Universities Council on Water Resources
Journal of Contemporary Water Research & Education
Issue 177, Pages 46-62, April 2023

Case Study Article

Informing Volunteer Water Quality Monitoring Program Design and Watershed Planning: Case Study of StreamSmart Data Analysis in the Upper White River Basin, Arkansas

*Erin Grantz[1], Brian E. Haggard[1,2], and BENG 4973/5973[2]

[1]*Arkansas Water Resources Center, University of Arkansas Division of Agriculture, Fayetteville, AR*
[2]*Biological and Agricultural Engineering Department, University of Arkansas, Fayetteville, AR*
**Corresponding Author*

Abstract: The watershed group H_2Ozarks founded the StreamSmart Citizen Science Program to establish baseline and long-term water quality data for the Upper White River Basin, Arkansas. StreamSmart volunteers collect water samples and conduct habitat and macroinvertebrate community assessments at >20 sites across a land use-land cover (LULC) gradient. Since 2020, H_2Ozarks has adaptively assessed the program to ensure that the investment in water quality data meets core goals, with particular interest in planning tools and aligning expectations of volunteer effort with the level of training and support. Study objectives were to use StreamSmart data to 1) facilitate understanding of water quality response to stressors in the basin using a range of methods (Spearman rank correlation, non-parametric changepoint analysis, and categorical and regression tree analysis) and 2) explore implications for program design and watershed planning. Water chemistry-LULC relationships were in-line with prior regional studies, as well as global patterns. Detected thresholds and hierarchy provide potential targets for managing LULC change to protect water quality, but further analysis is warranted to refine these relationships. Macroinvertebrate stressor-response was most detectable for sensitive and less sensitive taxa and for habitat index components, suggesting potential to streamline these programmatic elements. Study findings for StreamSmart should also be informative for other small-scale volunteer monitoring programs with limited resources, but which actively evaluate the types of data and program activities that yield a maximum scientific return on investment.

Keywords: *Volunteer monitoring, water quality, watershed management*

Human activities in watersheds influence the quality of adjacent and downstream water resources for beneficial uses like drinking water, aquatic life habitat, and recreation. At both global and local scales, greater extent of anthropogenic land use-land cover (LULC) types like urban and agriculture in watersheds is correlated with greater levels of nutrients, sediments, and salts in connected water bodies (Giovanetti et al. 2013; Lintern et al. 2017). Point sources, such as industrial or municipal wastewater discharges, also play a role, though these inputs are more easily quantified, regulated, and mitigated compared to non-point sources (Haggard 2010; Scott et al. 2011).

The water quality effects of human activities in watersheds extend beyond water chemistry, with in-stream and riparian habitat quality often becoming less stable and complex as watershed disturbance increases, supporting fewer sensitive species (White and Walsh 2020). The combined effects of water chemistry changes and habitat quality loss compound in the biological community response, with shifts to increased densities of individuals from pollution-tolerant taxa and overall reduced taxonomic richness (Xu et al. 2013).

It is increasingly recognized that a broad array of stakeholders must be mobilized to effectively address the effects of human activities on water

Research Implications

- StreamSmart water chemistry data responded to land use-land cover (LULC) gradients, most notably human development index thresholds and hierarchy that may provide useful targets for watershed planning.
- Highly predictive water chemistry-LULC relationships suggest that StreamSmart data can be combined with other datasets in knowledge "co-creation" around watershed management and planning.
- Macroinvertebrate and habitat stressor-response relationships were most detectable when considering sensitive groups and habitat components, like epifaunal substrate/cover, riffle/bend frequency, and channel flow status.
- Relationships between sensitive groups and habitat components may reflect volunteer biases, but also present an opportunity for StreamSmart to collect the same information with less volunteer time and effort.

quality (USEPA 2005). Watershed management planning is a stakeholder-driven process that takes a holistic approach to water quality protection and restoration in a specific watershed (USEPA 2008). Water quality monitoring is a core component of a watershed management plan, both to establish baseline conditions and to collect real-time information as watersheds evolve. Citizen science programs have an established history in water resources research and management, including water quality monitoring (Buytaert et al. 2014). These programs can be an entry point for stakeholders to community involvement and education around watershed management (Savan et al. 2003; Storey et al. 2016), and have been shown to produce water quality data of comparable quality to professionally collected datasets (Hoyer and Canfield 2021).

In Northwest Arkansas, the watershed group H_2Ozarks (formerly Ozark Water Watch) seeks to increase stakeholder awareness of water quality and watershed function by engaging the public in the StreamSmart Citizen Science Program. StreamSmart leverages volunteer monitoring to establish a baseline water quality database for the Upper White River Basin. The volunteers collect water samples and assess habitat quality and the aquatic macroinvertebrate community at more than 20 sites. The Upper White River Basin is rapidly urbanizing (NWARPC 2016) and is also the source water area for Beaver Lake, which provides drinking water for ~1 in 6 Arkansans. StreamSmart complements monitoring by other entities by providing more granular coverage of the basin.

Since 2020, StreamSmart has been adaptively assessing the program to ensure that the investment in water quality data meets core goals. The primary goal is to inform stakeholders about current water quality and any potential changes in the basin. But, H_2Ozarks and its partners, the Beaver Water District's source water protection program and the Beaver Watershed Alliance, also want to use StreamSmart data to inform nutrient reduction strategies, such as siting best management practices. Further, program changes have focused on making sure the expected volunteer time and effort investment matches the level of support that the group can provide. Training is an essential component of citizen science program success (Nerbonne and Vondracek 2003; Lewandoski and Specht 2015; San Llorente Capdevila et al. 2020), and the current StreamSmart training and staffing levels may not be sufficient for reliably generating complex data such as macroinvertebrate or habitat assessments (Fore et al. 2001). Volunteer interests and desired time investment are also considerations. Future changes may include scaling back macroinvertebrate work, and habitat assessments have already been discontinued.

The Arkansas Water Resources Center has provided lab services funded through Section 104(b) of the Water Resources Research Act of 1984 and technical support of volunteer training since StreamSmart began. In this study, we conducted a comprehensive stressor-response analysis of the StreamSmart volunteer water quality monitoring database. The study objectives were to 1) facilitate understanding water quality response to stressors in the basin using a range of stressor-response methods and 2) explore potential implications of study findings for the StreamSmart program design and watershed planning. Study findings for StreamSmart should also be informative for other volunteer monitoring programs of similar size, scale, and resource availability in evaluating, or re-evaluating, the types of data to collect for maximum return on investment.

Methods

Study Location and Site Characteristics

StreamSmart was founded in 2012 to monitor water quality in the Upper White River Basin (Figure 1). The basin is primarily forested (>60%), and pasture agriculture is the predominant anthropogenic LULC. However, rapid urbanization is also occurring, typically on prior pasture lands. StreamSmart volunteers have collected water samples, conducted habitat assessments, and collected information on aquatic macroinvertebrates at 23 sites since the program began (Table 1), with 14 sites currently active. StreamSmart selects and maintains core monitoring locations to encompass gradients of LULC types. Volunteers are required to attend a half-day comprehensive training covering best

