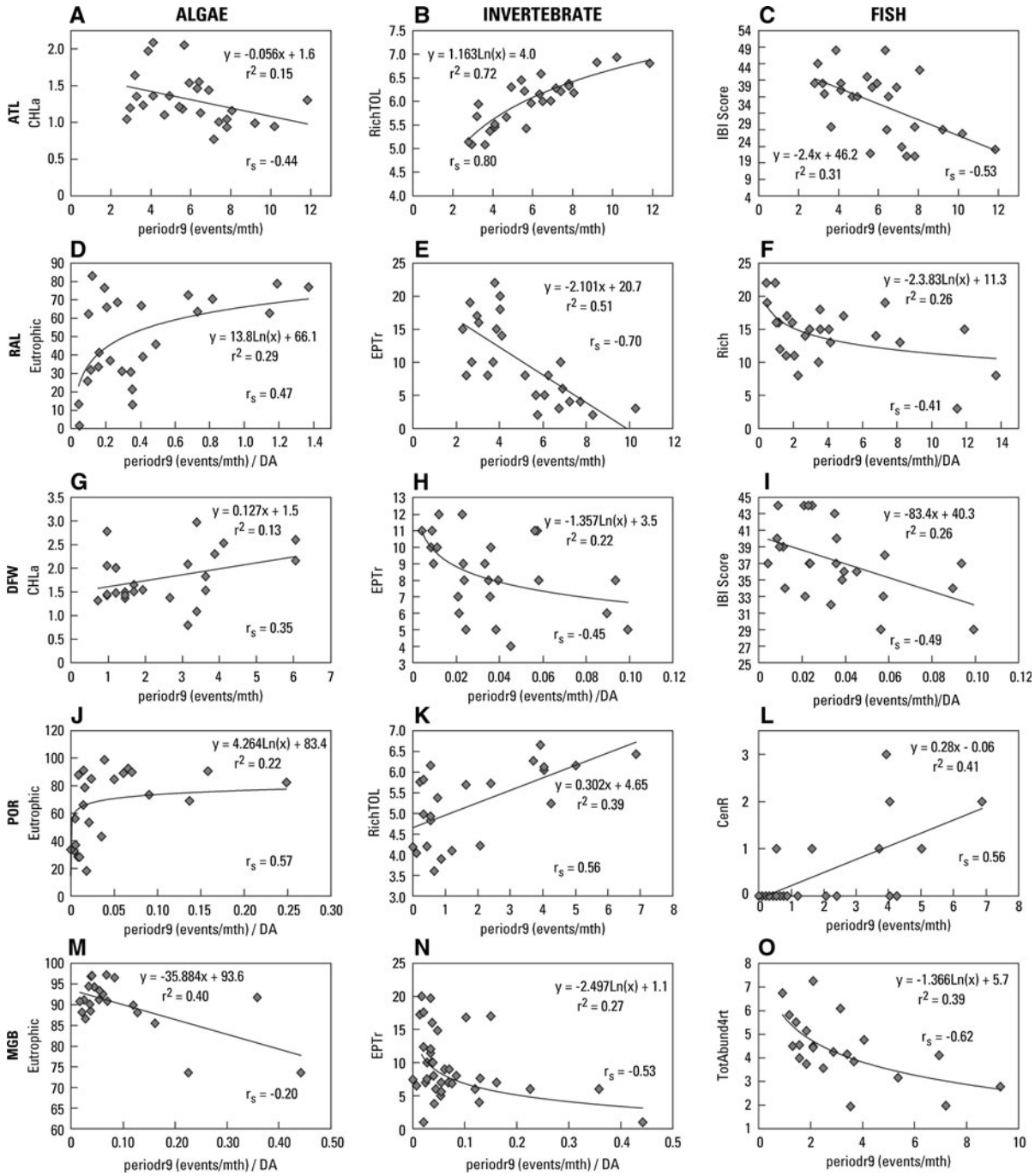


Fig. 7 Scatter plots of the high-flow event frequency hydrologic condition metric (periodr9) and an algal, invertebrate and fish metric in the five metropolitan areas. The selected linear and logarithmic model parameters and Spearman correlation coefficients (r_s) are indicated

largest numbers of significant relations were with the invertebrate community metrics, specifically EPTr and RichTol. Fish community metrics were significantly negatively correlated with the frequency of high-flow events (periodr9, periodf5) most strongly in the MGB metropolitan area. Fish IBI scores and abundance of gravel-nesting species (aGravel) declined in ATL; overall fish species richness

declined in RAL; and fish richness, centrarchid richness, and IBI scores declined in DFW in response to high-flow event frequency. Algal community metrics were relatively weakly associated with HCMs in the metropolitan study areas, although mean correlations were highest in association with HCMs that described high-flow event frequency and conditions in the POR metropolitan area. Chlorophyll *a* (CHLa) levels corresponded significantly to HCMs in ATL, RAL, and DFW study areas. The DFW metropolitan area had the lowest number of significant responses (23), and the pattern of responses was unique in that only three flow components (five HCMs) were significantly correlated with two invertebrate community metrics, whereas correlations were substantially more common in other metropolitan areas (in both invertebrate metrics and flow components). The only metropolitan area where the percentage of eutrophic diatoms response was not significantly correlated with any HCMs was DFW; however, CHLa levels in the DFW area were positively correlated with cumm_chg and max_totfall and inversely correlated with mxh_95/DA. The pattern of correlations between HCMs and measures of biology was distinct in the MGB metropolitan area. First, the total number of significantly correlated metrics was lower (36 correlations), slightly more than half as many as the next highest metropolitan area (RAL, 61 correlations) and less than half as many as were observed in POR (82 correlations). Second, in terms of individual biological communities, fish rather than invertebrates tended to respond more consistently and strongly to HCMs. Third, the biological metric response to basin urbanization (NLCD_IS and MA_UII) had similar, not greater, correlation strength as the response to HCMs for the fish and invertebrate community metrics. In general, correlation coefficients between biological metrics and urbanization (NLCD_IS and MA_UII) were substantially higher than with HCM relations, indication that other landscape factors (i.e., chemistry, temperature, and sediment) as well as hydrology influence changes in biological communities (Fig. 5; Appendix Table 2 in Supplementary material).

Metropolitan-area scale predictions of biota response to HCMs

At the metropolitan scale ($n < 30$), 14 regression tree models were determined to be predictive based on cross-validation. The MGB area had the fewest sites ($n = 22$) and no predictive models. In all metropolitan areas, however, notable biological endpoint separation was determined based on one HCM (Appendix Table 3 in Supplementary material). Overall, HCMs explained up to 73, 92, and 79% of the variance for the algal, invertebrate, and fish community metrics, respectively (Table 3). The frequency of high-flow events (periodr9) was the most common primary split variable (3 models). As with Spearman rank correlations, regression tree models derived for invertebrate metrics tended to be the strongest.

The sole predictive algal model (POR) had the least eutrophic diatom sites associated with decreased magnitude of high flow (max_totfall) and lower rate of flow change (cumm_chg/(DA), Fig. 8A, terminal node 1). The 12 most eutrophic diatom sites were associated with increased magnitude of high flow (max_totfall). The models most predictive ($r^2_{eq} \geq 0.70$) for invertebrate community metrics were produced in the RAL and ATL metropolitan areas (Table 3). The three most predictive invertebrate models had periodr9 as the primary split variable; periodr9 had negative correlation with EPTr and positive correlation with RichTOL. In the RAL area, sites with greatest EPTr were associated with decreased frequency of high-flow events (periodR9) and longer duration of high-flow events (mdh_95, Fig. 8B, terminal node 2). The models most predictive of change in fish community metrics were produced in POR, RAL, and DFW (Table 3). The most predictive of all community metric models were observed in the RAL metropolitan area with the invertebrate EPTr and RichTOL ($r^2_{eq} = 0.92$ and 0.83, respectively). Detailed regression tree results, including split node HCMs and biological endpoint values are included (Appendix Table 3 in Supplementary material).

The Spearman rank correlation and regression tree analyses identified numerous HCMs that improved with transformation of DA (Fig. 5; Appendix Table 2 in Supplementary material). Relations involving all the three biological community types and high-flow magnitude (max_totfall/DA) and rate of change (cumm_chg/DA) were improved when the HCMs were transformed by DA. Transformation of the frequency of high flow (periodr9/DA and periodf5/DA) and magnitude of average flow (skew*DA) improved biological metric relations but with less

Table 3 Summary of model results from the cross-validated regression tree analysis

Community metric	Metropolitan area												National scale "combined5"			National scale contextual models		
	ATL			RAL			DFW			POR								
	TN	r^2_{eq}	Primary split	TN	r^2_{eq}	Primary split	TN	r^2_{eq}	Primary split	TN	r^2_{eq}	Primary split	TN	r^2_{eq}	Primary split	TN	r^2_{eq}	Primary split
Algae																		
CHLa													4	0.45	max_totfall −	8	**0.72**	max_totfall −
Diat_rich																2	0.20	cumm_chg +
Eutrophic										3	**0.73**	max_totfall +	4	0.55	cumm_chg −	2	0.43	MAP −
Invertebrate																		
ABUND							2	0.47	cv_log*DA +	2	0.55	max_totfall/DA −	2	0.23	max_totfall −	7	0.64	max_totfall −
AbundTol													5	0.62	median +	5	0.61	median +
EPTr	3	**0.76**	day_pctchg −	4	**0.92**	periodr9 −							2	0.29	periodr9/prec −	6	**0.70**	periodr9/prec −
RichTOL	3	**0.76**	periodr9 +	3	**0.83**	periodr9 +				2	0.65	max_totfall/DA +	5	**0.74**	max_totfall/DA +	4	0.67	MAP −
Fish																		
TotAbund																		
Rich										3	**0.72**	mxh_75/DA −	2	0.21	cumm_chg +	6	0.52	cumm_chg +
CenR				4	**0.79**	median +	2	0.41	periodf5/DA −				2	0.40	max_totfall +	2	0.40	max_totfall +
aGravel	2	0.44	cumm_chg −										3	0.52	cumm_median +	3	0.52	cumm_median +
IBI				2	0.41	mdh_75/DA −	4	**0.79**	periodf5/DA −	–	–	–						

Table shows biological community metrics, 1-relative resubstituted error (r^2_{eq}), and primary split variable (hydrologic condition metric; HCM) selected for each model in each metropolitan area and for combined metropolitan area data sets. No predictive models for the MGB metropolitan area. Bolded coefficients identify $r^2_{eq} > 0.70$. The HCM prefix sign indicates direction of change. "TN" denoted number of terminal-end nodes. Model details including HCM split values and terminal node separation are contained in Appendix Table 3 (Supplementary material). National-scale contextual models formulated without drainage area-transformed HCM and with contextual variables [drainage area (DA), site-specific long-term mean annual precipitation (MAP), and water surface gradient (WSG)]. "–" denotes no model attempted