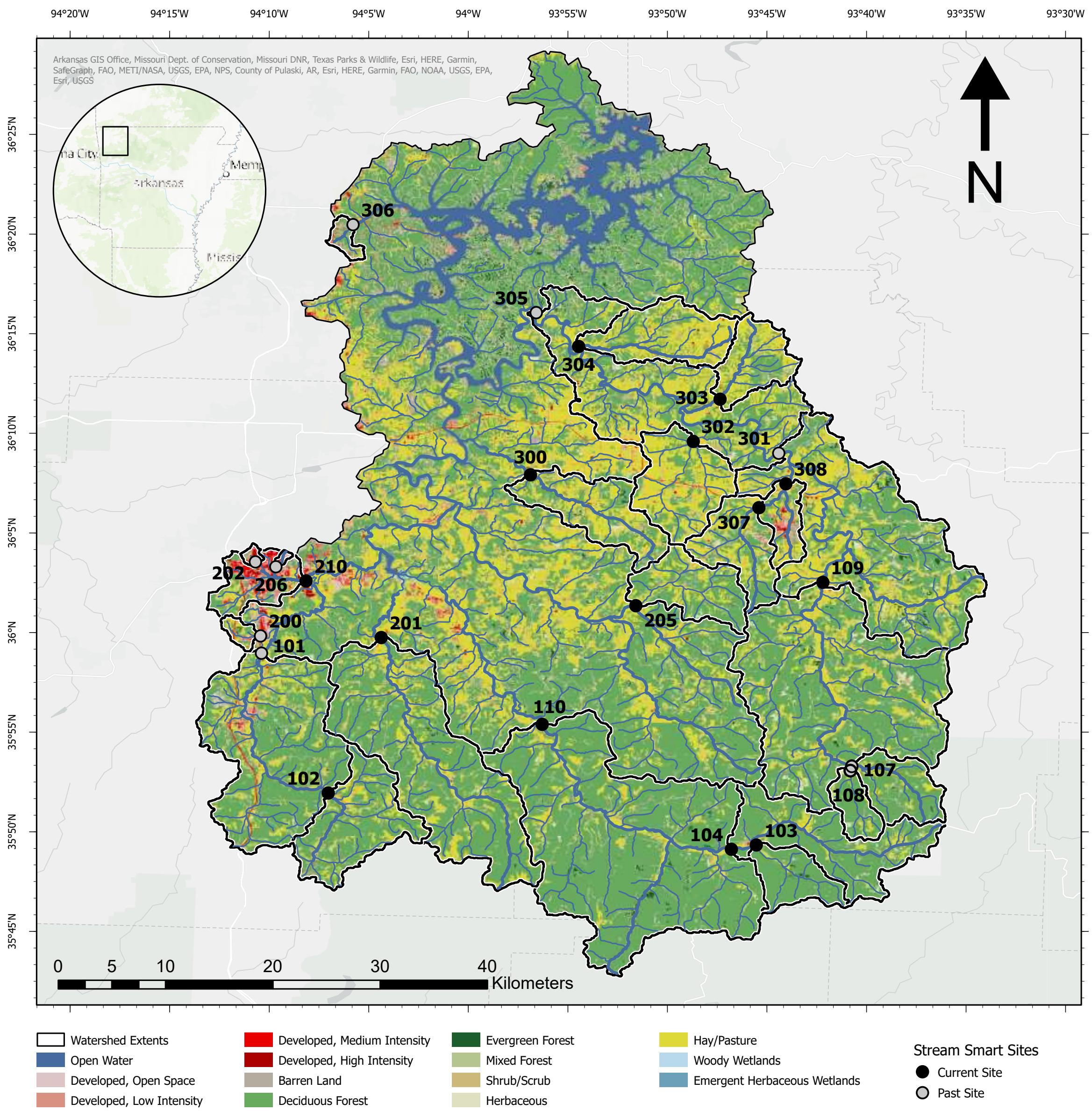

Figure 1. Map of the study area, the Upper White River Basin located in Northwest Arkansas, showing distribution of land use-land cover (LULC) characteristics in the watershed (Dewitz and USGS 2021), as well as the locations of StreamSmart volunteer monitoring sites (2012 – present).

practices for collecting water samples, kicking for macroinvertebrates, and habitat assessment terminology. The training includes a brief on-site demonstration by water quality professionals.

This analysis used data from 21 sites and focused on the period 2012 – 2020. Two sites were not included; site 110 was not established until 2020, and site 308 has a point-source discharger in the watershed (City of Huntsville municipal wastewater treatment plant). Signals of point-source pollution were expected to confound analysis, as it is recognized that non-point sources are dominant in the watershed, including channel erosion and runoff from farms, unpaved roads, and urban or urbanizing areas (Perez et al. 2015).

StreamSmart site sub-watershed areas and LULC data (MRLC 2018) were obtained from https://modelmywatershed.org/. Summary LULC categories were calculated as the percentage of each site's sub-watershed area (Table 1). Agricultural land (%) was the sum of pasture/hay, grassland/herbaceous, and cultivated crops; forest land (%) was the sum of deciduous, evergreen, and mixed forest categories, as well as shrub/scrub; and urban land (%) was the sum of all developed and barren land categories. A human development index (%) for each site was calculated as the sum of agricultural and urban land. Locations of poultry houses were obtained from the Arkansas Highway and Transportation Department cultural features GIS database (Arkansas GIS Office 2014), and poultry house density (houses/km^2) was calculated as number of houses divided by sub-watershed area, for each site.

Table 1. StreamSmart monitoring site information and watershed land use-land cover (LULC) characteristics, including poultry house density, agricultural, forest, and urban land, as well as the human development index (HDI).

Site #	Hydrologic Unit Code 10 Name	Latitude	Longitude	Site sub-watershed area (km^2)	Poultry house density (houses/km^2)	Agriculture (%)	Forest (%)	Urban (%)	HDI (%)
101	West Fork	35.982714	-94.173129	215	0.49	25.3	66.9	7.0	32.3
102	West Fork	35.865723	-94.117257	65	0.92	26.3	68.1	5.1	31.5
103	Headwaters	35.822256	-93.758937	29	0.17	4.0	92.6	1.7	5.7
104	Headwaters	35.818676	-93.779774	106	0.26	10.9	84.7	3.7	14.6
107	War Eagle	35.888319	-93.679017	50	0.04	15.5	81.1	2.7	18.2
108	War Eagle	35.887989	-93.678974	17	0	13.5	84.3	1.5	15.1
109	War Eagle	36.041958	-93.703225	273	0.29	22.2	73.6	3.3	25.5
200	West Fork	35.997178	-94.173949	5	1.20	40.0	36.8	29.6	69.6
201	Middle Fork	35.995825	-94.072894	174	0.67	27.3	68.8	2.8	30.1
202	West Fork	36.059103	-94.178209	1.63	0	1.2	19.0	79.8	81.0
205	Richland	36.022453	-93.859784	43	0.61	33.1	62.7	3.1	36.2
206	West Fork	36.055019	-94.161107	1.13	0	0	0.9	99.1	99.1
210	West Fork	36.043179	-94.135852	31	0	15.5	33.4	49.2	64.7
300	Beaver Reservoir	36.131947	-93.947956	52	1.00	51.0	42.6	5.5	56.5
301	War Eagle	36.149997	-93.740137	525	0.46	30.7	63.6	4.8	35.6
302	War Eagle	36.159851	-93.81169	56	1.18	66.8	27.9	5.3	72.1
303	War Eagle	36.195153	-93.789276	32	1.28	59.8	34.3	5.3	65.0
304	War Eagle	36.239342	-93.907653	50	1.00	61.7	31.8	5.8	67.4
305	War Eagle	36.267597	-93.94313	808	0.68	39.5	54.7	4.9	44.4
306	Beaver Reservoir	36.341208	-94.096513	7	0	15.1	55.4	31.0	46.1
307	War Eagle	36.104418	-93.75675	42	0.74	43.2	48.6	7.2	50.4
308	War Eagle	36.124453	-93.734211	61	0.62	43.4	44.6	11.3	54.8

Sample Collection and Analysis

Water Chemistry. Volunteer teams collected grab water samples quarterly (February, May, August, and November) at each site. Each sampling event was carried out at all sites within a two-week timeframe during base flow conditions. Samples were collected from the thalweg while facing upstream from the access point and taking care not to capture any disturbed sediments. Clean and acid-washed sample bottles were provided by the Arkansas Water Resources Center Water Quality Lab and were triple rinsed at the stream by volunteers prior to sample collection. Samples were stored on ice and in the dark until being returned to the lab within 36 hours to allow processing within 48 hours. Chain of custody was documented at each step.

Water samples were analyzed at the lab using standard procedures for the following water chemistry variables: alkalinity (mg/L $CaCO_3$), conductivity (uS/cm), pH, total dissolved solids (TDS, mg/L), nitrate+nitrite-nitrogen (NO_x-N, mg/L), total nitrogen (TN, mg/L), total phosphorus (TP, mg/L), and total suspended solids (TSS, mg/L). Analytical methods, detection and reporting limits, preservation, holding times, and quality assurance details are available at https://awrc.uada.edu/water-quality-lab/certification-and-quality-assurance/. The lab is certified under the State Environmental Laboratory Certification Program by the Arkansas Department of Energy and Environment – Environmental Quality Division.

Habitat Quality Assessment. At each site visit, StreamSmart volunteers completed the USEPA Rapid Bioassessment Protocols for Habitat Assessment rubric (Barbour et al. 1999), which uses visual assessment of habitat quality for aquatic life use. The rubric includes ten components: 1) epifaunal substrate/available cover, 2) embeddedness, 3) velocity/depth regime, 4) sediment deposition, 5) channel flow status, 6) channel alteration, 7) frequency of riffles (or bends), 8) bank stability, 9) vegetative protection, and 10) riparian vegetative zone width. Descriptions were provided for 5-point intervals to score each component (0 – 20); component scores were summed into a habitat quality index score.

Aquatic Macroinvertebrate Community Index. During May and August site visits, volunteers collected aquatic macroinvertebrates for identification and community assessment. Stream riffle cross sections were sampled at an angle moving upstream at three locations, avoiding bridges and road crossings. Riffle locations were sampled by kicking into a D-frame net within a 1 m^2 area for one minute after first setting aside any large substrate in a collection tub. The net was rinsed with stream water into a container after each kick. Large substrate and net contents were examined for macroinvertebrates, which were removed for identification using StreamSmart's simplified flow chart.