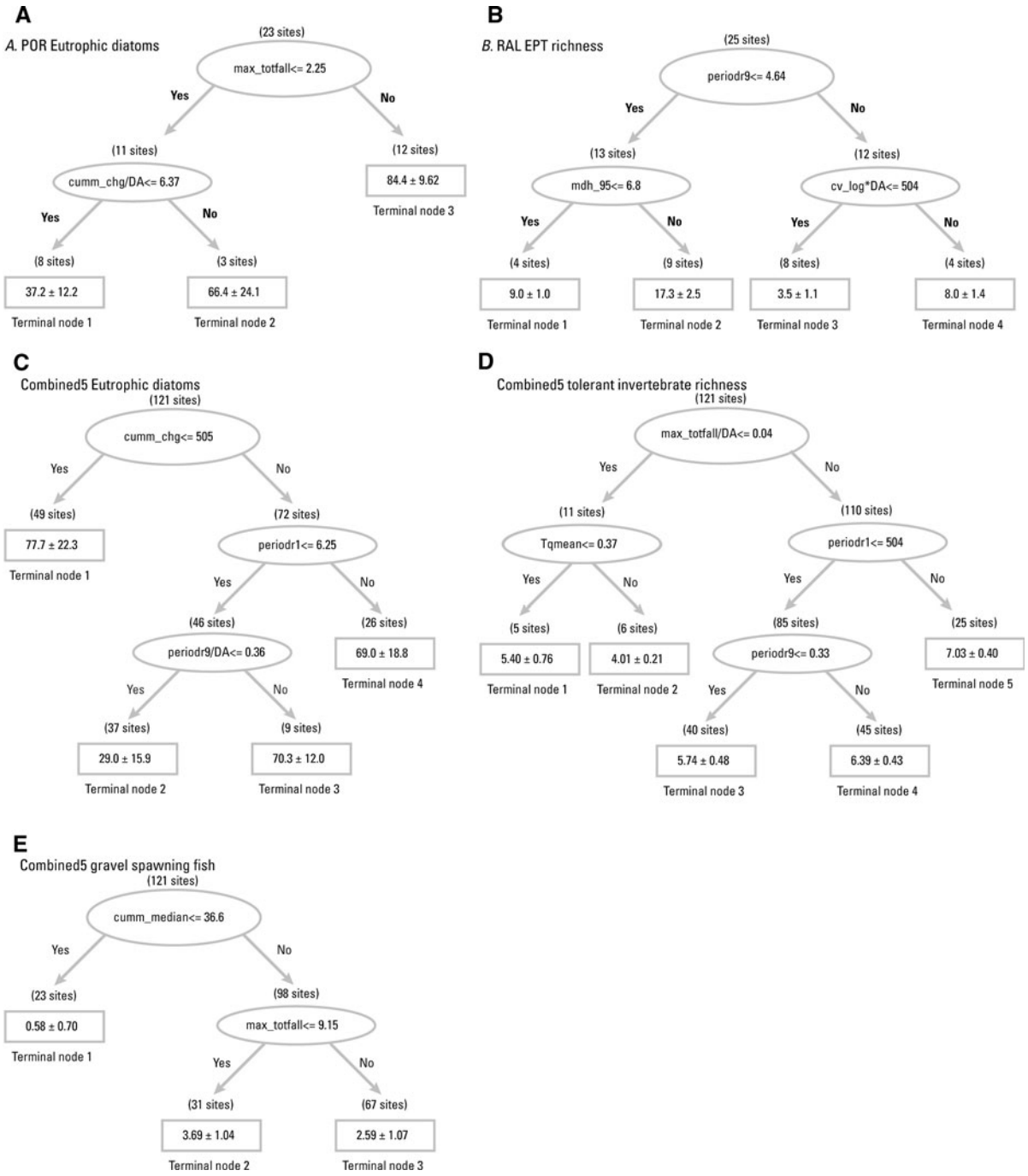


Fig. 8 Regression tree models of the influence of hydrologic condition metrics on eutrophic diatoms in the POR metropolitan area (**A**), EPT invertebrate richness in the RAL metropolitan area (**B**), and eutrophic diatoms (**C**), tolerant invertebrate richness (**D**), gravel-spawning fish abundance (**E**) on the combined5 national scale. The parentheses indicate the number of sites at a node. Terminal nodes are denoted by a rectangle with biological endpoint value (mean ±1 SD)

consistency (Fig. 5). One half of the predictive metropolitan area regression tree models used DA-transformed HCMs as the primary split variable (Table 3). Drainage area as an independent variable was listed in two of the predictive-model variable importance tables (data not shown); however, the DA

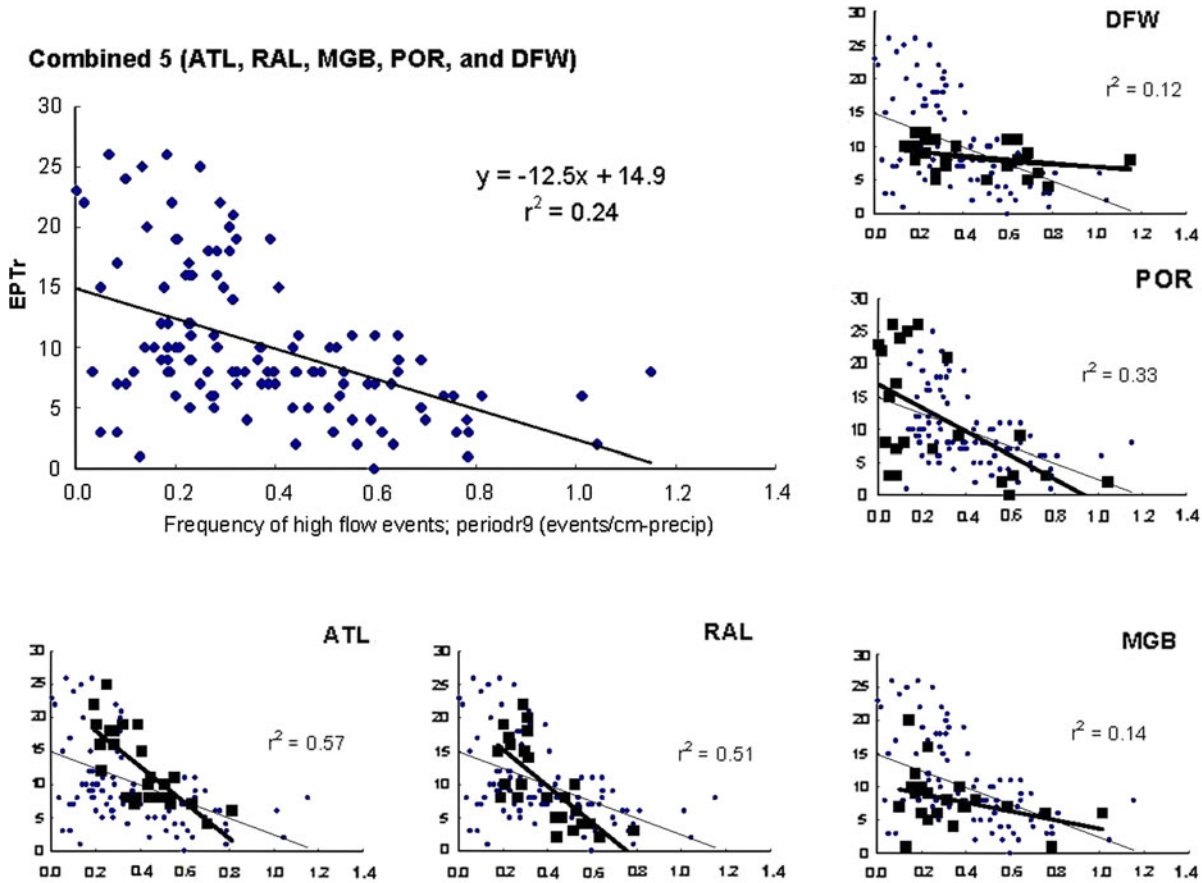


Fig. 9 Scatter plots of the invertebrate richness metric, EPTr, against the precipitation transformed frequency of high-flow event hydrologic condition metric for five metropolitan areas and the national-scale (combined5) data set (ATL, RAL, MGB, POR, and DFW) with linear regression lines. *Light symbols* and *dotted line* in individual metropolitan plots are to enhance comparison to the combined5 results. Small scatter plots have same scales as the combined5 scatter plot

variable was not used as a split variable in any metropolitan area models.

National-scale predictions using precipitation-transformed HCMs

The frequency of high-flow event metric (periodr9) had significant ($P < 0.01$) linear response to imperviousness in each of the five metropolitan areas; r^2 ranged from 0.39 (POR) to 0.83 (MGB, Fig. 3). To account for climate variability, periodr9 was divided by the mean monthly precipitation from one site representative of an entire metropolitan area (value shown in Fig. 3). For example, periodr9 values in POR were divided by the precipitation rate of 6.7 cm per month; the flashiest POR site had an average of 7.0 events per month (1.04 events/cm × 6.7 cm/month, Fig. 3). Despite collection of hydrologic data over a wide range of physiographic regions, climatic conditions, dissimilar time intervals, and lengths of record (Table 1; Fig. 2), the relation between imperviousness and periodr9, when transformed by a precipitation value, was significant ($P < 0.01$; $r^2 = 0.48$) for the combined5 data set. Individual metropolitan areas had similar rates of change. The national-scale (combined5) relation was substantially weaker without precipitation transformation ($r^2 = 0.28$). It is likely that transformation of periodr9 with site-specific precipitation (note the substantial long-term precipitation differences between sites within a given metropolitan area; Table 1) may result in stronger relations; however, site-specific precipitation during

the transducer data collection was not available. Spearmen rank correlations using the precipitation-transformed duration of high-flow metrics (mxh_95, mdh_95) did not show improvement with either imperviousness or biological metrics.

The precipitation-transformed periodr9 variable also had significant relation ($r^2 = 0.24$, $P < 0.01$) with EPTr at the national scale. Lowest EPTr values ($\leq$11) occurred at sites with periodr9 > 0.45 events/cm, a sector represented by all five metropolitan areas (Fig. 9). The greatest rate of change occurred in metropolitan areas with predominantly forested non-urban sites (RAL and ATL). The poorest relations and smallest rate of change occurred in metropolitan areas with agricultural non-urban sites (DFW and MGB). The poorer relation in POR may have resulted from a mixture of agriculture and forest in the non-urban areas. The seven sites in the lower left corner of the scatter plot in Fig. 9 are in the MGB (2) and POR (5) metropolitan areas. Inclusion of percent agriculture in the basin on the abscissa affected these points and improved the overall relation ($r^2 = 0.37$; data not shown).

Cross-validated regression tree models formulated on the combined5 data set consisted of 2–5 terminal-end nodes and explained up to 55, 74, and 52% of the variance for the algal, invertebrate, and fish community metrics, respectively (Table 3). In the algal eutrophic model (Fig. 8C) less frequent high-flow events (periodr9 transformed by precipitation and DA) resulted in decreased eutrophic diatoms (terminal node 2); sites with more frequent high-flow events had increased eutrophic diatoms (terminal node 3). In the 4-HCM invertebrate (RichTOL) model ($r^2_{eq} = 0.74$), reduced high-flow magnitude (max_totfall/DA) and increased Tqmean defined the most desirable sites (low RichTOL, terminal node 2, Fig. 8D). Sites with the highest RichTOL values had increased high-flow magnitude in conjunction with one of the two conditions—increased frequency of small flow pulses (periodr1, terminal node 5) or increased frequency of high-flow events (periodr9, terminal node 4). The gravel-spawning fish (aGravel) model, which used rate-of-change (cumm_median) and magnitude of high flow (max_totfall) metrics, accounted for 52% of the variance (Fig. 8E). Sites least favorable for gravel-spawning fish had the lowest rate of change in flow (terminal node 1); most favorable sites had a higher rate of change but lower magnitude of maximum flows (terminal node 2). Incorporation of contextual variables (long-term site-specific precipitation, DA, and watershed slope) as independent variables in the regression tree models improved five national-scale regression tree models (Table 3).

Relation between HCMs and urbanization

Two metrics that characterize stream flashiness, duration of high flow (mdh_95) and frequency of high-flow events (periodr9), were orthogonal in data structure (Fig. 6) and correlated with two common measures of urban development, watershed percent imperviousness and percent developed land, in all metropolitan areas (Table 4). The MGB area had the strongest correlation with mdh_95 ($r_s = -0.82$) and periodr9 ($r_s = 0.92$). In all metropolitan areas, periodr9 had a stronger correlation with imperviousness and development than mdh_95. No metropolitan area indicated an impervious threshold for periodr9 (Fig. 3).

In each metropolitan area, density of roads (RDARDEN) was more strongly related to one or both of the HCMs than imperviousness. In the ATL, RAL, and MGB areas, the distant-weighted urban land variable (pURBANdw) had stronger relation to flow duration (mdh_95) than overall percent development. In ATL, increased percent wetland in the basin (NLCD1_9, $r_s = 0.82$), greater proportion of like adjacencies of wetland patches (PLA_C9, $r_s = 0.79$), mean area of wetland patches (PAM_C9, $r_s = 0.75$), and largest patch index (LPI_C9, $r_s = 0.71$) were more strongly correlated to quality flow duration (increased mdh_95) than percent developed land or imperviousness in the basin ($r_s = -0.40$ and -0.41, respectively).

Examined at a national scale, the mean area of developed patches (PAM_C2, $r_s = 0.76$) had the strongest correlation to high-flow event frequency (periodr9), which suggests that large contiguous areas of developed land may produce more frequent high flows. Other patch variables that had similar response also described larger contiguous areas of developed land—largest developed patch index (LPI_C2), a measure of connectivity, and the proportion of like adjacencies of developed patches (PLA_C2), a measure of the clustering of developed patches. Consistent with the metropolitan-area scale, high-flow duration was associated with reduced density of roads and reduced developed land in the riparian

Table 4 Spearman rank correlation coefficients (r_s) for the selected GIS variables and two biologically important HCMs for five individual metropolitan areas, five metropolitan areas combined, and combined metropolitan areas with total impervious surface less than 10%

	ATL ($n = 26$)		RAL ($n = 25$)		POR ($n = 23$)		MGB ($n = 22$)		DFW ($n = 25$)		combined5 ($n = 121$)		combined5 $\leq$ 10% imperv ($n = 72$)	
	mdh_95	periodr9	mdh_95	periodr9	mdh_95	periodr9	mdh_95	periodr9	mdh_95	periodr9	mdh_95	periodr9	mdh_95	periodr9
Urban indicator variables														
Watershed total impervious surface (%) NLCD_IS	−0.41	0.83	−0.69	0.80	0.15	0.65	−0.82	0.92	−0.26	0.78	−0.41	0.71	−0.20	0.29
Watershed total developed land (%) P_NLCD1_2	−0.40	0.90	−0.74	0.84	0.15	0.67	−0.78	0.86	−0.27	0.78	−0.42	0.72	−0.16	0.36
Road area density (km/km^2) RDARDEN	−0.39	0.90	**−0.74**	**0.85**	0.17	**0.74**	**−0.83**	**0.90**	−0.31	0.77	**−0.46**	0.68	−0.23	0.23
Watershed urban land, distance weighted (%) pURBANdw	**−0.42**	0.90	**−0.80**	0.80	0.15	**0.70**	**−0.82**	0.85	−0.30	0.76	**−0.46**	0.68	**−0.26**	0.26
Impervious surface in the riparian zone (%) P_NLCD1_BIS	**−0.43**	0.76	−0.68	0.81	0.14	0.65	**−0.79**	**0.87**	−0.25	0.75	**−0.44**	0.64	**−0.25**	0.10
Developed land in the riparian zone (%) P_NLCD1_B2	**−0.46**	0.85	−0.73	0.83	0.15	0.67	0.76	0.85	−0.26	0.73	**−0.44**	0.65	−0.23	0.17
Land-cover/landscape variables														
Watershed total forest land (%) p_NLCD1_4	0.27	−0.79	0.69	**−0.85**	−0.02	−0.45	0.29	−0.51	0.29	−0.21	−0.02	−0.35	**−0.42**	−0.20
Watershed total grassland (%) P_NLCD1_7	0.20	−0.79	0.74	−0.78	−0.14	−0.42	0.37	−0.44	0.18	−0.72	0.17	−0.34	0.10	0.14
Watershed total agriculture land (%) P_NLCD1_8	0.34	−0.76	0.60	−0.60	−0.29	−0.38	**0.78**	−0.82	0.18	−0.57	**0.50**	−0.44	**0.44**	0.04
Watershed total wetland (%) P_NLCD1_9	**0.82**	−0.31	0.40	−0.43	0.07	−0.19	0.46	−0.58	0.28	0.23	0.33	−0.29	**0.26**	−0.10
Mean high range organic matter (%) ORMH	0.04	−0.25	**0.79**	−0.45	−0.30	−0.47	−0.32	0.50	−0.02	0.16	−0.15	−0.18	**−0.37**	**−0.44**
Hydrologic soil group, infiltration rate 1–4 mm/h (%) P_HSG_3	na	na	−0.33	0.41	0.09	0.22	0.30	0.21	−0.26	0.06	−0.33	−0.05	−0.15	**−0.42**
Soil organic carbon top 100 cm depth (%) SOC100CM	**0.56**	−0.74	0.67	−0.53	−0.22	−0.22	−0.40	0.52	0.02	0.09	−0.09	−0.23	−0.09	**−0.46**
Soil organic carbon, top 30 cm depth (%) SOC30CM	0.36	−0.50	**0.82**	−0.48	−0.22	−0.28	−0.11	0.20	−0.03	0.14	−0.08	−0.29	−0.11	**−0.54**
Land-cover configuration variables														
Developed largest patch index (%) LPI_C2	**−0.40**	0.89	−0.71	**0.86**	0.14	0.66	−0.77	0.86	−0.24	0.77	−0.40	**0.74**	−0.11	0.36

Table 4 continued

	ATL ($n = 26$)		RAL ($n = 25$)		POR ($n = 23$)		MGB ($n = 22$)		DFW ($n = 25$)		combined5 ($n = 121$)		combined5 $\leq$ 10% imperv ($n = 72$)	
	mdh_95	periodr9	mdh_95	periodr9	mdh_95	periodr9	mdh_95	periodr9	mdh_95	periodr9	mdh_95	periodr9	mdh_95	periodr9
Forest largest patch index (%) LPI_C4	−0.12	0.62	0.62	**−0.84**	−0.05	−0.44	0.22	−0.41	−0.26	−0.30	−0.10	−0.41	**−0.46**	−0.35
Agriculture largest patch index (%) LPI_C8	0.27	−0.72	0.44	−0.64	−0.26	−0.34	**0.82**	−0.73	0.05	−0.31	0.41	−0.51	**0.40**	−0.07
Wetlands largest patch index (%) LPI_C9	**0.71**	−0.18	0.14	−0.35	0.08	0.19	0.31	−0.41	0.37	−0.23	0.24	−0.27	0.19	−0.08
Mean area of developed patches (ha) PAM_C2	**−0.50**	0.68	−0.51	0.64	−0.02	0.28	−0.54	0.75	**−0.58**	0.74	−0.36	**0.76**	−0.01	**0.44**
Mean area of forest patches (ha) PAM_C4	−0.17	−0.58	0.53	−0.69	0.15	−0.38	0.34	−0.51	0.38	−0.52	−0.01	−0.40	**−0.33**	−0.34
Mean area of agriculture patches (ha) PAM_C8	0.33	−0.68	0.66	−0.58	−0.23	−0.57	0.49	−0.52	**0.43**	−0.32	0.38	−0.53	**0.26**	−0.18
Mean area of wetland patches (ha) PAM_C9	**0.75**	−0.26	0.41	−0.37	0.15	0.20	0.46	−0.48	0.21	−0.31	0.32	−0.24	0.16	0.11
Proportion of adjacent developed patches (%) PLA_C2	−0.06	0.84	**−0.75**	0.77	0.11	0.61	**−0.82**	0.79	**−0.43**	**0.81**	−0.39	**0.75**	−0.08	**0.53**
Proportion of adjacent forest patches (%) PLA_C4	0.37	−0.76	0.68	−0.77	0.00	−0.37	0.33	−0.43	0.09	−0.34	−0.05	−0.35	**−0.44**	−0.17
Proportion of adjacent agriculture patches (%) PLA_C8	**0.47**	−0.90	**0.77**	−0.81	0.06	−0.44	0.55	−0.71	0.20	−0.44	0.35	−0.54	**0.26**	−0.25
Proportion of adjacent wetlands patches (%) PLA_C9	**0.79**	−0.34	0.62	−0.42	0.17	0.27	0.67	−0.72	0.36	−0.28	0.36	−0.20	**0.28**	0.07

Bolded coefficients identify variables with stronger correlation than watershed total developed correlation with HCM and minimal $P > 0.05$

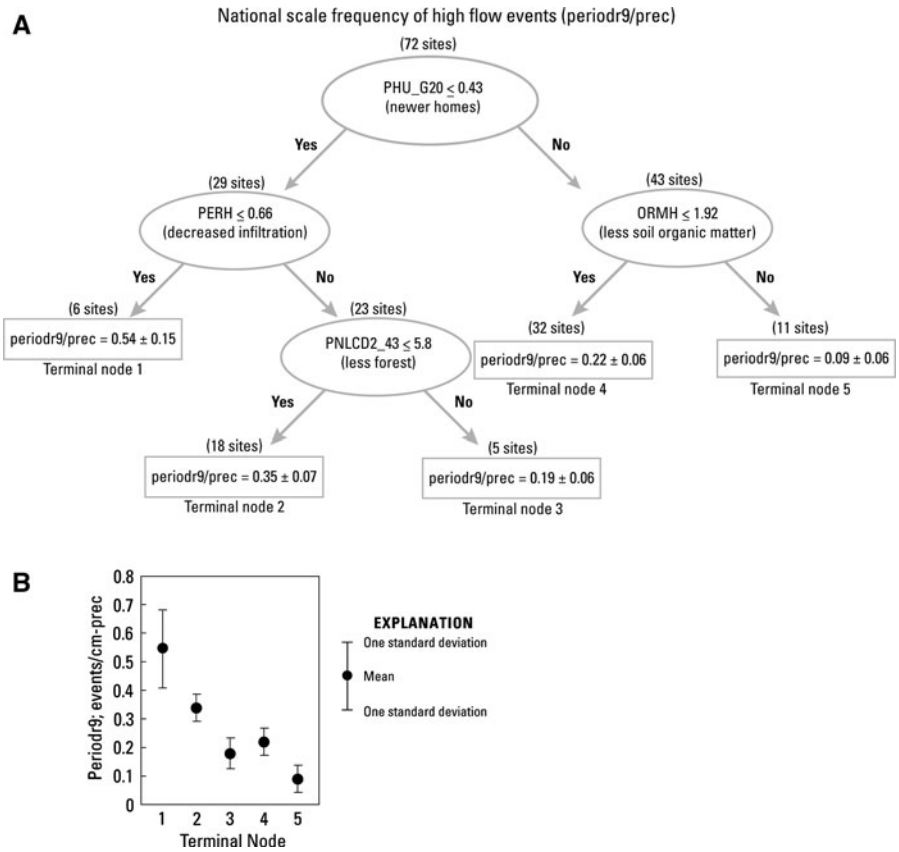


Fig. 10 (**A**) Regression tree model of the influence of land-cover and land-use variables on the precipitation transformed hydrologic condition metric for the 72 sites with watershed imperviousness ≤10% in the national-scale (combined5) data set and (**B**) regression tree terminal node means and standard deviations. Terminal nodes are denoted by a *rectangle* with mean periodr9/prec (±1 SD). The parentheses indicate the number of sites at a node. For example, if PHU_G20 > 0.43 and ORMH ≤ 1.92, then 32 sites meet these criteria with a mean flashiness of 0.22

zone. In a subset of 72 sites with the lowest imperviousness (≤10%), the relations between development (or imperviousness) and high-flow duration and event frequency were much weaker (Table 4). Duration and frequency of high-flow events were weakly correlated to imperviousness ($r_s = -0.20$ and $r_s = 0.29$, respectively). In these watersheds, soil properties, such as organic carbon, organic matter, and soil infiltration rate (SOC100cm, SOC30cm, ORMH, and HSG_3), had stronger relations with periodr9 than percent developed land. A decrease in like adjacencies of development (PLA_C2) or an increase of soil organic matter in the top 30 cm (SOC30cm) had the strongest correlation with improved (reduced) frequency of high-flow events.

The regression tree model for the subset of low impervious sites yielded additional insight into periodr9 (Fig. 10). The primary split variable (node 1) was the portion of housing built before 1979 (PHU_G20 < 0.43). Sites with an increased number of older homes (PHU_G20 > 0.43) were associated with a reduced frequency of high-flow events. This relation may be a reflection of reduced soil-infiltration capacity as a result of inadvertent soil compaction (Gregory et al., 2006), which is more prevalent in recent construction practices (heavy equipment). Sites with highest periodr9 consisted of newer developments and less permeable soils (terminal node 1); the periodr9 value of 0.54 was associated with an EPTr value of ~7 (Fig. 9). At the 23 newer development sites (split for terminal nodes 2 and 3), percent forest in the basin was the split variable that determined more frequent (terminal node 2) and less frequent (terminal node 3) high-flow event sites, a finding consistent with increased infiltration capacity of woodlands. The 11 least flashy sites had older housing developments and increased soil organic matter (terminal node 5); the periodr9 value (0.09) was associated with an EPTr value (~14) roughly twice that of terminal node 1.

Discussion and summary

We used a multifaceted approach, employing biological metrics calculated from algal, invertebrate, and fish community data sets along a gradient of urban intensity, to identify a limited number of biologically relevant HCMs. The biological metrics selected for use in this study are among those used by State and national assessment programs and have been shown in some instances to be regionally applicable. Regression tree models and Spearman rank correlation were used to select a primary set of HCMs (one from each of five flow-component categories) that were most important in the relations with the three stream biological communities.

Although our investigation used short hydrologic periods of records ($\leq$1 year) of hourly cross-sectional area time series collected over a wide range of wet–dry precipitation conditions (Fig. 2) and physiographic regions (Table 1), the resulting HCMs explained up to 73, 92, and 79% of variance for specific algal, invertebrate, and fish community metrics, respectively. In addition, our biologically relevant HCMs often were consistent with results (both in data structure and biological relations) from previous investigations that used long-term discharge records to compute hydrologic metrics, worked with non-urban basins that were often larger in size, and used hydrologic indices based on flow rather than cross-sectional area. For example, Olden & Poff (2003) used PCA to select a subset of hydrologic indices that explained the dominant portion of variation in a set of 171 hydrologic indices using a method that minimized multicollinearity while representing important attributes of the flow regime. Their investigation used long-term (36-year) daily mean streamflow data at 420 sites representing non-urbanized watersheds of $<$5000 km^2 located across a range of climatic and geologic settings extending across the continental United States. They concluded that magnitude of average flow (skew), high-flow magnitude (10% exceedance), and frequency of high-flow event (pulse count) metrics were the most transferable indices between stream types—three indices that explained a substantial portion of the variation for all 171 indices regardless of the stream type. Although our study differed substantially in site selection and type of hydrologic data, our biologically driven (rather than HCM data structure) approach selected similar flow-component metrics—magnitude average flow (skew), high-flow magnitude (max_totfall), and high-flow frequency (periodr9). In another study of 83 small to large (mean flow 0.5–500 m^3/s) streams across New Zealand (Clausen & Biggs, 1997), skew had the strongest correlation to the primary PCA axis of hydrologic metrics, significant negative correlation with periphyton richness and diversity, and positive correlation with invertebrate diversity and richness. In our national investigation, we also found skew to be important (correlation and regression tree models) within all the three biological groups (Fig. 5; Appendix Tables 2 and 3 in Supplementary material). Skew, whether calculated as a third moment (our investigation) or estimated as the ratio of Qmean/Qmedian (Clausen & Biggs, 1997; Olden & Poff, 2003), captured the effects of both high-flow pulses and extreme high-flow magnitude; both components have biological relevance.