The macroinvertebrate community index was designed by StreamSmart specifically for non-expert volunteers. Pre-defined taxonomic units, approximately at the Family level, were marked as present (1) or absent (0) in a rubric. Taxa were grouped into categories based on relative sensitivity to habitat and water quality degradation (i.e., sensitive, less sensitive, and tolerant). Sensitive taxa included caddisfly larvae, hellgrammites, mayfly nymphs, gilled snails, riffle beetle adult, stonefly nymphs, and water penny larvae. Less sensitive taxa were beetle larvae, clams, crane fly larvae, crayfish, damselfly nymphs, dragonfly nymphs, scuds, sowbugs, fishfly larvae, alderfly larvae, and watersnipe fly larvae. Tolerant taxa were aquatic worms, blackfly larvae, leeches, midge larvae, and pouch snails. The macroinvertebrate community index was the sum of the count of present taxa after weighting each sensitivity group using multipliers of 3, 2, and 1 for sensitive, less sensitive, and tolerant taxa, respectively. Site water quality was classified as excellent (≥22), good (17-22), fair (11-16), or poor (<11) based on the index score.

Data Analysis

Site medians were calculated for all water chemistry, habitat, and macroinvertebrate variables, including habitat components and macroinvertebrate sensitivity groups, for use in stressor-response analysis. Site median calculation and all subsequently described analyses were carried out using R 4.1.2 (R Core Team 2021). We explored stressor-response relationships using Spearman rank-order correlation and non-parametric changepoint analysis (nCPA) (King and Richardson 2003; Qian et al. 2003). Correlation is commonly used to describe monotonic water chemistry-LULC relationships, facilitating comparison with preceding studies in the basin (Giovanetti et al. 2013). However, these relationships can also be non-linear, such as thresholds stressor values associated with disproportionate water quality response. Non-parametric changepoint analysis divides data at a threshold value in the explanatory variable by minimizing deviance within groups. For a changepoint to be detected, groups on both sides of the threshold had to have at least three observations. For water chemistry, stressor-response analysis focused on LULC. For macroinvertebrates, potential stressors included water chemistry, habitat quality index and component scores, as well as LULC. Correlations and changepoints were considered significant if $p<0.10$.

Water quality can also relate to watershed stressors in a hierarchy, where a relationship may only be observed, or is much stronger, if other primary conditions are met. We explored potential hierarchy in water chemistry responses to LULC using categorical and regression tree analysis (CART; De'Ath and Fabricius 2000) with the rpart package in R (Therneau and Atkinson 2019). Data were insufficient to explore stressor hierarchy and structure for macroinvertebrates. Similar to nCPA, CART divides and groups data to minimize deviance. However, CART can consider multiple variables simultaneously and recursively partitions data into subsets based on identified thresholds. These data subsets may then be split again based on secondary or tertiary thresholds. Control parameters in CART were tuned to require groups to contain at least three observations. Splits in final models were required to reduce deviance by at least 5% (i.e., complexity parameter ≥0.05). We used urban, agriculture, and human development index (but not forest) as model inputs to simplify results interpretation.

Results

StreamSmart Site Medians

Site medians for alkalinity, conductivity, and TDS each spanned an order of magnitude (6 – 150 mg/L $CaCO_3$, 24 – 542 µS/cm, and 30 – 303 mg/L, respectively) (Table 2). Auto-correlation among these three variables was evident, with the least and greatest medians aligning across sites. Site median pH ranged from slightly less than neutral (6.6) to alkaline (8.0). For nutrients, TP varied within a narrow range (0.010 – 0.038 mg/L), except for site 308, where municipal wastewater treatment plant influence was evident (0.11 mg/L). TN medians, in contrast, varied over an order of magnitude (0.12 – 3.8 mg/L). All TSS medians were less than the lab's reporting limit of 10 mg/L. The minimum habitat quality index (99) and macroinvertebrate community index (7.5) medians were both observed at site 210, while site 301 had the greatest median habitat quality index (145), and site 205 had the greatest median macroinvertebrate community index (20).

Water Chemistry-LULC Relationships

All water chemistry variables, except TSS ($p > 0.10$), were correlated with the level of total anthropogenic watershed disturbance (Table 3; Figure 2A-F), increasing with increasing human development index ($p < 0.001$, rho = 0.61 – 0.95) and decreasing with increasing forest ($p < 0.001$, rho = -0.66 – -0.94). All water chemistry analytes were significantly (positively) correlated with urban land (p = < 0.001 – 0.080, rho = 0.39 – 0.87), but correlation analysis was not effective for describing water chemistry-LULC relationships with agricultural land, with the exception of TN ($p = 0.014$, rho = 0.53). Water quality medians were highly variable among sites with the least agriculture, which diluted otherwise linear signals above ~10% agriculture. As with agriculture, only TN was significantly correlated with poultry house density ($p = 0.097$, rho = 0.37). Pasture land and poultry houses are often spatially paired in the basin.

Changepoints in the human development index and forest land were found for all analytes (p < 0.001 – 0.034, R^2 = 0.37 – 0.82), except TSS ($p > 0.10$), suggesting that water chemistry values tended to be greater above a human development index threshold range = 27.8 – 45.3% and tended to be less above a forest land threshold range = 29.9 – 71.2%. Thresholds ranging from 4 – 5% urban land were also detected for all variables (p < 0.001 – 0.068, R^2 = 0.31 – 0.75), except TSS.

In contrast to correlation analysis, agricultural land thresholds were identified for both TN ($p < 0.001$, $R^2 = 0.56$) and TP ($p = 0.070$, $R^2 = 0.36$), estimated as 43.2% (CI = 16.0 – 49.9%) and 4.0% (CI = 1.2 – 45.3%), respectively. The nCPA models for alkalinity, conductivity, pH, and TDS tended to have greater explanatory power ($R^2 = 0.67 – 0.82$) relative to nutrients ($R^2 = 0.31 – 0.56$), as well as greater confidence in the threshold estimate (i.e., narrower CI).

Macroinvertebrate Community Relationships

Total macroinvertebrate community index scores were correlated with urban land (p = 0.047, rho = -0.56) and three components of the habitat quality index, but not the cumulative index (Table 4; Figure 3A-F). These components were channel flow status (p = 0.032, rho = 0.60), epifaunal substrate/available cover (p = 0.035, rho = 0.59), and frequency of riffles/bends (p = 0.045, rho = 0.56). Changepoints were also identified, suggesting greater macroinvertebrate community scores occurring above thresholds of 11 in both channel flow status sub-scores (CI = 10.0 – 13.3) and epifaunal substrate/available cover sub-scores (CI = 9.5 – 14.0). The sub-score of 11 ranks just above the mid-point in the possible range (i.e., 0 – 20), and is the lowest value considered to represent "sub-optimal" conditions.

These same relationships were also observed when sub-scores for the sensitive (epifaunal substrate/available cover and frequency of

Table 2. Sample counts for total sampling events (n_{events}) and macroinvertebrate collections (n_{MI}) at StreamSmart monitoring sites. Site medians for each water chemistry variable, as well as habitat quality index (HQI) assessment and macroinvertebrate community index scores.

Site	n_{events}	Alk (mg/L $CaCO_3$)	Cond (µS/cm)	pH	TDS (mg/L)	TN (mg/L)	TP (mg/L)	TSS (mg/L)	HQI	n_{MI}	Macro Index
101	22	64	193	7.8	99	0.34	0.012	2.1	125	0	-
102	24	32	103	7.7	54	0.34	0.014	2.2	124	6	15
103	20	6	24	6.6	30	0.12	0.014	1.2	133	8	19
104	19	10	32	6.8	31	0.28	0.014	1.8	120	8	12
107	8	20	55	7.0	39	0.13	0.013	0.6	135	0	-
108	7	12	40	6.8	32	0.12	0.012	0.7	135	0	-
109	7	30	81	7.0	60	0.53	0.015	1.7	141	2	14
200	14	138	527	7.8	312	0.29	0.012	3.0	108	0	-
201	23	43	118	7.5	65	0.47	0.010	1.4	129	6	15
202	10	143	542	7.7	303	1.2	0.023	1.9	134	0	-
205	13	32	105	7.3	65	0.93	0.012	1.1	143	3	20
206	26	150	502	8.0	284	2.9	0.038	1.3	105	0	-
210	21	132	476	7.8	254	0.92	0.020	2.7	99	6	8
300	29	135	396	7.7	218	3.5	0.026	1.2	128	8	21
301	24	64	194	7.7	97	1.3	0.026	4.5	146	0	-
302	28	134	352	8.0	200	3.4	0.025	1.3	145	9	15
303	30	100	260	7.4	153	3.4	0.020	0.4	120	10	14
304	28	139	348	7.3	207	3.8	0.020	1.1	145	10	9
305	27	84	227	7.7	126	1.9	0.019	3.9	140	0	-
306	29	140	337	7.9	181	1.8	0.016	3.2	111	1	14
307	21	76	239	7.6	127	1.0	0.020	1.4	125	2	12
308	21	100	436	7.8	233	2.7	0.11	2.1	136	2	10

Table 3. Water chemistry and land use-land cover (LULC) relationships based on results of Spearman rank correlation analysis and non-parametric changepoint analysis (nCPA) on StreamSmart site medians. For both tests, a result is significant if p<0.10. The Spearman rank correlation coefficient (rho) describes the relationship strength and ranges from -1 to 1, with positive and negative values denoting positive and inverse correlations, respectively. The results of nCPA include a changepoint (CP) value with a confidence interval (CI) encompassing the lower (5%) and upper (95%) confidence estimates around the threshold values. The mean of the water chemistry variable values distributed below (left) and above (right) the LULC threshold are also provided.