We felt it was not necessary (or possible) to identify one or two HCMs applicable to every geographic setting and for all aspects of the aquatic biology of a stream; a position consistent with Konrad & Booth's (2005) view that not one or two aspects of an altered regime cause(s) biological consequence. Selection of our primary HCMs (mdh_95, periodr9, day_pctchg, max_totfall, and skew) was founded on abundance and richness metrics for three taxa groups—biological endpoints often used in the biological condition gradient (Davies & Jackson, 2006), U.S. Environmental Protection Agency's (USEPA), bioassessment criteria (Barbour et al. 1999), and the USEPA Causal Analysis/Diagnosis Decision Information System (CADDIS) framework (U.S. Environmental Protection Agency, 2000). We found consistent tendency between the primary HCMs and indicators of quality stream condition at both the metropolitan area and national scales. Increased duration of high flow or decreased frequency of high flow events, magnitude of high flow, or rate of change was consistently associated with increased fish richness/IBI, increased invertebrate EPTr, or decreased tolerant invertebrate abundance/richness. Our primary HCMs are entirely consistent with a recent finding across 67 small-to-medium upland streams in the Northeastern United States that macroinvertebrate assemblage and biotic decline are related to all five flow components—magnitude, frequency, duration, timing, and rate of

change (Kennen et al., 2009). The primary HCMs identified in this study, while not all encompassing, may have wide application.

With a proliferation of hydrologic indices used in the literature, there is a need to identify a limited number of easily computed HCMs that have biological relevance and apply under a variety of settings (Olden & Poff, 2003). Three of the five primary HCMs (skew, day_pctchg, and max_totfall) and three magnitude of average flow metrics (cv_log, median, and Tqmean) that are readily available from published USGS sources or obtained by simple spreadsheet manipulation (no user-written macro required) have been identified in our study as biologically relevant. In a study of 111 sites across the Western United States (Konrad et al., 2008), median, arguably the most readily available statistic, has been correlated with increased EPTr dominance of fish assemblages by trophic and habitat specialists, lower periphyton density, richness and diversity, and lower invertebrate density and richness (Konrad & Booth, 2005). At 83 river sites across New Zealand, the absolute flow magnitude (as depicted by mean or median not transformed by DA) was significantly related to most of the periphyton and invertebrate variables (Clausen & Biggs, 1997). We had similar, albeit relatively weak, biological response to median cross-sectional area (Fig. 5; Appendix Table 2 in Supplementary material). Tqmean and cv_log, two HCMs associated with biology in our study and previously identified as biologically relevant in studies that used long-term daily streamflow statistics (Konrad & Booth, 2005; Konrad et al., 2008), also are easily computed. A rate-of-change metric was determined to have negative correlation with invertebrate abundance, non-insect and intolerant richness in a multiregional investigation in the Western United States (Konrad et al., 2008), and with insectivore fish in 33 streams across the Tennessee River Valley (Knight et al., 2008). Our simply computed rate-of-change metric (day_pctchg), based on daily cross-sectional area data, had substantial correlation with biological metrics from all the three of our taxa groups (Fig. 5) as well as a dominant metric representing overall HCM data structure (PCA axis 1, Fig. 6). In addition, day_pctchg was highly correlated to the Richard-Baker flashiness index, a metric used to quantify flow regime change in 515 Midwestern United States streams (Baker et al., 2004).

We concur with the Clausen & Biggs (1997) suggestion that flood-disturbance frequency is the most *overall* ecologically useful flow variable. In our study with respect to HCM data structure, periodr9 was located in a position ~45° between the two primary PCA axes (Fig. 6), a compromise position that represented variables dominant in the HCM overall structure. The high-flow frequency metric was located in a similar position (quadrant between PCA axis 1 and PCA axis 2) in studies in both New Zealand and the continental United States (Clausen & Biggs, 1997; Olden & Poff, 2003). Flood disturbance has been widely viewed as the principal axis of conceptual models of stream benthic community structure (Clausen & Biggs, 1997). Across New Zealand, the flood-disturbance frequency index was determined to have significant relations with four of the six periphyton and invertebrate variables (Clausen & Biggs, 1997). High-flow frequency also has been associated with changes in fish assemblage (Konrad & Booth, 2005), reduced insectivore fish (Knight et al., 2008) and reduced biotic index and clinger invertebrates in King County, western Washington (Cassin et al., 2005). Increased frequency of high flows also is associated with increased richness of tolerant invertebrate taxa and decreased EPT richness in Northeastern United States upland streams (Kennen et al., 2009). Our frequency of high-flow event metric (periodr9), although area based and derived from short data record, had similar associations across the United States. Increased periodr9 values were associated with reduced EPTr (or increased RichTol) in the five metropolitan areas along with decreased fish metrics (IBI, Richness, or TotAbund) in four metropolitan areas (Fig. 7). The high-flow frequency metric, with minor precipitation adjustment, accounted for 24% of the EPTr on a national scale (Fig. 9). We did not feel that nine times the median change was a definitive value but rather indicated general importance of a metric that represented substantial flow change. Periodr9 was highly correlated to periodr7 (removed during our data reduction) and similar to the retained variable periodf5 (Fig. 6). In an assessment of high-flow frequency, Konrad & Booth (2005) found that use of a threshold of events greater than the 10th percentile was less sensitive to climatic conditions than a threshold of three times the median flow. The threshold for our frequency of high-flow event variable (periodr9) was based on a median change (Appendix

Table 1 in Supplementary material) rather than overall median value (of the area time series); therefore, it was not directly comparable to many of the previously published metrics. The finding of transferability across stream types (Olden & Poff, 2003), biological relevance in previous investigations, simple relation to all taxa in the five metropolitan areas in this study, and high occurrence of periodr9 in our regression tree models underscored the importance of the frequency of high-flow event metric. We did not, however, discriminate the frequency of high-flow events from stressors (habitat, temperature, and chemistry) commonly associated with urbanization.

We found that the flashiness metric (periodr9) had significant response to urbanization and imperviousness in each of the five metropolitan areas, a finding consistent with McMahon et al. (2003), who determined that stage-based flashiness responded to urbanization. Relations between imperviousness and periodr9 were strongest in individual metropolitan areas; however, transformation of periodr9 with mean precipitation for the metropolitan area resulted in a geographically generalized linear model ($r^2 = 0.48$) for the five metropolitan areas. Formation of the relation from hydrologic records collected non-concurrently in five metropolitan areas representing a range of climatic and physiographic settings (Table 1) and during dry (POR) and wet (MGB, RAL, and ATL) periods suggests broad applicability. Recent studies suggest that directly connected imperviousness has stronger association with stream alteration, algal, invertebrate, and fish assemblage integrity than simple total imperviousness (Wang et al., 2001; Taylor et al., 2004; Walsh, 2004; Walsh et al., 2005a; Roy & Shuster, 2009). It is likely, therefore, that high-flow event frequency (periodr9) response to watershed imperiousness would have been stronger, had our data set permitted determination of connected imperviousness (Fig. 3).

Each metropolitan area had significant relations ($P < 0.05$) between HCMs and metrics chosen from each of the three groups of aquatic communities; however, the strongest and greatest number of relations were in metropolitan areas (ATL, RAL, and POR) in which non-urban sites consisted of forest rather than agricultural land use. Higher correlations were observed in the invertebrate community, mainly with HCMs (periodr9, periodf5) that described the frequency of high-flow events (Fig. 5). Locally derived fish IBI scores were correlated with metrics related to magnitude of average flow (ATL), duration of high flow (ATL, MGB, and RAL), high-flow event frequency (ATL, MGB, and DFW), frequency of small flow pulses (MGB), and rate of change (ATL) and indicated potential issues related to the importance of a stream's hydrologic context when deriving and reporting fish IBI scores as part of a water-quality assessment (Appendix Table 2 in Supplementary material). A national-scale EPTr relation that used precipitation-transformed periodr9 ($P < 0.01$, $r^2 = 0.24$) indicated that sites with highest EPTr had reduced frequency of high-flow events (periodr9 < 0.4 events/cm, Fig. 9).

Cross-validated regression tree models that used a limited number of HCMs (1–4) to predict biological metrics accounted for a substantial amount of variation and, in some cases, were highly predictive despite substantial inherent variability in biological data collection and determination of HCMs. Regression tree models in individual metropolitan areas accounted for as much as 73, 92, and 79% of the variance for algae, invertebrates, and fish, respectively. Predictive regression tree models were substantially stronger than the bivariate HCMs to biological relations, often incorporating HCMs from several flow components. On a national scale, although biological sampling occurred over a 4-year period and encompassed sites with a wide range of DA, precipitation, and elevation, the predictive models accounted for substantial variance. Incorporation of contextual variables often strengthened the predictive models.

Although our HCMs were computed from cross-sectional area time series rather than flow data, our results nevertheless may have wide applicability. Investigators in possession of cross-sectional area data (or stage data readily converted to area) can use our relations and models directly. Practitioners with access to stream-discharge data may use our predictive models and high-flow event frequency (periodr9) relations as a guide for formulating their discharge-based HCM relations. For example, based on Fig. 8A, the magnitude of high-flow (max_totfall) and rate-of-change (cumm_chg) metrics could be computed from the discharge data to develop an algal (eutrophic) model in POR (or a setting similar to POR). The rate-of-change (day_pctchg) and duration of high-flow (mxh_95) metrics could be the basis of an EPTr model

in ATL (Appendix Table 3 in Supplementary material). The rate-of-change metric (day_pctchg) could be calculated simply with daily-flow data and spreadsheet to reflect relations in ATL, RAL, POR, and MGB (Appendix Table 2 in Supplementary material). From preliminary work in RAL, discharge-based HCMs may have stronger relations with stream invertebrate communities and basin imperviousness than cross-sectional-based HCMs (data not shown).

Extensive study has shown that land-cover composition can account for much of the variability in stream ecological conditions; however, few studies have considered the relation between land-cover configuration and ecological conditions (Alberti et al., 2007; Schueler et al., 2009). We found numerous land-cover patterns that were associated strongly with two biologically relevant HCMs (mdh_95, periodr9) that characterize the natural flow regime and quantify flashiness (Baker et al., 2004; Walsh et al., 2005b). Although mdh_95 and periodr9 were related to imperviousness, the two HCMs had stronger responses to several land-cover characteristic and pattern variables, an indication that management options in addition to wholesale impervious-surface reduction may be possible. Increased overall percent wetland and adjacent patches of wetland (ATL) and distance between developed land and stream (ATL, RAL, and MGB) were examples of land-cover patterns more strongly associated with improved HCMs than overall percent development or imperviousness in the basin. In each of the individual metropolitan areas, periodr9 had strong association to density of roads, a finding consistent with previous investigations—roads are an important stressor in urbanizing landscapes (Jones et al., 2000; Trombulak & Frissell, 2000; Alberti et al., 2007). In our national data set (combined5), reduction of road density, developed land in the riparian zone, urban patch size, connectivity, and like adjacency were all more strongly associated with improved (increased) high-flow duration (mdh_95) and decreased frequency of high-flow events (periodr9) than basin imperviousness (Table 4). Our HCM response to land cover was consistent with a finding of strong correlations between the same land-cover pattern metrics (density of roads; percent wetland in the basin and wetland-patch characteristics) and invertebrate metrics in an analysis of the nine metropolitan areas (Cuffney et al., 2008). In the more geographically focused Puget Sound study, Alberti et al. (2007) concluded that reduction of these same urban pattern variables was strongly related to increased benthic integrity index in 42 subbasins in the Puget Sound lowland region (Alberti et al., 2007). Our study of five metropolitan areas provided national-scale concurrence with studies by King et al. (2005), Walsh et al. (2005b), and Alberti et al. (2007) that spatial configuration of urban land use is relevant to aquatic ecosystems, and effects of urbanization on stream flashiness may be mitigated when developed land patches are intermixed across the basin rather than agglomerated into large patches (McMahon et al., 2003) in addition to drainage infrastructure consideration (Walsh et al., 2005a).

Many investigators and researchers have recommended a degradation threshold of 10% impervious surface cover but suggested that poorly designed urban areas may impact aquatic systems at levels much less than 10% imperviousness (Walsh et al., 2005b). In an analysis of nine metropolitan areas, Cuffney et al. (2008) determined that at 10% impervious land cover, the invertebrate assemblages had undergone about one-third of the change that occurred over the entire urban gradient, a finding that underscored the need to understand underlying mechanisms of specific stressors (such as frequency of high-flow events) in areas of low imperviousness. Although duration and frequency of high flow had substantial relations to overall watershed imperviousness in our national data set (combined5), we found low correlation between the two HCMs (mdh_95, periodr9) and imperviousness in the 72 low (<10%) impervious sites (Table 4). At sites of low imperviousness, soil characteristics had greater association with the periodr9 than imperviousness. This suggests that, in lesser-developed watersheds (≤10% imperviousness), in addition to consideration of impervious connectedness, it may be especially beneficial to preserve areas with high-infiltration soils; whereas in developed watersheds, the beneficial hydrologic effects of high-infiltration soils may be negated by increased runoff from impervious surfaces. Previous investigators have suggested that impervious cover may not be the best metric to predict stream-quality indicator for sites with less than 10% impervious cover and that forest may have an increased role in these watersheds (Schueler et al., 2009). Our national-scale regression tree model

offered initial confirmation that age of development (surrogate for soil compaction and reduced infiltration), soil properties, and percent forest were more important in delineating quality hydrologic conditions (and potentially better EPTr values) than overall percent imperviousness in lesser-developed sites (Fig. 10).

Previous investigators have identified the need for stream ecology to help guide urban land-use development (Feminella & Walsh, 2005). For example, to minimize fish assemblage alteration in urban streams in Georgia, Roy et al. (2005) recommended ordinances be adopted to reduce the frequency and magnitude of high-flow events, especially during low to medium precipitation events. With increased understanding, HCM–biology relations may provide one link between watershed change (or mitigation) and stream biology. Deterministic rainfall runoff models have been used to predict hydrograph and HCM change resulting from proposed watershed development scenarios (Cassin et al., 2005); however, these models can be resource- and time-intensive. Our exploratory findings between readily available GIS land-cover characteristics and biologically relevant HCMs offer initial support for a potentially less costly approach. The Ecological Limits of Hydrologic Alteration (ELOHA), a framework published by an international group of river scientists, recommended use of scientifically defined flow-ecology links in an adaptive management context that included outcomes such as channel restoration and fishery management (Poff et al., 2009). Although few studies have made the definitive link between hydrologic alteration and specific biological responses (Konrad & Booth, 2005), results from our national-scale urban gradient study provide examples of possible links and are consistent with and contribute to the recommended ELOHA and CADDIS (U.S. Environmental Protection Agency, 2000) frameworks.

Use of a limited number of HCMs, along with other important stressors, may lead to development of geographically generalized biological models. Development of models that account for pre-urbanized disturbance and also include stream stressors, such as chemistry, habitat, and temperature, in addition to HCMs may result in robust connection between stressors and stream biology. Although elusive, mechanistic understanding of the links between urbanization and stream ecosystem degradation is essential if ecologists hope to understand the nature of ecological response to disturbance and the desire to contribute to the development of scenarios that can guide planning decisions (Paul & Meyer, 2001). At a time when development on the urban fringe is inevitable, planners need to look to design urban land in ways that will minimize the effects on ecological systems by mimicking natural processes or allowing natural processes to occur (Lloyd et al., 2002). Understanding at the stressor level as outlined in conceptual frameworks, such as Walsh et al. (2005b) and CADDIS model (U.S. Environmental Protection Agency, 2000), may lead to development decisions and improved management actions. Relations developed on a regional or metropolitan-area scale are likely to be stronger than geographically generalized models; however, broad spatial models are able to provide much needed understanding in unmonitored areas.

Acknowledgments Numerous USGS scientists and technicians collected, compiled, and managed the large amount of data used in this article. Forrest Bilek and Austin Baldwin provided substantial assistance in data analysis. Laura Nelson, Rebecca Carvin, and Jeffrey Corbett provided illustration and manuscript assistance. Kay Hedrick, Rebecca Deckard, Jerry McMahon, Rodney Knight, and Timothy Fletcher provided invaluable review of the draft manuscript. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U. S. Government.

References

Alberti, M., D. Booth, K. Hill, B. Coburn, C. Avolio, S. Coe & D. Spirandelli, 2007. The impact of urban patterns on aquatic ecosystems: an empirical analysis in Puget lowland sub-basins. Landscape and Urban Planning 80: 345–361.

Arnold, C. L. & C. J. Gibbons, 1996. Impervious surface coverage: the emergence of a key environmental indicator. Journal of the American Planning Association 62: 243–258.

Bahls, L. L., 1993. Periphyton Bioassessment Methods for Montana Streams. Water Quality Bureau, Department of Health and Environmental Sciences, Helena, MT.

Baker, D. B., R. P. Richards, T. T. Loftus & J. W. Kramer, 2004. A new flashiness index: characteristics and applications to midwestern rivers and streams. Journal of the American Water Resources Association 40: 503–522.

Barbour, M. T., J. Gerritsen, B. D. Snyder & J. B. Stribling, 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, 2nd ed. EPA 841-B-99-002. U.S.

Environmental Protection Agency, Office of Water, Washington, DC.

Booth, D. B. & C. R. Jackson, 1997. Urbanization of aquatic systems: degradation thresholds, stormwater detection, and the limits of mitigation. Journal of the American Water Resources Association 33: 1077–1090.

Booth, D. B., D. Hartley & R. Jackson, 2002. Forest cover, impervious-surface area, and the mitigation of stormwater impacts. Journal of the American Water Resources Association 38: 835–845.

Booth, D. B., J. R. Karr, S. Schauman, C. P. Konrad, S. A. Morley, M. G. Larson & S. J. Burges, 2004. Reviving urban streams: land use, hydrology, biology, and human behavior. Journal of the American Water Resources Association 40: 1351–1364.

Breiman, L., J. H. Friedman, R. A. Olshen & C. J. Stone, 1984. Classification and Regression Trees. Chapman and Hall, New York.

Cablk, M., D. White & A. R. Kiester, 2002. Assessment of spatial autocorrelation in empirical models in ecology. In Scott, J. M., P. J. Heglund & M. L. Morrison (eds), Morrison Predicting Species Occurrences: Issues of Accuracy and Scale. Island Press, Washington DC: 429–440.

Cassin, J., H. Berge, J. Burkey, R. Fuerstenberg, B. Murray, D. St. John, L. Tear & K. Whiting, 2005. Hydrological and biological indicators of flow alteration in urban streams. In Proceedings of the AWRA Annual Water Resources Conference, Seattle, Washington, November 7–10, 2005.

Clausen, B. & B. J. F. Biggs, 1997. Relationships between benthic biota and hydrological indices in New Zealand streams. Freshwater Biology 38: 327–342.

Clausen, B. & B. J. F. Biggs, 2000. Flow variables for ecological studies in temperate streams: groupings based on covariance. Journal of Hydrology 237: 184–197.

Coles, J. F., T. F. Cuffney, G. McMahon & K. M. Beaulieu, 2004. The effects of urbanization on the biological, physical, and chemical characteristics of coastal New England streams. U. S. Geological Survey Professional Paper 1695.

Coles, J. F., A. H. Bell, B. C. Scudder & K. D. Carpenter, 2009. The effects of urbanization and other environmental gradients on algal assemblages in nine metropolitan areas across the United States. U. S. Geological Survey Scientific Investigations Report 2009-5022.

Cuffney, T. F. & J. F. Falcone, 2009. Derivation of nationally consistent indices representing urban intensity within and across nine metropolitan areas of the conterminous United States. U. S. Geological Survey Scientific Investigations Report 2008-5095.

Cuffney, T. F., M. E. Gurtz & M. R. Meador, 1993. Methods for collecting benthic invertebrate samples as part of the National Water-Quality Assessment Program: U. S. Geological Survey Open-File Report 93-406.

Cuffney, T. F., G. McMahon, I. A. Waite, J. T. May & R. A. Brightbill, 2008. Comparison of benthic macroinvertebrate response models to urban intensity in nine metropolitan areas of the conterminous United States. In NABS 56th Annual Meeting, Salt Lake City, Utah, May 25–28, 2008.

Cummins, K. W., 1973. Trophic relations of aquatic insects. Annual Review of Entomology 18: 183–206.

Data Description Inc, 1996. Data Desk, Version 6.1. Data Description, Inc, Ithaca, NY.

Davies, S. P. & S. K. Jackson, 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecological Applications 16: 1251–1266.

De'ath, G. & K. E. Fabricius, 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81: 3178–3192.

Efron, B. & R. Tibshirani, 1991. Statistical-data analysis in the computer-age. Science 253: 390–395.

Falcone, J., J. Stewart, S. Sobieszcyk, J. Dupree, G. McMahon & G. Buell, 2007. A comparison of natural and urban characteristics and the development of urban intensity indices across six geographic settings. U. S. Geological Survey Scientific Investigations Report 2007-5123.

Feminella, J. W. & C. J. Walsh, 2005. Urbanization and stream ecology: an introduction to the series. Journal of the North American Benthological Society 24: 585–587.

Giddings, E. M. P, A. H. Bell, K. M. Beaulieu, T. F. Cuffney, J. F. Coles, L. R. Brown, F. A. Fitzpatrick, J. Falcone, L. A. Sprague, W. L. Bryant, M. C. Peppler, C. Stephens & G. McMahon, 2009. Selected physical, chemical, and biological data used to study urbanizing streams in nine metropolitan areas of the United States, 1999-2004. U. S. Geological Survey Data Series 423.

Goldstein, R. M. & M. R. Meador, 2005. Multilevel assessment of fish species traits to evaluate habitat degradation in streams of the upper midwest. North American Journal of Fisheries Management 25: 180–194.

Gregory, J. H., M. D. Dukes, P. H. Jones & G. L. Miller, 2006. Effect of urban soil compaction on infiltration rate. Journal of Soil and Water Conservation 61: 117–124.

Jones, J. A., F. J. Swanson, B. C. Wemple & K. U. Snyder, 2000. Effects of roads on hydrology, geomorphology, and disturbance patches in stream networks. Conservation Biology 14: 76–85.

Kennard, M. J., B. D. Harch, B. J. Pusey & A. H. Arthington, 2006. Accurately defining the reference condition for summary biotic metrics: a comparison of four approaches. Hydrobiologia 572: 151–170.

Kennen, J. G., K. Riva-Murray & K. M. Beaulieu, 2009. Determining hydrologic factors that influence stream macroinvertebrate assemblages in the northeastern US. Ecohydrology. doi:10.1002/eco.99.

King, R. S., M. E. Baker, D. F. Whigham, D. E. Weller, T. E. Jordan, P. F. Kazyak & M. K. Hurd, 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. Ecological Applications 15: 137–153.

Knight, R. R., M. B. Gregory & A. K. Wales, 2008. Relating streamflow characteristics to specialized insectivores in the Tennessee River Valley: a regional approach. Ecohydrology 1: 394–407.

Konrad, C. P. & D. B. Booth, 2005. Hydrologic changes in urban streams and their ecological significance. In Brown L. R., R. H. Gray, R. H. Hughes & M. R. Meador (eds), Effects of Urbanization on Stream Ecosystems. American Fisheries Society, Symposium 47, Bethesda, MD: 157–177.

Konrad, C. P., A. M. D. Brasher & J. T. May, 2008. Assessing streamflow characteristics as limiting factors on benthic

invertebrate assemblages in streams across the western United States. Freshwater Biology 53: 1983–1998.

Lange-Bertalot, H., 1979. Pollution tolerance of diatoms as a criterion for water quality estimation. Nova Hedwigia 64: 285–304.

Lloyd, S. D., T. H. F. Wong & B. Porter, 2002. The planning and construction of an urban stormwater management scheme. Water Science and Technology 45: 1–10.

McMahon, G., J. D. Bales, J. F. Coles, E. M. P. Giddings & H. Zappia, 2003. Use of stage data to characterize hydrologic conditions in an urbanizing environment. Journal of the American Water Resource Association 39: 1529–1546.

Moulton, S. R., II, J. G. Kennen, R. M. Goldstein & J. A. Hambrook, 2002. Revised protocols for sampling algal, invertebrate, and fish communities as part of the National Water-Quality Assessment Program. U. S. Geological Survey Open-File Report 02-150.

Olden, J. D. & N. L. Poff, 2003. Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. River Research and Applications 19: 101–121.

Paul, M. J. & J. L. Meyer, 2001. Streams in the urban landscape. Annual Review of Ecology and Systematics 32: 333–365.

Poff, N. L. & J. D. Allan, 1995. Functional-organization of stream fish assemblages in relation to hydrological variability. Ecology 76: 606–627.

Poff, N. L., B. Richter, A. H. Arthington, S. E. Bunn, R. J. Naiman, E. Kendy, M. Acreman, C. Apse, B. P. Bledsoe, M. Freeman, J. Henriksen, R. B. Jacobson, J. Kennen, D. M. Merritt, J. O'Keeffe, J. Olden, K. Rogers, R. E. Tharme & A. Warner, 2009. The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards. Freshwater Biology 55: 147–170.