		--- Spearman ---		------------------------ nCPA ------------------------				
Water Chemistry	**Geospatial**	**p**	**rho**	**p**	**CP (CI)**	**R^2**	**mean left**	**mean right**
Alk (mg/L $CaCO_3$)	% Agriculture	0.36	-	0.11	-	-	-	-
Alk (mg/L $CaCO_3$)	% Forest	<0.001	-0.92	<0.001	59.1 (42.7-61.2)	0.82	125	31
Alk (mg/L $CaCO_3$)	% HDI	<0.001	0.93	<0.001	40.3 (38.4-54.6)	0.82	31	125
Alk (mg/L $CaCO_3$)	% Urban	<0.001	0.87	<0.001	5.2 (4.3-5.4)	0.75	33	123
Alk (mg/L $CaCO_3$)	PHD (house/km^2)	0.60	-	0.100	-	-	-	-
Cond (µS/cm)	% Agriculture	0.28	-	0.24	-	-	-	-
Cond (µS/cm)	% Forest	<0.001	-0.94	<0.001	59.1 (41.0-61.2)	0.72	382	94
Cond (µS/cm)	% HDI	<0.001	0.95	<0.001	45.3 (39.2-57.7)	0.74	107	398
Cond (µS/cm)	% Urban	<0.001	0.87	<0.001	5.2 (4.3-18.4)	0.69	98	379
Cond (µS/cm)	PHD (house/km^2)	0.46	-	0.28	-	-	-	-
pH	% Agriculture	0.66	-	0.40	-	-	-	-
pH	% Forest	<0.001	-0.70	<0.001	71.2 (58.6-74.9)	0.77	7.67	6.84
pH	% HDI	<0.001	0.72	<0.001	27.8 (22.6-33.2)	0.77	6.84	7.67
pH	% Urban	<0.001	0.79	<0.001	4.3 (2.7-4.5)	0.69	7.00	7.71
pH	PHD (house/km^2)	0.80	-	0.32	-	-	-	-
TDS (mg/L)	% Agriculture	0.26	-	0.22	-	-	-	-
TDS (mg/L)	% Forest	<0.001	-0.93	<0.001	45.6 (41.0-61.2)	0.74	241	77
TDS (mg/L)	% HDI	<0.001	0.95	<0.001	45.3 (40.3-61.0)	0.75	63	224
TDS (mg/L)	% Urban	<0.001	0.88	<0.001	5.2 (4.5-19.0)	0.67	60	212
TDS (mg/L)	PHD (house/km^2)	0.43	-	0.19	-	-	-	-
TN (mg/L)	% Agriculture	0.014	0.53	<0.001	43.2 (16.0-49.9)	0.56	0.86	3.53
TN (mg/L)	% Forest	<0.001	-0.77	0.008	59.1 (32.6-65.2)	0.49	2.20	0.45
TN (mg/L)	% HDI	<0.001	0.73	0.005	40.3 (33.5-64.9)	0.49	0.45	2.20
TN (mg/L)	% Urban	0.018	0.51	0.047	5.2 (4.0-5.4)	0.34	0.61	2.06
TN (mg/L)	PHD (house/km^2)	0.097	0.37	0.019	1.0 (0.1-1.0)	0.46	0.89	2.88
TP (mg/L)	% Agriculture	0.60	-	0.070	4.0 (1.2-45.3)	0.36	0.025	0.017
TP (mg/L)	% Forest	0.001	-0.66	0.015	29.9 (10.0-65.8)	0.42	0.029	0.016
TP (mg/L)	% HDI	0.003	0.61	0.034	33.9 (33.5-72.1)	0.37	0.013	0.021
TP (mg/L)	% Urban	0.013	0.53	0.068	4.3 (4.0-49.2)	0.31	0.013	0.021
TP (mg/L)	PHD (house/km^2)	0.99	-	0.66	-	-	-	-
TSS (mg/L)	% Agriculture	0.58	-	0.26	-	-	-	-
TSS (mg/L)	% Forest	0.77	-	0.45	-	-	-	-
TSS (mg/L)	% HDI	0.71	-	0.42	-	-	-	-
TSS (mg/L)	% Urban	0.080	0.39	0.21	-	-	-	-
TSS (mg/L)	PHD (house/km^2)	0.61	-	0.81	-	-	-	-

riffles/bends) or less sensitive macroinvertebrate group (channel flow status) were the response variable (Table 4; Figure 3A-F). The strength of relationships, variability explained, and value of changepoints for macroinvertebrate group sub-

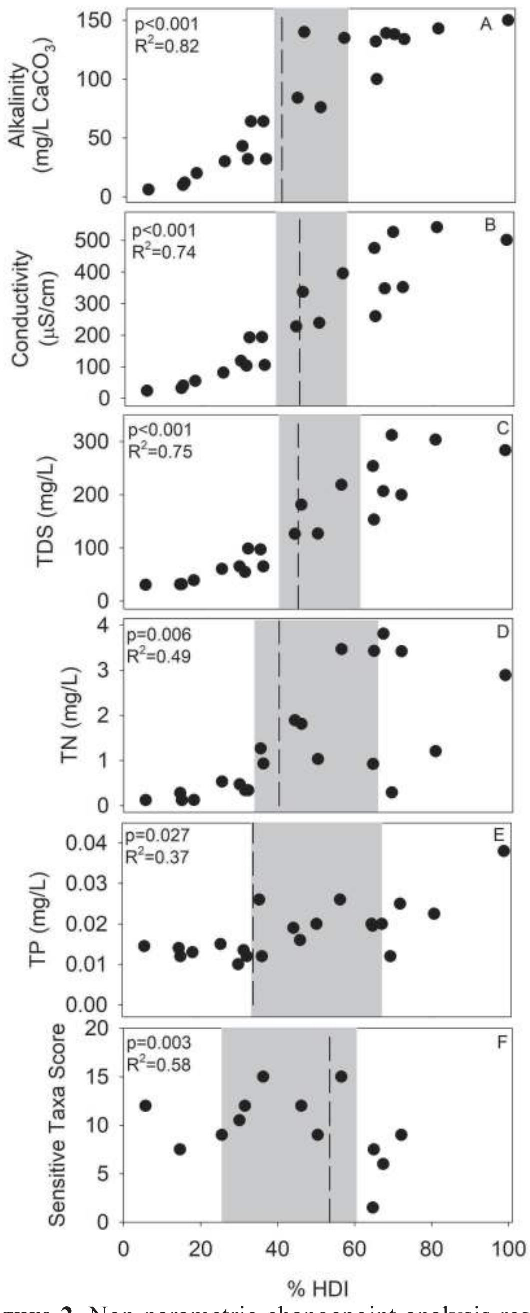

Figure 2. Non-parametric changepoint analysis results for A) alkalinity, B) conductivity, C) total dissolved solids, D) total nitrogen (TN), E) total phosphorus (TP), and F) macroinvertebrate community sensitive taxa showing thresholds and associated confidence interval in the human development index (HDI, %) as dashed lines and gray shaded areas, respectively.

scores was in range with the values observed using the total community index. Further, additional stressors were identified using group sub-scores. These included thresholds in forest land ($p = 0.033$, $R^2 = 0.48$) of 38.5% (CI = 33.8 – 57.1%) and the human development index ($p = 0.099$, $R^2 = 0.19$) of 53.5% (CI = 35.5 – 60.6%), which suggested greater sensitive taxa presence when forest was greatest and the human development index was least. For less sensitive taxa, correlations with agricultural land ($p = 0.098$, rho = 0.48) and TSS ($p = 0.051$, rho = -0.55) were identified, as well as thresholds in bank stability scores ($p = 0.091$, $R^2 = 0.39$) of 11.5 (CI = 10.5 – 15.0) and the total habitat quality index ($p = 0.098$, $R^2 = 0.46$) of 126.5 (CI = 115.5 – 131.0). Both changepoints suggested greater less sensitive taxa presence when habitat quality was greater. For tolerant taxa, a changepoint in sediment deposition scores of 14.5 (CI = 11.0 – 15.0) was detected that suggested greater presence with less sediment deposition.