Porter, S. D., 2008. Algal attributes: an autecological classification of algal taxa collected by the National Water-Quality Assessment Program. U. S. Geological Survey Data Series 329.

Richter, B. D., J. V. Baumgartner, J. Powell & D. P. Braun, 1996. A method for assessing hydrologic alteration within ecosystems. Conservation Biology 10: 1163–1174.

Roy, A. H. & W. D. Shuster, 2009. Assessing impervious surface connectivity and applications for watershed management. Journal of the American Water Resources Association 45: 198–209.

Roy, A. H., M. C. Freeman, B. J. Freeman, S. J. Wenger, W. E. Ensign & J. L. Meyer, 2005. Investigating hydrologic alteration as a mechanism of fish assemblage shifts in urbanizing streams. Journal of the North American Benthological Society 24: 656–678.

Schueler, R. T., L. Fraley-McNeal & K. Cappiella, 2009. Is impervious cover still important? Review of recent research. Journal of Hydrologic Engineering 14: 309–315.

Steen, P. J., T. G. Zorn, P. W. Seelbach & J. S. Schaeffer, 2008. Classification tree models for predicting distributions of Michigan stream fish from landscape variables. Transactions of the American Fisheries Society 137: 976–996.

Steinberg, D. & P. Colla, 1997. CART – Classification and Regression Trees. Salford Systems, San Diego, CA.

Tate, C. M., T. F. Cuffney, G. McMahon, E. M. P. Giddings, J. F. Coles & H. Zappia, 2005. Use of an urban intensity index to assess urban effects on streams in three contrasting environmental settings. In Brown, L. R., R. H. Gray, R. M. Hughes & M. R. Meador (eds), Effects of Urbanization on Stream Ecosystems. American Fisheries Society Symposium 47, Bethesda, MD: 291–315.

Taylor, S. L., S. C. Roberts, C. J. Walsh & B. E. Hatt, 2004. Catchment urbanization and increased benthic algal biomass in streams: linking mechanisms to management. Freshwater Biology 49: 835–851.

Trombulak, S. C. & C. A. Frissell, 2000. Review of ecological effects of roads on terrestrial and aquatic communities. Conservation Biology 14: 18–30.

U. S. Environmental Protection Agency, 2000. Stressor Identification Guidance Document (CADDIS model). EPA 822/B-00/025. U.S. Environmental Protection Agency, Office of Research and Development, Washington DC.

van Dam, H., A. Mertens & J. Sinkeldam, 1994. A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. Netherlands Journal of Aquatic Ecology 28: 117–133.

Walsh, C. J., 2004. Protection of in-stream biota from urban impacts: minimize catchment imperviousness or improve drainage design? Marine & Freshwater Research 55: 317–326.

Walsh, C. J., T. D. Fletcher & A. R. Ladson, 2005a. Stream restoration in urban catchments through redesigning stormwater systems: looking to the catchment to save the stream. Journal of the North American Benthological Society 24: 690–705.

Walsh, C. J., A. H. Roy, J. W. Feminella, P. D. Cottingham, P. M. Groffman & R. P. Morgan, 2005b. The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society 24: 706–723.

Wang, L., J. Lyons, P. Kanehl & R. Bannerman, 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management 28: 255–266.

Appendix Table 1. Definitions and summary statistics of hydrologic condition metrics (HCM) based on continuous hourly cross-sectional area values used in analysis. Statistics are for five metropolitan areas combined (ATL, RAL, DFW, POR, and MGB). Variables retained for analysis in the reduced data sets are denoted by bold and shaded text. Flow components as defined in Olden & Poff (2003)

Variable	Definition	Unit	Max	Mean	Median	Min
	Magnitude of average flow condition					
cv	coefficient of variation of values	unitless	377.95	63.30	50.05	3.63
skew	skew of values	unitless	24.71	5.71	4.88	0.61
cv_log[1,3]	Coefficient of variation of values, values are equal to log(1+value)	unitless	124.62	22.02	17.55	0.72
coef_disp	(75th percentile value - 25th percentile value)/median value	unitless	4.00	0.41	0.33	0.02
mean	mean value	m2	146.33	13.73	4.37	0.57
median	median (50th percentile) value	m2	142.40	13.07	3.30	0.20
Tqmean	fraction of time value exceeded mean value	unitless	0.68	0.32	0.32	0.12
mean/DA	mean value, divided by drainage area	m2/km2	3.58	0.24	0.09	0.02
median/DA	median (50th percentile) area value , divided by drainage area	m2/km2	3.48	0.23	0.07	0.01
cv*DA	cv times drainage area	km2	19780	3478	1758	159
skew*DA	skew times drainage area	km2	2772	329	223	20
cv_log*DA	cv_log times drainage area	km2	5046	1081	704	45
coef_disp*DA	coef_disp times drainage area	km2	127	22	13	1
	Magnitude of low flow condition					
pct_25a	25th percentile value	m2	139.40	12.33	2.80	0.00
pct_10a	10th percentile value	m2	138.90	11.74	2.30	0.00
pct_5a	5th percentile value	m2	138.60	11.53	2.00	0.00
pct_25n	25th percentile value, divided by median value	unitless	0.99	0.84	0.87	0.00
pct_10n	10th percentile value divided by median value	unitless	0.99	0.74	0.77	0.00
pct_5n	5th percentile value, divided by median value	unitless	0.99	0.70	0.73	0.00
	Magnitude of high flow condition					
pct_99a	99th percentile value	m2	192.40	27.82	13.00	1.80
pct_95a	95th percentile value	m2	180.30	18.01	7.60	1.10
pct_90a	90th percentile value	m2	149.50	15.54	6.10	0.80
pct_75a	75th percentile value	m2	146.30	13.93	4.10	0.60
pct_99n	99th percentile value divided by median value	unitless	49.88	5.00	3.45	1.06
pct_95n	95th percentile value divided by median value	unitless	10.00	2.31	2.00	1.03
pct_90n	90th percentile value, divided by median value	unitless	7.00	1.72	1.59	1.01
pct_75n	75th percentile value, divided by median value	unitless	4.00	1.26	1.20	1.01
max_totrise	The maximum change in value during one rising period	m2	270.30	40.93	25.10	0.80
max_totfall	The maximum change in value during one falling period	m2	298.90	38.35	24.20	0.60
max_totrise/DA	max_totrise divided by drainage area	m2/km2	15.89	1.01	0.55	0.03
max_totfall/DA	max_totfall divided by drainage area	m2/km2	15.81	0.96	0.48	0.01
	Frequency of small flow pulses					
periodr1[1]	Frequency of rising events, where total rise is greater than or equal to 1 times the median total rise (number of rises)	event/month	154.47	65.78	65.50	11.00
periodf1[1]	Frequency of falling events, where total rise is greater than or equal to 1 times the median total rise (number of rises)	event/month	160.47	81.60	80.05	25.85
periodr1/DA[1]	periodr1/(mth) divided by drainage area	event/month/km2	13.79	2.00	1.38	0.16
periodf1/DA[1]	periodf1/(mth) divided by drainage area	event/month/km2	16.16	2.39	1.66	0.39
	Frequency of high flow events					
periodr3[1]	Frequency of rising events, where total rise is greater than or equal to 3 times the median total rise (number of rises)	event/month	16.52	7.28	6.53	0.65
periodr5[1]	Frequency of rising events, where total rise is greater than or equal to 5 times the median total rise (number of rises)	event/month	14.14	5.14	4.59	0.44
periodr7[1]	Frequency of rising events, where total rise is greater than or equal to 7 times the median total rise (number of rises)	event/month	13.32	4.28	3.87	0.22
periodr9[1]	Frequency of rising events, where total rise is greater than or equal to9 times the median total rise (number of rises)	event/month	11.84	3.78	3.45	0.00
periodf3[1]	Frequency of falling events, where total fall is greater than or equal to 3 times the median total fall (number of falls)	event/month	21.70	7.81	7.07	0.22
periodf5[1]	Frequency of falling events, where total fall is greater than or equal to 5 times the median total fall (number of falls)	event/month	13.64	4.63	4.11	0.11
periodf7[1]	Frequency of falling events, where total fall is greater than or equal to 7 times the median total fall (number of falls)	event/month	11.51	3.65	3.27	0.00

Variable	Definition	Unit	Max	Mean	Median	Min
periodf9[1]	Frequency of falling events, where total fall is greater than or equal to9 times the median total fall (number of falls)	event/month	10.68	3.18	2.66	0.00
periodr3/DA[1]	periodr3/(mth) divided by drainage area	event/month/km2	2.08	0.25	0.12	0.01
periodr5/DA[1]	periodr5/(mth) divided by drainage area	event/month/km2	1.65	0.18	0.08	0.01
periodr7/(mth)/DA[1]	periodr7/(mth) divided by drainage area	event/month/km2	1.45	0.15	0.06	0.00
periodr9/DA[1]	periodr9/(mth) divided by drainage area	event/month/km2	1.37	0.14	0.05	0.00
periodf3/DA[1]	periodf3/(mth) divided by drainage area	event/month/km2	1.96	0.25	0.12	0.01
periodf5/DA[1]	periodf5/(mth) divided by drainage area	event/month/km2	1.59	0.16	0.07	0.00
periodf7/DA[1]	periodf7/(mth) divided by drainage area	event/month/km2	1.39	0.13	0.05	0.00
periodf9/DA[1]	periodf9/(mth) divided by drainage area	event/month/km2	1.30	0.12	0.04	0.00
	Duration of low flow events					
mxl_25	Duration of longest pulse less than the 25th percentile value	hours	1168	355.91	332	0
mxl_10[1,3]	Duration of longest pulse less than the 10th percentile value	hours	526	146.65	123	0
mxl_5	Duration of longest pulse less than the 5th percentile value	hours	329	71.85	54	0
mdl_25[1,3]	Median duration of low pulses less than the 25th percentile value	hours	183	7.8	3.0	0
mdl_10	Median duration of low pulses less than the 10th percentile value	hours	271	6.9	2.5	0
mdl_5	Median duration of low pulses less than the 5th percentile value	hours	69	5.0	2.5	0
	Duration of high flow events					
mxh_75[1,2,4]	Duration of longest pulse greater than the 75th percentile value	hours	1220	330.7	280	53
mxh_90[1,2,4]	Duration of longest pulse greater than the 90th percentile value	hours	480	126.6	108	29
mxh_95[1,2,4]	Duration of longest pulse greater than the 95th percentile value	hours	249	71.2	64	19
mdh_75[1,3]	Median duration of high pulses greater than the 75th percentile value	hours	365	13.2	6	1
mdh_90[1,3]	Median duration of high pulses greater than the 90th percentile value	hours	148	17.3	10	1
mdh_95[1,3]	Median duration of high pulses greater than the 95th percentile value	hours	63	14.1	10	1
mxh_75/DA	mxh_75 divided by drainage area	hours/km2	112.29	10.11	6.05	0.56
mxh_90/DA	mxh_90 divided by drainage area	hours/km2	17.54	3.73	2.19	0.40
mxh_95/DA	mxh_95 divided by drainage area	hours/km2	9.60	1.97	1.25	0.19
mdh_75/DA	mdh_75 divided by drainage area	hours/km2	4.54	0.28	0.17	0.01
mdh_90/DA	mdh_90 divided by drainage area	hours/km2	1.96	0.37	0.25	0.01
mdh_95/DA	mdh_95 divided by drainage area	hours/km2	1.22	0.31	0.20	0.02
max_durrise	maximum duration of consecutive periods of rising values	hours	32	12.3	11	4
med_durrise	median duration of consecutive periods of rising values	hours	1	1.0	1	1
max_durfall	maximum duration of consecutive periods of falling values	hours	69	22.7	20	3
med_durfall	median duration of consecutive periods of falling values	hours	1	1.0	1	1
max_durrise/DA	max_durrise divided by drainage area	hours/km2	1.42	0.31	0.23	0.05
max_durfall/DA	max_durfall divided by drainage area	hours/km2	3.05	0.57	0.42	0.03
	Rate of change					
cumm_chg	Sum of the absolute value of the total rise and fall values	m2	3309.10	854.59	611.20	88.80
cumm_median	Sum of the absolute value of the total rise and fall values , divided by median value	unitless	3636.00	332.76	183.30	4.46
day_pctchg	Sum of the absolute value of the relative change in daily values, SUM (|value(t) - value(t-1)|/value(t)), where value(t) is the daily mean value on day t	unitless *Based on daily flow	334.33	41.20	27.19	0.94
rb_flash[1,3]	Version of Richards-Baker flashiness index, = sum of the absolute values of change in mean daily value, divided by the sum of the daily mean values for the POR; (SUM|value(t) - value(t-1)|) / (SUM value(t)), where value(t) is the daily mean value on day t	unitless *Based on daily flow	0.59	0.18	0.12	0.01
cumm_chg/DA	cumm_chg divided by drainage area	m2/ km2	286.71	24.87	11.83	2.22
cumm_median*DA	cumm_median times drainage area	km2	101143	13897	7040	182
day_pctchg*DA	day_pctchg times drainage area	km2	12781.46	2196.99	819.52	38.36
rb_flash*DA	rb_flash times drainage area	km2	53.04	10.18	4.29	0.31

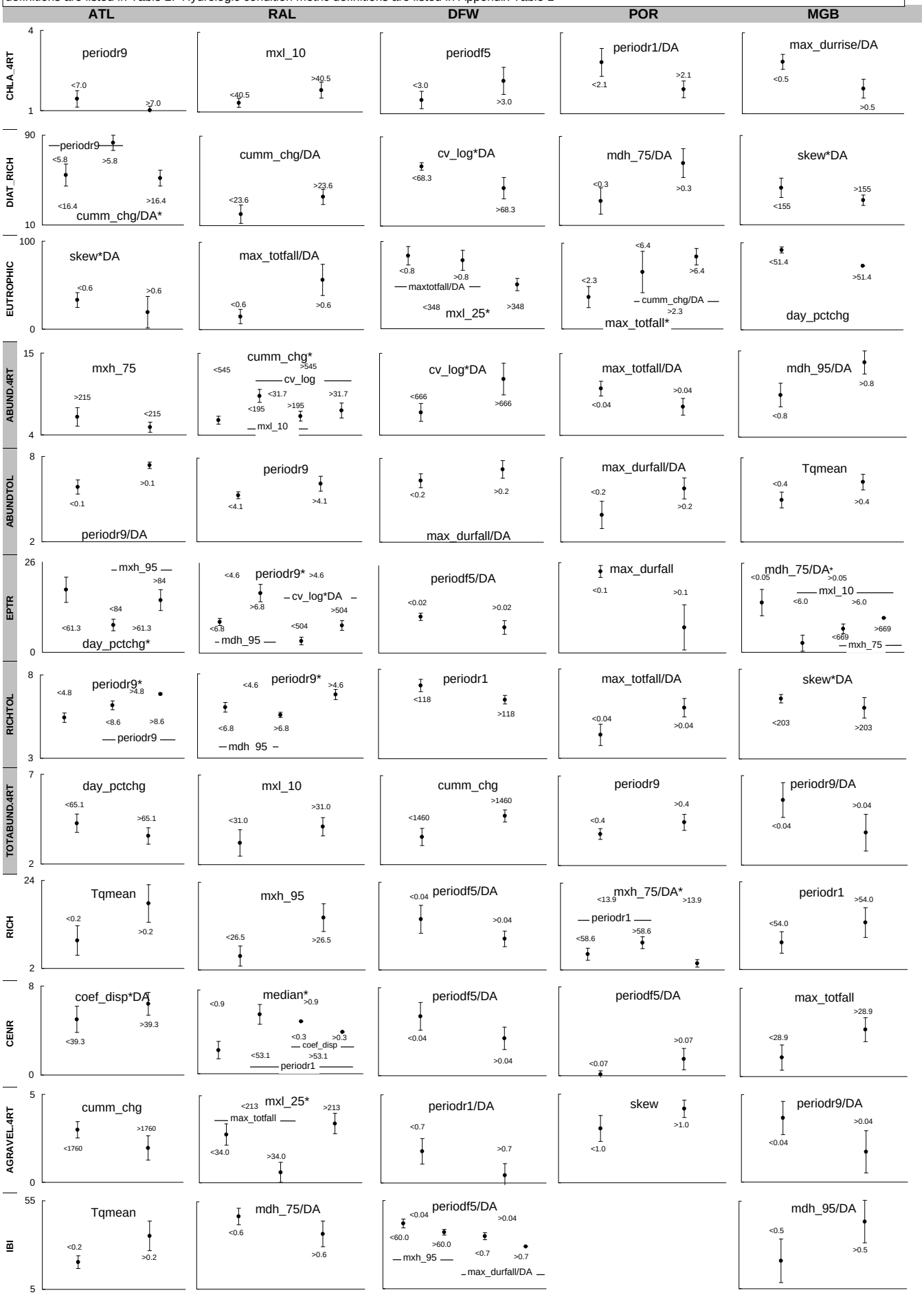
[1] Variable transformed by study area mean precipitation for use in combined metropolitan area data set.
[2] Variable transformed by study area maximum precipitation for use in combined metropolitan area data set.
[3] Variable transformed by study area mean soil moisture for use in combined metropolitan area data set.
[4] Variable transformed by study area maximum soil moisture for use in combined metropolitan area data set.
Unit abbreviations: m, meters; m^2, square meters, km^2, square kilometers

Appendix Table 2. Spearman correlation coefficients (ξ) between selected biological community metrics and hydrologic condition metrics (HCM) and impervious surface (NLCD_IS). Bolded coefficients indicate significance (p<0.05) within metropolitan areas. Metropolitan area totals are the number of significant correlations between individual biological community metrics and HCM but does not include significant correlations with impervious surface (NLCD_IS). Flow categories are from Olden & Poff (2003)

Hydrologic Condition Metric		Algae			Invertebrates				Fish				
Flow Category	HCM	CHLa	Diat_rich	EU-TROPHIC	ABUND.4rt	AbundTOL	EPTR	RichTOL	TotAbund	Rich	CenR	aGravel	IBI Score[1]
ATL (n=26)													
Magnitude average flow	coef_disp*DA	0.21	0.14	0.31	-0.02	-0.23	0.01	-0.05	-0.05	0.34	**0.42**	0.05	0.34
	skew	0.00	-0.10	**-0.42**	-0.12	-0.07	0.34	**-0.41**	0.09	-0.13	-0.25	0.10	0.04
	skew*DA	0.18	-0.04	**-0.43**	-0.12	-0.36	**0.55**	**-0.54**	0.12	0.03	-0.10	0.13	0.16
	Tqmean	0.15	0.29	0.06	-0.27	**-0.55**	**0.59**	**-0.59**	**0.51**	**0.44**	**0.44**	**0.49**	**0.53**
Magnitude high flow	pct_99a	-0.08	0.07	0.23	0.31	0.29	**-0.51**	**0.54**	**-0.47**	-0.23	-0.15	**-0.55**	**-0.44**
Frequency small flow pulses	periodr1	-0.16	0.09	-0.03	-0.04	0.03	0.19	-0.16	0.23	-0.13	**-0.40**	0.16	0.00
Frequency high flow events	periodf5	-0.34	0.15	0.32	0.32	**0.52**	**-0.64**	**0.57**	-0.03	0.05	-0.03	-0.16	-0.25
	periodf5/DA	-0.30	-0.05	0.09	0.23	**0.47**	**-0.39**	0.34	-0.02	-0.17	-0.28	-0.03	-0.25
	periodr9	**-0.44**	0.11	0.17	0.10	**0.48**	**-0.77**	**0.80**	-0.26	-0.27	-0.18	**-0.41**	**-0.53**
	periodr9/DA	-0.36	-0.16	0.05	0.05	**0.52**	**-0.62**	**0.57**	-0.26	-0.39	-0.30	-0.23	**-0.45**
Duration low flow	mdl_10	-0.29	-0.06	-0.01	**0.39**	0.15	-0.36	**0.39**	**-0.44**	-0.22	-0.07	-0.30	**-0.40**
	mxl_10	0.08	-0.14	-0.07	0.04	**-0.41**	0.03	0.04	-0.24	-0.29	0.08	-0.17	**-0.26**
	mxl_25	0.08	-0.09	-0.08	-0.11	**-0.42**	0.13	-0.13	-0.13	-0.10	0.09	0.03	-0.06
Duration high flow	max_durrise	-0.11	0.37	0.38	0.37	0.11	-0.31	**0.39**	-0.26	-0.01	0.08	-0.32	-0.22
	mdh_75	-0.15	0.17	0.27	0.36	0.14	**-0.51**	**0.53**	-0.21	0.03	0.09	-0.24	-0.23
	mdh_75/DA	-0.27	0.08	0.22	0.34	0.27	**-0.59**	**0.61**	-0.22	0.02	0.03	-0.21	-0.26
	mdh_95	**0.49**	0.09	0.29	0.18	-0.08	0.13	-0.06	0.19	0.37	0.28	0.24	**0.45**
	mdh_95/DA	**0.42**	0.05	0.27	0.26	-0.03	0.19	-0.17	0.21	0.28	0.24	0.38	**0.52**
Rate of change	cumm_chg	-0.36	-0.06	-0.04	0.06	**0.47**	**-0.68**	**0.68**	**-0.47**	-0.24	-0.23	**-0.52**	**-0.55**
	cumm_chg/DA	-0.33	-0.31	-0.28	0.01	**0.43**	**-0.46**	**0.40**	**-0.41**	-0.37	**-0.40**	-0.32	**-0.44**
	cumm_median	-0.26	-0.30	-0.18	-0.05	0.38	**-0.43**	0.37	-0.38	-0.32	-0.27	-0.23	-0.31
	day_pctchg	-0.26	-0.12	0.10	0.12	**0.48**	**-0.67**	**0.66**	**-0.45**	-0.22	-0.10	-0.34	-0.35
	NLCD_IS	-0.39	0.15	0.27	0.23	**0.64**	**-0.81**	**0.81**	-0.26	-0.17	-0.08	-0.39	**-0.47**
ATL metro area HCM totals (66)		3	0	2	1	10	13	14	6	1	4	3	9
RAL (n=25)													
Magnitude average flow	coef_disp*DA	-0.08	-0.05	**-0.51**	0.24	-0.20	**0.43**	-0.35	-0.08	0.21	0.14	-0.18	0.16
	cv_log	-0.26	0.04	0.12	0.01	0.07	-0.18	0.01	0.19	-0.23	**-0.59**	0.09	-0.16
	median	0.29	-0.12	-0.26	-0.09	-0.19	0.35	-0.12	0.02	0.34	**0.48**	0.03	0.26
	skew	-0.23	0.27	**0.52**	0.02	0.15	**-0.42**	0.20	-0.09	-0.21	-0.13	-0.01	-0.18
Magnitude low flow	pct_5n	0.07	0.06	0.17	0.11	0.11	-0.17	0.21	0.00	0.08	**0.61**	-0.06	0.00
Magnitude high flow	max_totfall/DA	-0.01	0.20	**0.61**	0.14	0.33	**-0.56**	**0.44**	0.01	-0.37	-0.26	-0.01	-0.37
Frequency small flow pulses	periodr1/DA	-0.16	0.16	0.35	0.02	0.29	**-0.50**	0.35	0.03	-0.31	**-0.49**	0.10	-0.24
Frequency high flow events	periodf5	-0.03	0.15	0.25	0.11	**0.69**	**-0.69**	**0.70**	0.14	-0.19	-0.01	-0.13	-0.25
	periodf5/DA	-0.11	0.15	**0.44**	0.04	**0.55**	**-0.75**	**0.63**	0.01	-0.38	-0.31	-0.04	-0.34
	periodr9	-0.01	0.14	0.30	0.18	**0.72**	**-0.70**	**0.73**	0.12	-0.25	0.03	-0.15	-0.29
	periodr9/DA	-0.07	0.13	**0.47**	0.03	**0.52**	**-0.71**	**0.61**	0.03	**-0.41**	-0.31	-0.02	-0.34
Duration low flow	mdl_10	-0.05	0.03	-0.27	-0.16	-0.20	**0.41**	-0.16	0.04	0.06	0.09	-0.01	0.16
	mxl_10	**0.43**	-0.32	-0.10	-0.30	-0.04	0.18	-0.01	0.12	0.34	0.18	0.21	0.26
	mxl_25	0.31	-0.23	-0.33	-0.33	-0.37	**0.49**	-0.26	0.17	0.38	-0.05	0.33	0.31
Duration high flow	max_durfall/DA	-0.04	0.12	**0.45**	-0.02	0.27	**-0.46**	0.40	-0.02	**-0.48**	-0.35	0.00	-0.38
	max_durrise	0.30	-0.30	**-0.47**	0.12	-0.39	**0.58**	-0.33	0.09	0.30	**0.43**	0.00	0.29
	max_durrise/DA	0.04	-0.04	0.34	0.01	0.06	-0.29	0.20	-0.04	-0.39	-0.29	0.06	-0.27
	mdh_75	0.01	-0.06	-0.26	-0.09	-0.36	**0.52**	**-0.43**	-0.14	0.01	0.17	-0.29	-0.11
	mdh_75/DA	-0.15	0.04	0.32	-0.18	-0.02	-0.01	-0.02	-0.29	**-0.52**	-0.22	-0.35	**-0.59**
	mdh_95	0.37	-0.31	**-0.43**	-0.09	**-0.55**	**0.79**	**-0.61**	0.09	0.29	0.21	0.12	0.30
	mxh_95	-0.01	-0.14	-0.30	-0.30	**-0.70**	**0.76**	**-0.70**	-0.08	0.26	0.10	0.16	0.32
Rate of change	cumm_chg	0.07	0.12	0.24	0.20	**0.52**	**-0.51**	**0.58**	0.29	-0.05	0.03	-0.02	-0.12
	cumm_chg/DA	0.00	0.18	**0.49**	0.14	**0.54**	**-0.74**	**0.67**	0.06	-0.35	-0.19	-0.01	-0.34
	cumm_median	-0.13	0.05	0.31	0.23	**0.47**	**-0.55**	**0.43**	0.02	**-0.43**	-0.34	-0.11	-0.32
	day_pctchg	0.00	0.05	0.25	0.18	**0.47**	**-0.46**	**0.42**	0.16	-0.31	-0.36	-0.05	-0.28
	NLCD_IS	-0.24	0.37	**0.58**	0.18	**0.69**	**-0.89**	**0.82**	0.03	-0.40	-0.09	-0.14	-0.45
RAL metro area HCM totals (61)		1	0	9	0	10	19	12	0	4	5	0	**1**
POR (n=23)													
Magnitude average flow	coef_disp	0.12	0.05	0.10	-0.18	0.12	-0.42	**0.46**	-0.05	0.05	0.27	-0.16	-
	cv_log	0.23	0.14	0.35	-0.35	0.36	**-0.62**	**0.65**	0.05	-0.01	0.40	-0.02	-
	cv_log*DA	0.17	0.03	0.10	-0.16	0.13	-0.41	**0.46**	0.05	0.13	0.29	-0.07	-
	median	-0.16	-0.12	-0.31	**0.45**	-0.39	**0.50**	**-0.49**	-0.11	0.15	-0.27	-0.24	-
	median_DA	-0.15	-0.01	-0.15	0.35	-0.23	**0.44**	**-0.43**	-0.09	0.04	-0.19	-0.26	-
	skew	0.00	0.20	**0.54**	**-0.43**	**0.46**	**-0.48**	**0.49**	0.17	**-0.43**	**0.44**	0.28	-
	Tqmean	-0.12	-0.10	**-0.53**	0.12	**-0.45**	**0.52**	**-0.53**	-0.17	0.02	**-0.63**	-0.11	-
Magnitude low flow	pct_5n	-0.28	-0.10	-0.24	0.31	-0.24	**0.44**	**-0.48**	-0.14	-0.09	-0.22	0.03	-
Magnitude high flow	max_totfall	0.28	0.38	**0.59**	**-0.49**	**0.46**	**-0.52**	**0.54**	0.17	-0.02	**0.51**	-0.06	-
	max_totfall/DA	0.19	**0.47**	**0.62**	**-0.57**	**0.61**	**-0.62**	**0.61**	0.17	-0.19	**0.56**	-0.02	-
Frequency small flow pulses	periodr1	-0.12	-0.21	**-0.54**	0.33	-0.27	**0.46**	**-0.55**	-0.04	0.41	-0.28	-0.17	-
Frequency high flow events	periodf5	0.19	0.41	**0.49**	**-0.49**	**0.58**	**-0.62**	**0.68**	0.27	-0.08	**0.60**	0.01	-
	periodf5/DA	0.15	**0.48**	**0.54**	**-0.58**	**0.67**	**-0.64**	**0.69**	0.27	-0.21	**0.62**	0.07	-
	periodr9	0.27	0.23	0.41	-0.32	**0.48**	**-0.52**	**0.56**	0.31	0.06	**0.56**	0.04	-
	periodr9/DA	0.25	0.34	**0.57**	**-0.46**	**0.61**	**-0.60**	**0.65**	0.34	-0.07	**0.62**	0.09	-
Duration low flow	mxl_10	0.14	0.34	0.24	**-0.45**	0.19	-0.13	0.14	0.05	-0.01	0.19	-0.03	-
Duration high flow	max_durfall/DA	0.09	0.41	**0.60**	**-0.45**	**0.63**	**-0.62**	**0.62**	0.18	-0.28	**0.57**	0.04	-