Categorical and Regression Tree Models

Hierarchy in LULC characteristics was detected for TN, TP, conductivity, and TDS (Figure 4A-C; TDS not shown). For the remaining variables, secondary splits in the data were either not identified or did not reduce relative error beyond the primary split. For TN, the primary LULC predictor was agricultural land, with all of the greatest TN concentrations (n=4, 3.5 mg/L, on average) observed above a threshold of 47% (Figure 4A). For sites with less than 47% agriculture, a secondary split was observed in the human development index, explaining an additional 15% of dataset variability, with the least TN concentrations (n=8; 0.29 mg/L, on average) occurring below 34% human development. A tertiary split in agricultural land = 15.3% was observed in TN concentrations at sites with ≥ 34% human development, but no more than 47% agriculture, that divided intermediate TN concentrations into two groups based on the relative contribution of urban versus agricultural land to the human development index.

For TP, CART identified two thresholds in the human development index, with 71% and 34% as the primary and secondary thresholds, respectively (Figure 4B). This result differed from nCPA, which identified 34% as the most meaningful changepoint, with 71% as the upper bounds of the confidence interval. The CART model suggested that the greatest TP concentrations (n=3, TP = 0.028 mg/L on average) were associated with human development ≥ 71%, accounting for 42% of variability in the dataset. The secondary split associated the smallest TP concentrations (n=8, TP = 0.013 mg/L, on average) with human development < 34% and intermediate TP concentrations (n=10, TP = 0.019 mg/L, on average) with a human development index range of 34 – 71%.

For conductivity, the primary split was at 45% human development (Figure 4C). A secondary split in the human development index = 32% was also observed, suggesting that the least conductivity (n=7, 65 µs/cm, on average) was associated with human development < 32%, while intermediate conductivities (n=4, 180 µs/cm, on average) occurred within a range of 32 – 45%. For sites with a human development index ≥ 45%, a secondary split was also observed at 18% urban. Similar to the CART model for TN, this threshold separated median conductivities for sites above a human development threshold based on relative contributions of agricultural and urban lands to the overall index. The greatest conductivities (n=5, 477 µS/cm, on average) were observed at sites where urban land was > 18% (of at least 45% human development), while conductivities at sites with urban land below that threshold were about 1/3 less (n=5, 319 µS/cm, on average).

Discussion

Implications for StreamSmart and Other Volunteer Monitoring Programs

Our synthesis of stressor-response approaches showed a number of relationships in the StreamSmart database, including thresholds and hierarchy. These findings show the importance of considering multiple types of relationships in stressor-response analysis, and this process could

Table 4. Select macroinvertebrate community index and sensitivity group sub-score relationships with potential biological stressors, including the habitat quality index and components, land use-land cover (LULC), and water chemistry based on results of Spearman rank correlation analysis and non-parametric changepoint analysis (nCPA) on StreamSmart site medians. For both tests, a result is significant if p<0.10, and only statistically significant relationships are shown due to the large number of stressor-response pairs. The Spearman rank correlation coefficient (rho) describes the relationship strength and ranges from -1 to 1, with positive and negative values denoting positive and inverse correlations, respectively. The results of nCPA include a changepoint (CP) value with a confidence interval (CI) encompassing the lower (5%) and upper (95%) confidence estimates around the threshold values. The mean of the water chemistry variable values distributed below (left) and above (right) the LULC threshold are also provided.

		--- Spearman ---		----------------- nCPA -----------------		
Macro metric	**Stressor**	**p**	**rho**	**p**	**R^2**	**CP (CI)**
Total	% Urban	0.047	-0.56	-	-	-
Total	Channel flow status	0.032	0.60	0.032	0.49	11.0 (10.0-13.3)
Total	Epifaunal substrate/available cover	0.035	0.59	0.011	0.48	11.0 (9.5-14.0)
Total	Frequency of riffles/bends	0.045	0.56	-	-	-
Sensitive	% Forest	0.21	-	0.033	0.48	38.5 (33.8-57.1)
Sensitive	% HDI	0.23	-	0.099	0.19	53.5 (25.5-60.6)
Sensitive	Epifaunal substrate/available cover	0.011	0.68	0.003	0.58	12.3 (9.5-14.0)
Sensitive	Frequency of riffles/bends	0.085	0.50	-	-	-
Less Sensitive	% Agriculture	0.098	0.48	-	-	-
Less Sensitive	Bank stability	0.13	-	0.091	0.39	11.5 (10.5-15.0)
Less Sensitive	Channel flow status	0.052	0.55	0.083	0.32	11.5 (10.0-12.8)
Less Sensitive	HQI	0.10	-	0.098	0.46	126.5 (115.5-131.0)
Less Sensitive	TSS	0.051	-0.55	-	-	-
Tolerant	Sediment deposition	0.008	0.70	0.008	0.52	14.5 (11.0-15.0)

be applied to data exploration by other volunteer monitoring groups. We found that it was possible to detect water quality dynamics with only 21 sites for water chemistry and as few as 14 sites for macroinvertebrate data. However, the small number of StreamSmart sites is a limitation, as evidenced by large confidence intervals around many of the threshold estimates (such as in Figure 2D-F). The StreamSmart program can increase statistical power and return on investment by joining program data with datasets from other local, state, or national entities for further analysis and refinement (Stepenuck and Genskow 2018). StreamSmart would also benefit by introducing scientific knowledge "co-creation" to the volunteer experience, which promotes stakeholder buy-in to watershed planning (Buytaert et al. 2014).

Thresholds and CART results may be especially useful for planning tools, as they not only show that water chemistry and LULC are related, but also offer potential target values for managing LULC change to protect water quality. Detected water-chemistry-LULC relationships were consistent with prior studies in the region (Giovanetti et al. 2013; McCarty et al. 2018), as well as global patterns (Lintern et al. 2017), and showed greater nutrients, salts, and sediments with greater human activity in the basin.

Stressor-response relationships were detected for the StreamSmart macroinvertebrate data, but fewer than for water chemistry. The total community index decreased with increasing urban land and decreases in several individual habitat components. Many preceding studies have observed biodiversity loss and community shifts related to habitat quality (Santucci et al. 2005; Stone et al. 2005; Liao et al. 2018), watershed LULC (Weijters et al. 2009; Kuemmerlen et al. 2015), and nutrients (Evans-White et al. 2013). These include volunteer monitoring studies, which also observed relationships with urban land (Fore et al. 2001). The small number of sites

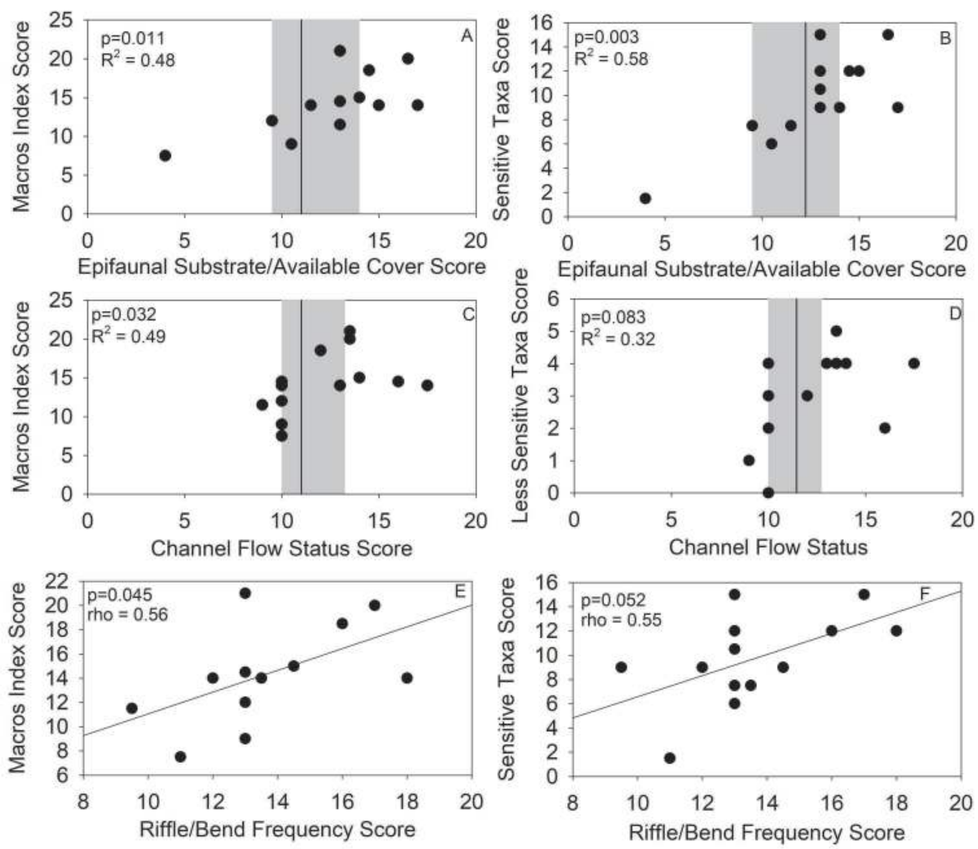

Figure 3. Side-by-side comparison of stressor-response relationships using the full macroinvertebrate community index versus sensitive or less sensitive taxa groups, for habitat component stressors A-B) epifaunal substrate/available cover, C-D) channel flow status, and E-F) riffle/bend frequency score.

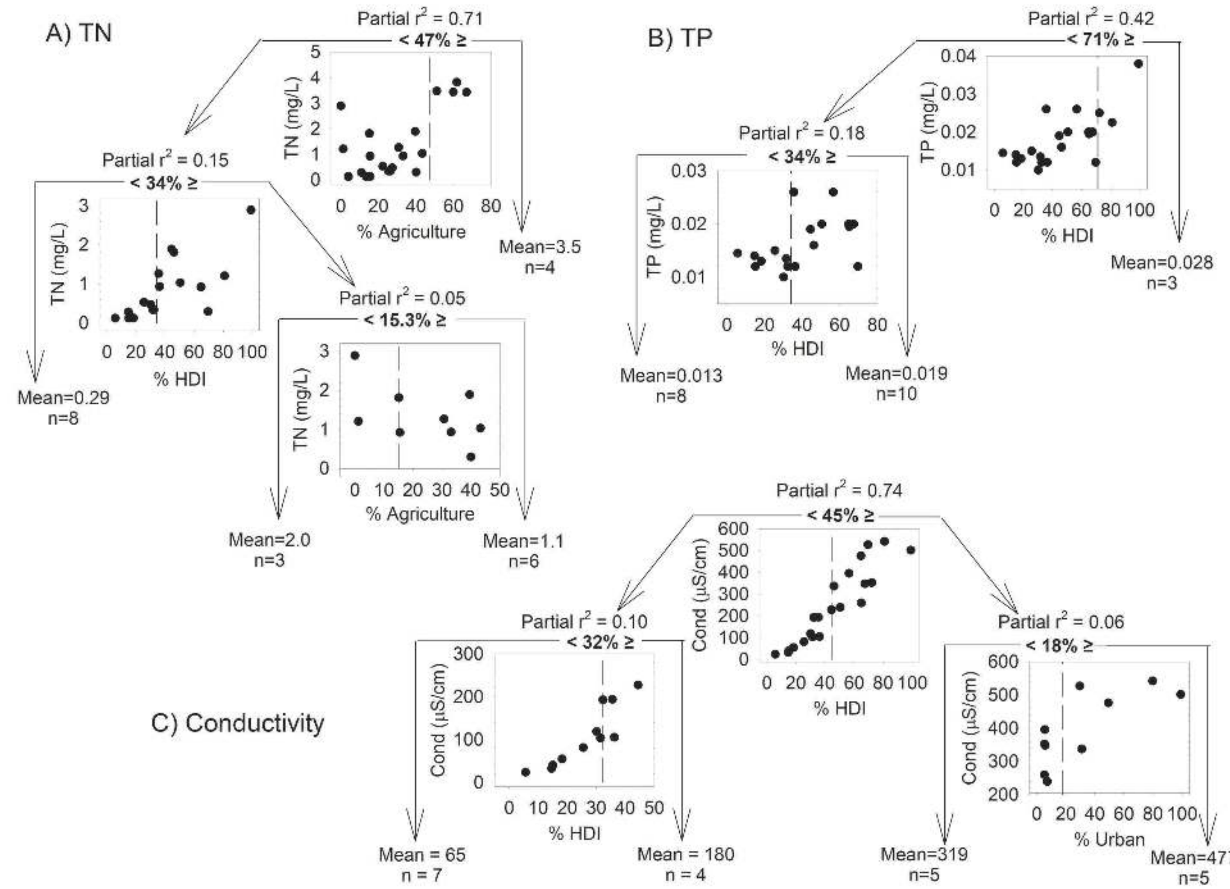

Figure 4. Categorical and regression tree models for A) TN, B) TP, and C) conductivity based on the watershed land use-land cover characteristics agriculture land (%), urban land (%), and agricultural and urban land combined in a human development index (HDI, %).

with macroinvertebrate data likely limited study findings. StreamSmart's unique index also presents a challenge for increasing statistical power and scope by incorporating outside datasets. More traditional monitoring programs use more complex, but standard metrics, such as the USEPA Rapid Bioassessment Protocols Macroinvertebrate Index of Biotic Integrity (RBIBI; Barbour et al. 1999). Validating the StreamSmart index to the RBIBI or other professional indices may be possible (Engel and Voshell, Jr. 2002). Alternately, StreamSmart could explore partnerships with other groups that also want to use a simplified index.

StreamSmart could also go in a different direction and use study findings to streamline the macroinvertebrate and habitat assessments by including just a few key pieces of information. Analysis showed the same meaningful stressor-response relationships were detected when macroinvertebrate sensitivity groups were the response variable, as well as when habitat components were the stressors. These findings suggest that StreamSmart could obtain the same information by assessing only the components that showed meaningful stressor-response relationships. This approach could also be applied by other volunteer monitoring groups with concerns about providing adequate training for more complex assessments, or about keeping the volunteer experience focused on fun and a minimal time investment.

It is not known why the sensitive macroinvertebrate groups or specific habitat components were especially predictive. But, a possible explanation is natural biases (Nerbonne et al. 2008) and inherent properties of the different macroinvertebrates (Nerbonne and Vondracek 2003) that mean volunteers do a better job with these variables. Many sensitive and less sensitive taxa (e.g., mayflies, stoneflies, crayfish) have traits, such as high motility or large bodies, that make them easier to detect, as well as anthropomorphically charismatic features such as visible eyes, legs, gills, and pincers. Even those that are smaller or lack as many charismatic features, such as caddisfly larvae, display engaging and highly identifiable behaviors, such as constructing casings from gravel or leaf litter. By contrast, tolerant taxa tend to have less motility and smaller bodies, as well as association with detritus, that can make both detection and identification more difficult, tedious, and unappealing (Peeters et al. 2022).

Only sedimentation was identified as a stressor for tolerant taxa, and this result was the opposite of the expected relationship, suggesting greater tolerant taxa presence with less sedimentation (i.e., greater scores). However, other studies have shown that community tolerance increases with greater TSS and turbidity (Chase et al. 2017). The StreamSmart data relationship makes the valid point that tolerant taxa are part of the diverse communities associated with good habitat quality. But, it also shows that the function of tolerant taxa in the StreamSmart index is redundant. In general, it may not be possible to capture community tolerance dynamics with simplified identification protocols because information on community tolerance is tied up in metrics using counts and percentages of individuals (Xu et al. 2013), or identification below the Family level is needed (Dusabe et al. 2022).

It is also possible that some habitat components may have a disproportionately large effect in the river networks of the Upper White River Basin. Components identified as stressors were epifaunal substrate/available cover, riffle/bend frequency, and channel flow status, which do have commonality as descriptors of the stream channel itself, rather than bank or riparian attributes. However, another explanation is differences in volunteers' relative understanding of the different components. For example, the strongest habitat predictor was epifaunal substrate/available cover, which may already be encompassed by local knowledge of ideal in-stream habitat for game fishing.

Implications for Watershed Management Planning

Water chemistry-LULC relationships identified in this study can be used to inform programming decisions by H_2Ozarks, Beaver Watershed Alliance, and the Beaver Water District source water protection program, with the caveat that these findings would ideally be refined and strengthened by bringing in additional data sources. For all water chemistry variables, the most meaningful anthropogenic LULC relationship was with the human development index, which combines agricultural and urban lands. Attempts to separately describe agricultural effects showed the advantages of basing management tools for mixed urban-agriculture watersheds on the human development index. Watershed LULC characteristics are interrelated and agricultural land was inversely auto-correlated with both urban and forest land. The opposite effects of these LULC characteristics on water chemistry created noise for low-range agriculture sites that prevented detection of an exclusively agricultural effect on water chemistry, with the exception of TN.