Hydrologic Condition Metric		Algae			Invertebrates				Fish				
Flow Category	HCM	CHLa	Diat_rich	EU-TROPHIC	ABUND.4rt	AbundTOL	EPTR	RichTOL	TotAbund	Rich	CenR	aGravel	IBI Score[1]
	mxh_75	-0.32	0.21	**-0.45**	0.06	0.15	0.01	-0.08	-0.16	-0.17	-0.02	-0.10	-
	mxh_95	-0.21	0.18	**-0.48**	0.01	0.07	0.06	-0.18	-0.26	-0.05	-0.11	-0.21	-
Rate of change	cumm_median	0.22	0.22	0.36	**-0.53**	**0.62**	**-0.73**	**0.70**	0.29	-0.10	**0.52**	0.24	-
	day_pctchg	0.38	0.23	**0.51**	**-0.43**	**0.48**	**-0.70**	**0.73**	0.19	0.05	**0.58**	-0.01	-
	NLCD_IS	0.20	**0.42**	**0.56**	**-0.58**	**0.76**	**-0.78**	**0.75**	0.34	-0.12	**0.50**	0.30	-
POR metro area HCM totals (82)		0	2	12	11	11	16	18	0	1	11	0	
MGB (n=22)													
Magnitude average flow	cv_log	-0.17	-0.17	-0.29	-0.13	0.04	**-0.49**	0.03	0.03	0.00	0.06	-0.18	0.08
	median	0.03	0.17	0.29	-0.07	-0.06	**0.48**	-0.06	0.06	0.20	0.13	0.20	-0.03
	skew*DA	0.11	**-0.50**	0.32	-0.28	-0.34	0.30	-0.29	0.01	0.27	0.11	0.30	-0.02
	Tqmean	0.25	**0.48**	-0.16	0.14	0.18	0.02	-0.13	0.16	-0.10	-0.40	-0.11	-0.18
Magnitude high flow	pct_99a	-0.27	-0.15	0.10	-0.33	-0.15	0.34	-0.05	0.11	**0.50**	**0.52**	0.26	0.30
Frequency small flow pulses	periodr1	-0.24	0.26	-0.33	-0.11	-0.33	0.15	-0.08	0.39	**0.64**	0.38	0.35	**0.61**
	periodr1/DA	-0.27	0.23	**-0.47**	-0.25	-0.21	-0.29	0.02	-0.09	0.09	0.31	-0.09	0.27
Frequency high flow events	periodf5	-0.11	0.13	-0.14	-0.18	0.17	-0.29	0.23	-0.42	-0.20	**0.47**	-0.24	-0.29
	periodf5/DA	-0.20	0.21	-0.34	-0.18	0.08	**-0.46**	0.28	**-0.53**	-0.35	0.32	-0.34	-0.29
	periodr9	-0.08	0.01	-0.01	-0.30	0.25	-0.36	0.18	**-0.62**	**-0.48**	0.33	**-0.48**	**-0.49**
	periodr9/DA	-0.18	0.03	-0.20	-0.33	0.11	**-0.53**	0.17	**-0.67**	**-0.54**	0.25	**-0.57**	-0.35
Duration low flow	mdl_10	-0.17	-0.20	**0.57**	-0.18	0.01	0.19	0.02	0.17	0.17	0.34	-0.15	0.11
	mxl_10	0.10	-0.28	**0.55**	-0.06	-0.09	**0.46**	-0.23	0.21	0.26	-0.12	0.02	0.30
Duration high flow	mdh_75/DA	-0.06	0.09	-0.21	0.24	0.23	-0.40	**0.49**	-0.06	-0.23	0.15	0.00	-0.26
	mdh_95	0.01	-0.08	0.13	0.31	-0.15	0.28	-0.08	**0.54**	**0.52**	-0.20	0.39	**0.58**
	mdh_95/DA	-0.12	0.06	-0.12	0.33	-0.25	0.05	-0.01	**0.47**	0.40	-0.19	0.20	**0.67**
	mxh_95	-0.33	-0.27	0.39	**-0.48**	-0.13	0.32	-0.34	0.25	**0.45**	0.12	0.12	0.09
	mxh_95/DA	**-0.48**	-0.01	-0.06	-0.40	-0.05	-0.24	0.09	-0.10	0.03	0.02	-0.32	0.02
Rate of change	cumm_chg	0.00	0.04	-0.29	-0.22	-0.04	-0.09	0.00	-0.26	0.10	**0.46**	0.00	-0.02
	cumm_median	0.02	-0.14	**-0.49**	-0.03	-0.01	**-0.54**	0.04	-0.30	-0.20	0.20	-0.22	-0.01
	day_pctchg	0.02	-0.19	-0.42	-0.11	0.05	**-0.57**	0.08	-0.31	-0.24	0.19	-0.26	-0.03
	NLCD_IS	-0.18	-0.02	-0.08	-0.33	0.15	**-0.52**	0.16	**-0.73**	**-0.58**	0.37	**-0.61**	**-0.51**
MGB metro area HCM totals (36)		1	2	4	1	0	7	1	5	6	3	2	4
DFW (n=25)													
Magnitude average flow	cv_log	0.05	-0.03	0.08	0.17	0.02	-0.05	0.24	-0.18	-0.16	-0.30	0.21	**-0.42**
	skew	0.40	-0.38	-0.19	0.12	-0.03	**0.41**	-0.28	0.26	-0.14	-0.10	0.03	0.14
	skew*DA	**0.44**	-0.23	-0.04	0.17	-0.13	0.38	-0.22	0.30	0.18	0.10	0.20	0.37
Magnitude high flow	max_totfall	**0.54**	-0.10	-0.04	0.16	0.13	0.10	-0.21	0.20	-0.08	-0.05	0.03	0.01
	max_totfall/DA	0.33	-0.11	-0.15	-0.19	0.31	-0.17	-0.04	0.13	**-0.43**	-0.23	-0.34	-0.19
	pct_99a	**0.48**	0.31	-0.12	0.06	0.26	-0.21	-0.10	0.22	0.06	0.23	0.05	0.20
Frequency high flow events	periodf5/DA	0.21	0.15	-0.10	0.11	0.23	**-0.46**	0.08	-0.07	**-0.59**	**-0.50**	-0.24	**-0.60**
	periodr9/DA	0.13	0.11	-0.11	-0.02	0.05	**-0.45**	-0.03	0.00	**-0.48**	-0.36	-0.23	**-0.49**
Duration low flow	mxl_10	0.21	-0.17	0.12	-0.01	-0.31	0.06	-0.12	**0.48**	0.07	-0.17	-0.10	0.12
	mxl_25	**0.44**	0.03	-0.39	-0.05	0.18	0.08	-0.07	0.28	0.11	-0.12	-0.03	0.34
Duration high flow	max_durfall/DA	-0.13	-0.03	-0.01	-0.39	**0.46**	-0.18	0.36	-0.26	**-0.48**	-0.33	-0.28	-0.25
	max_durrise/DA	-0.09	-0.09	-0.18	-0.16	**0.45**	-0.10	0.25	-0.30	-0.27	-0.15	-0.09	-0.26
	mxh_75/DA	-0.27	0.15	-0.02	-0.21	0.24	-0.28	0.01	-0.17	-0.15	-0.10	-0.26	**-0.42**
	mxh_95/DA	**-0.56**	-0.24	0.08	-0.33	0.17	0.09	0.26	**-0.41**	-0.06	-0.01	-0.21	-0.27
Rate of change	cumm_chg	**0.54**	0.20	-0.11	0.17	-0.01	-0.05	-0.21	0.27	-0.08	0.02	0.08	0.08
	cumm_chg/DA	0.34	0.13	-0.25	-0.16	0.24	-0.32	-0.04	0.18	**-0.41**	-0.13	-0.28	-0.20
	NLCD_IS	**0.58**	0.19	-0.13	0.29	-0.14	-0.35	-0.28	0.27	-0.07	0.06	0.11	-0.14
DFW metro area HCM totals (23)		6	0	0	0	2	3	0	2	5	1	0	4

[1] No fish IBI scores were used in POR metropolitan area.

[2] Totals do not include significant correlations between HCM and NLCD_IS.

Appendix Table 3. Summary of results from the metropolitan scale regression tree model analysis. Hydrologic condition metric decision variables and values and biologic endpoints (mean ± 1 SD) are denoted. Primary (initial) split node has asterisk. For example, in the RAL RichTol model if periodr9 > 4.6 RichTol mean = 6.8 ± 0.3; for periodr9 < 4.6 and mdh_95 < 6.8 RichTol mean = 6.0 ± 0.3; for periodr9 < 4.6 and mdh_95 > 6.8 RichTol mean = 5.1 ± 0.2. Biologic metric definitions are listed in Table 2. Hydrologic condition metric definitions are listed in Appendix Table 1