Similarly, uniform nCPA results suggesting consistent and highly predictive thresholds at minimal levels of urban land (4 – 5%) likely reflect that urbanization in the basin tends to occur on prior pasture, where the watershed human development index may already be near or above thresholds for greater water chemistry effects. Indeed, water chemistry-human development index relationships showed little scatter that would evidence such a disproportionate effect of urban land. The CART models for TN and conductivity, however, add nuance to this interpretation, suggesting that a disproportionate effect of urban and agricultural lands may be present, but only in watersheds with human development at or above thresholds of 34 – 45%. Further, urban land thresholds associated with disproportionate effects in CART models were closer to 20% than the 4 – 5% suggested by nCPA.

Thresholds in the human development index may be especially useful watershed management and planning tools because they delineate a level, or range, that suggests an increase in potential risks to water quality due to human activities. Watershed organizations in the basin can prioritize among candidate sites for best management practices by screening for watershed human development index greater than thresholds to determine both the greatest need for mitigation and the greatest potential for return on investment. Human development index thresholds differed between water chemistry variables, though not substantially, ranging from approximately 30 – 45%. These differences could reflect different controls on salt, nutrient, or sediment concentrations, but also data limitations related to a small sample size.

The CART models provide further context for prioritization, with both similarities and differences between variables. For the salts that contribute to conductivity, forest land maintenance of at least 55% (i.e., human development index = 45%) is a potential path to keeping levels low in least-developed watersheds. Additional benefits may accrue if forest is maintained >70%. For more developed watersheds (human development index ≥45%), mitigations may have the greatest efficacy by targeting urban and urbanizing areas with low-impact development or green infrastructure that reduces or slows runoff from impervious surfaces (Carey et al. 2013). However, agricultural conservation practices should also provide benefit.

For nutrients, multiple TP response thresholds to the human development provide forest maintenance targets for both the least and most developed watersheds (~70% and ~30% forest, respectively), as well as multiple human development index ranges to delineate best management practice candidate sites. The CART model for TN differed from other analytes by having a primary split in agricultural land. All sites having median TN concentration > 3.0 mg/L also had > 50% agriculture, making implementation of agricultural conservation practices, specifically in areas where agriculture is greatest, a potential priority route to TN reduction, with the primary goal of reducing runoff and leaching of nutrients from animal manures (Quinn and Stroud 2002). However, a similar disproportionate urban versus agricultural effect to the conductivity model was observed for TN among sites with the greatest human development index, after screening out the sites with the most agriculture. Thus, investment in urban-oriented practices also has potential to make a difference for watershed areas with this LULC profile.

Sediment-LULC relationships were less clear than for other water chemistry variables, with a monotonic TSS response to urban land, but no responses to agriculture or the human development index. The overall low-range of TSS site medians may limit potential to observe LULC effects on in-stream sediment concentrations. StreamSmart data are collected, by design, only under base flow conditions, which has been shown to be effective for identifying sub-watersheds with greater nutrient concentrations, while also allowing for broadening monitoring coverage (McCarty and Haggard 2016). This analysis suggests potential limits to this approach for TSS; a more pronounced TSS gradient, as well as LULC linkages, might better be detected under storm flow conditions.

Conclusions

Study results show that StreamSmart volunteers are providing a valuable water quality monitoring service. The value of a high-quality water chemistry baseline dataset and LULC relationships for the Upper White River Basin accrues return on investment by StreamSmart volunteers, H_2Ozarks, and partner organizations, with the potential to inform watershed management planning through "co-creation" of decision-making support tools with other monitoring entities. The volunteer monitoring macroinvertebrate community and habitat quality data were also informative about water quality dynamics in the basin, though potentially limited by sample size and complexities around combining with outside data. Stressor-response relationships detected for sensitive and less sensitive macroinvertebrate groups, as well as habitat components, suggest potential bias in how the StreamSmart volunteers conduct these assessments, but also offer an approach to streamlining these complex tasks. StreamSmart, as well as similar volunteer monitoring programs, can leverage this information to improve the volunteer experience and find the best avenues for communication with stakeholders through the aspects of watershed science, habitat quality, and biodiversity that are already encompassed by local knowledge.

Acknowledgements

The following student authors, listed alphabetically, contributed to preliminary data analysis and manuscript development: Lillie Haddock, Tarah Inema, Solomon Isu, Mahmood Jebur, Ian Kennedy, Kyle Lawrence, Marguerita Leavitt, Oliva Liedel, Marret Lineberry, John Richins, and Holly Wren. We thank our partners at H_2Ozarks, Beaver Water District, and Beaver Watershed Alliance, especially the StreamSmart program coordinator Erin Scott. Funding for this project was provided through Section 104(b) of the Water Resources Research Act of 1984 administered by the United States Geological Survey, as well as the University of Arkansas System Division of Agriculture.

Author Bio and Contact Information

Mrs. Erin Grantz (corresponding author) is a program manager and data scientist with the Arkansas Water Resources Center, and she is currently working on her Ph.D. in Environmental Dynamics at the University of Arkansas. She drafted the manuscript and helped manage the data analysis for BENG 4973/5973, including the development of the classification and regression trees. She has an M.S. and B.S. in Crop, Soil and Environmental Sciences from the University of Arkansas, and her graduate research focuses on water quality and watershed stoichiometry. She may be contacted at egrantz@uark.edu or 1371 West Altheimer Drive, Don Tyson Center for Agricultural Sciences, Fayetteville, AR 72704, USA.

Dr. Brian Haggard has been focused on water quality issues since his graduate studies. He is currently the director of the Arkansas Water Resources Center and a professor in the Biological and Agricultural Engineering Department with the University of Arkansas System Division of Agriculture. Prior to his current positions, he worked for the US Geological Survey and USDA Agricultural Research Service in hydrology and water quality. He conceptualized using this data in BENG 4973/5973 and the statistical analysis. He has a Ph.D. in Biosystems Engineering from Oklahoma State University, a M.S. in Agronomy from the University of Arkansas, and a B.S. in Life Sciences from the Missouri University of Science and Technology. He may be contacted at haggard@uark.edu or Engineering Hall 203, University of Arkansas, Fayetteville, AR 72701, USA.

References

Arkansas GIS Office. 2014. All Cultural Feature (point). Available at: https://gis.arkansas.gov/prodcut/all-cultural-feature-point/. Accessed March 3, 2023.

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99-002. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

Buytaert, W., Z. Zulkafli, S. Grainger, L. Acosta, C.A. Tilashwork, J. Bastiaensen, B. De Bievre, J. Bhusal, J. Clark, A. Dewulf, M. Foggin, D.M. Hannah, C. Hergarten, A. Isaeva, T. Karpouzoglou, B. Pandeya, D. Paudel, K. Sharma, T. Steenhuis, S. Tilahun, G. Van Hecken, and M. Zhumanova. 2014. Citizen science in hydrology and water resources: Opportunities for knowledge generation, ecosystem service management, and sustainability development. *Frontiers in Earth Science* 2. Available at: https://www.frontiersin.org/articles/10.3389/feart.2014.00026/full. Accessed February 22, 2023.

Carey, R.O., G.J. Hochmuth, C.J. Martinez, T.H. Boyer, M.D. Dukes, G.S. Toor, and J.L. Cisar. 2013. Evaluating nutrient impacts in urban watersheds: Challenges and research opportunities. *Environmental Pollution* 173: 138-149. Available at: https://doi.org/10.1016/j.envpol.2012.10.004. Accessed February 22, 2023.

Chase, J.W., G.A. Benoy, and J.M. Culp. 2017. Combined effects of nutrient enrichment and inorganic sedimentation on benthic biota in an experimental stream. *Water Quality Research Journal* 52(3): 151-165.

De'Ath, G. and K.E. Fabricius. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology* 81(11): 3178-3192. DOI: 10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2.

Dewitz, J. and U.S. Geological Survey. 2021. National Land Cover Database (NLCD) 2019 Products (ver. 2.0, June 2021): U.S. Geological Survey data release. Available at: https://doi.org/10.5066/P9KZCM54. Accessed February 22, 2023.

Dusabe, M.C., T.A. Neubauer, F.A. Muvundja, B.L. Hyangya, and C. Albrecht. 2022. Family-level bio-indication does not detect the impacts of dams on macroinvertebrate communities in a low-diversity tropical river. *Frontiers in Environmental Science* 10: 902246. DOI: 10.3389/fenvs.2022.902246.

Engel, S.R. and J.R. Voshell Jr. 2002. Volunteer biological monitoring: Can it accurately assess the ecological condition of streams? *American Entomologist* 48: 164-177.

Evans-White, M.A., B.E. Haggard, and J.T. Scott. 2013. A review of stream nutrient criteria development in the United States. *Journal of Environmental Quality* 42(4): 1002-1014. DOI: 10.2134/jeq2012.0491.

Fore, L.S., K. Paulsen, and K. O'Laughin. 2001. Assessing the performance of volunteers in monitoring streams. *Freshwater Biology* 46: 109-123.

Giovannetti, J., L.B. Massey, B.E. Haggard, and R.A. Morgan. 2013. Land use effects on stream nutrients at Beaver Lake Watershed. *Journal of the American Water Works Association* 105(1): E1-E10. Available at: https://awwa.onlinelibrary.wiley.com/doi/abs/10.5942/jawwa.2013.105.0005. Accessed February 22, 2023.

Haggard, B.E. 2010. Phosphorus concentrations, loads, and sources within the Illinois River drainage area, Northwest Arkansas, 1997 – 2008. *Journal of Environmental Quality* 39: 2113-2120. DOI: 10.234/jeq2010.0049.

Hoyer, M.V. and D.E. Canfield, Jr. 2021. Volunteer-collected water quality data can be used for science and management. *Lake and Reservoir Management* 37: 235-245. DOI: 10.1080/10402381.2021.1876190.

King, R.S. and C.J. Richardson. 2003. Integrating bioassessment and ecological risk assessment: An approach to developing numerical water-quality criteria. *Environmental Management* 31(6): 795-809. DOI: 10.1007/s00267-002-0036-4.

Kuemmerlen, M., B. Schmalz, Q. Cai, P. Haase, N. Fohrer, and S.C. Jähnig. 2015. An attack on two fronts: Predicting how changes in land use and climate affect the distribution of stream macroinvertebrates. *Freshwater Biology* 60(7): 1443-1458. DOI: 10.1111/fwb.12580.

Lewandowski, E. and H. Specht. 2015. Influence of volunteer and project characteristics on data quality of biological surveys. *Conservation Biology* 29(3): 713-723. Available at: https://doi.org/10.1111/cobi.12481. Accessed February 22, 2023.

Liao, H., E. Sarver, and L.-A.H. Krometis. 2018. Interactive effects of water quality, physical habitat, and watershed anthropogenic activities on stream ecosystem health. *Water Research* 130: 69-78. Available at: https://www.sciencedirect.com/science/article/pii/S0043135417309934?via%3Dihub. Accessed February 22, 2023.

Lintern, A., J.A. Webb, D. Ryu, S. Liu, U. Bende-Michl, D. Water, P. Leahy, P. Wilson, and A.W. Western. 2017. Key factors influencing differences in stream water quality across space. *WIRES Water* 5: e1260. Available at: https://doi.org/10.1002/wat2.1260. Accessed February 22, 2023.

McCarty, J.A. and B.E. Haggard. 2016. Can we manage non-point source pollution using nutrient concentrations during seasonal baseflow? *Agricultural and Environmental Letters* 1(1): 160015. DOI: 10.2134/ael2016.03.0015.

McCarty, J.A., M.D. Matlock, J.T. Scott, and B.E. Haggard. 2018. Risk indicators for identifying critical source areas in five Arkansas watersheds. *Transactions of the ASABE* 61(3): 1025-1032. DOI: 10.13031/trans.12699.

Multi-Resolution Land Characteristics Consortium (MRLC). 2018. National Land Cover Database 2011 (NLCD 2011). Multi-Resolution Land Characteristics Consortium (MRLC). Available at: https://data.nal.usda.gov/dataset/national-land-cover-database-2011-nlcd-2011. Accessed February 22, 2023.

Nerbonne, J.F. and B. Vondracek. 2003. Volunteer macroinvertebrate monitoring: Assessing training needs through examining error and bias in untrained volunteers. *Journal of the North American Benthological Society* 22: 152-163.

Nerbonne, J.F., B. Ward, A. Ollila, M. Williams, and B. Vondracek. 2008. Effect of sampling protocol and volunteer bias when sampling for macroinvertebrates. *Journal of the North American Benthological Society* 27: 640-646. DOI: 10.1899/07-101.1.

Northwest Arkansas Regional Planning Commission (NWARPC). 2016. 2040 Northwest Arkansas Metropolitan Transportation Plan. Available at: https://www.nwarpc.org/datacatalog/2040-northwest-arkansas-metropolitan-transportation-plan/. Accessed February 22, 2023.

Peeters, E.T.H.M., A.A.M. Gerritsen, L.M.S. Seelen, M. Begheyn, F. Rienks, and S. Teurlincx. 2022. Monitoring biological water quality by volunteers complements professional assessments. *PLoS ONE* 17: e0263899. DOI: 10.1371/journal.pone.0263899.

Perez, A.P., K. Higgins, and M. Freyaldenhoven. 2015. Nonpoint Source Pollution in the Beaver Reservoir/Upper White River Watershed. FSPPC125. University of Arkansas Cooperative Extension Service. Available at: https://www.uaex.uada.edu/publications/pdf/FSPPC125.pdf. Accessed February 23, 2023.

Qian, S.S., R.S. King, and C.J. Richardson. 2003. Two statistical methods for the detection of environmental thresholds. *Ecological Modelling* 166(1-2): 87-97. Available at: https://doi.org/10.1016/S0304-3800(03)00097-8. Accessed February 22, 2023.

Quinn, J.M. and M.J. Stroud. 2002. Water quality and sediment and nutrient export from New Zealand hill-land catchments of contrasting land use. *New Zealand Journal of Marine and Freshwater Research* 36(2): 409-429.

R Core Team. 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/. Accessed February 23, 2023.

San Llorente Capdevila, A., A. Kokimova, S.S. Ray, T. Avellán, J. Kim, and S. Kirschke. 2020. Success factors for citizen science projects in water quality monitoring. *Science of the Total Environment* 728: 137843

Santucci, Jr., V.J., S.R. Gephard, and S.M. Pescitelli. 2005. Effects of multiple low-head dams on fish, macroinvertebrates, habitat, and water quality in the Fox River, Illinois. *North American Journal of Fisheries Management* 25: 975-992. DOI: 10.1577/M03-216.1.

Savan, B., A.J. Morgan, and C. Gore. 2003. Volunteer environmental monitoring and the role of the universities: The case of Citizens' Environment Watch. *Environmental Management* 31(5): 561-568. Available at: https://doi.org/10.1007/s00267-002-2897-y. Accessed February 23, 2023.

Scott, J.T., B.E. Haggard, A.N. Sharpley, and J.J. Romeis. 2011. Change point analysis of phosphorus trends in the Illinois River (Oklahoma) demonstrates the effects of watershed management. *Journal of Environmental Quality* 40: 1249-1256. DOI: 10.2134/jeq2010.0476.

Stepenuck, K.F. and K.D. Genskow. 2018. Characterizing the breadth and depth of volunteer water monitoring programs in the United States. *Environmental Management* 61: 46-57. DOI: 10.1007/s00267-017-0956-7.

Stone, M.L., M.R. Whiles, J.A. Webber, K.W.J. Williard, and J.D. Reeve. 2005. Macroinvertebrate communities in agriculturally impacted Southern Illinois streams: Patterns with riparian vegetation, water quality, and in-stream habitat quality. *Journal of Environmental Quality* 34: 907-917. DOI: 10.2134/jeq2004.0305.

Storey, R.G., A. Wright-Stow, E. Kin, R. Davies-Colley, and R. Stott. 2016. Volunteer stream monitoring: Do the data quality and monitoring experience support increased community involvement in freshwater decision making? *Ecology and Society* 21(4): 32. Available at: https://doi.org/10.5751/ES-08934-210432. Accessed February 23, 2023.

Therneau, T. and B. Atkinson. 2019. rpart: Recursive Partitioning and Regression Trees. R package version 4.1-15. Available at: https://CRAN.R-project.org/package=rpart. Accessed February 23, 2023.

United States Environmental Protection Agency (USEPA). 2005. Community-Based Watershed Management: Lessons from the National Estuary Program. EPA-842-B-05-003. Available at: https://www.epa.gov/nep/community-based-watershed-management-handbook. Accessed February 23, 2023.

United States Environmental Protection Agency (USEPA). 2008. Handbook for Developing Watershed Plans to Restore and Protect Our Waters. EPA-841-B-08-002. Available at: https://www.epa.gov/nps/handbook-developing-watershed-plans-restore-and-protect-our-waters. Accessed February 23, 2023.

Weijters, M.J., J.H. Janse, R. Alkemade, and J.T. Verhoeven. 2009. Quantifying the effect of catchment land use and water nutrient concentrations on freshwater river and stream biodiversity. *Aquatic Conservation: Marine and Freshwater Ecosystems* 19(1): 104-112. DOI: 10.1002/aqc.989.

White, J.Y. and C.J. Walsh. 2020. Catchment-scale urbanization diminishes effects of habitat complexity on instream macroinvertebrate assemblages. *Ecological Applications* 30(8): e02199. DOI: 10.1002/eap.2199.

Xu, M., Z. Wang, X. Duan, and B. Pan. 2013. Effects of pollution on macroinvertebrates and water quality bio-assessment. *Hydrobiologia* 729(1): 247-259. DOI: 10.1007/s10750-013-1504-y.