Appendix G – Stream Classification

Contents

This appendix contains a report written by Julie Zimmerman (previous of The Nature Conservancy) in 2010 and titled *Susquehanna River stream classification analysis: determining a stream classification system to apply to the Potomac River and Chesapeake Bay*. The analysis described in the report was intended to help determine the need for stream classifications in the ELOHA analysis of the MPRWA project.

The disc directory contains only this document.

Introduction

The Potomac River, larger Chesapeake Bay watershed, and other large watersheds have the need for a stream classification system that can be applied to large river basins. The need for stream classifications include: (1) general ecological descriptions of stream types, (2) the development of management units for water management plans, and (3) developing relationships between hydrologic alteration and ecological response (through the ELOHA process). When developing a stream classification system to use in ecological analyses, such as developing relationships between hydrologic alteration and ecological response, the goal is to group streams into types that are ecologically and/or hydrologically similar with the assumption that species in each type will respond similarly to perturbation. In addition, the number of stream types generated by a stream classification system should be manageable, appropriate for the objectives of the analysis, and account for needs of statistical analyses, such as sufficient sample size of biological data.

There are generally two approaches useful when classifying streams, as well as hybrid approaches. The first approach is to develop classes based on unaltered hydrology, assuming that aquatic species and communities will respond to similar hydrologic signatures. For example, streams with high groundwater input, cold temperatures, and relatively stable flows will have species assemblages with similar traits and similar tolerance to certain types of hydrologic alteration. These assemblages are assumed to differ from the assemblages found in streams with high flood magnitudes in the spring, low summer flows, and high overall flow variability. The second approach to stream classification is to develop classes based on stream or landscape attributes, such as temperature, stream gradient, geology, and stream size. Again, the assumption is that species assemblages will reflect the combination of attributes in a stream class. Often, stream and landscape attributes will predict the hydrologic signature of a stream class. For example, streams with cold water temperature, small stream size, and high percent limestone geology would tend to be streams with high groundwater input and relatively stable flows.

The first approach to stream classification (hydrology-based) requires data on unaltered daily stream flow that can be used in a statistical clustering procedure to group streams based on hydrologic metrics. The method is statistically-based and data-driven, but requires ecological knowledge and interpretation to develop the final set of stream classes (i.e., to decide how many classes is the "right" number, what set of classes make ecological sense). In addition, some kind of hydrologic foundation is required to provide unaltered streamflow data (e.g., a rainfall-runoff model). The second approach (attribute-based) also requires data, but it is more widely available. The Nature Conservancy's (TNC) Eastern Regional Office recently developed an attribute-based stream classification system for the Eastern Region that can be adapted to provide attribute-based stream classifications using a wide variety of stream- and landscape-level variables (Olivero & Anderson 2008). This method also requires ecological knowledge and interpretation to develop the final set of stream classes, but it is easier to implement because no statistical analyses or model-derived data are necessary.

The goal for this analysis is to compare a set of stream classification systems to choose the appropriate system for application to the Potomac River basin. The objectives of the stream classification system are to (1) classify streams that are hydrologically and ecologically similar, (2) allow for statistical analyses of hydrologic alteration and ecological response by stream type, and (3) ultimately units for management of water withdrawals and land use, based on the results of hydrologic alteration - ecological response analyses. All stream classification systems examined in this analysis were based on GIS-derived stream and landscape-based attributes, using combinations of variables available in the Northeast Aquatic Habitat Classification System (NEAHCS) developed by TNC and the Northeast Association of Fish and Wildlife Agencies (Olivero & Anderson 2008). Our approach was to re-analyze statistical relationships between hydrologic alteration and ecological response that were developed for the Susquehanna River without using any stream classification system (Apse *et al.* 2008; Case study 5), this time applying various stream classification systems to determine a system (if any) that improved the analyses and could be applied to the Potomac River. In all, we compared 9 different stream classifications for (1) the total number of stream classes and sample size of biological data points in each class for the Susquehanna analysis, (2) the total number of stream classes and likely sample size of biological data points in each stream class when applied to the Potomac River, and (3) the relative ability to strengthen the flow alteration-ecological response curves developed for the Susquehanna River (Apse *et al.* 2008; Case study 5).

Methods

Stream classification systems included in analysis

The Northeast Aquatic Habitat Classification System (Olivero & Anderson 2008) determined that the key variables in a stream classification system were stream size, gradient, geology, and temperature. The NEAHCS report suggested that the most useful stream classification for the eastern U.S. would include these four variables, and provided the number of categories that should be considered for each variable. In addition, the report prioritized the variables (in the order: size, gradient, geology, and temperature) and outlined how classes within each variable could be collapsed to simplify the stream classification system to fit specific project objectives. We chose to examine a simplified version of the stream classification system suggested by the NEAHCS report (size, 4 levels; gradient, 4 levels; geology, 3 levels; and temperature, 3 levels) as well as a classification system using the three of the four most important variables (size, 4 levels; gradient, 4 levels; and geology, 3 levels).

In addition to the classifications suggested by the NEAHCS report, we were interested in grouping streams that were hydrologically similar, and felt that a variable that represented groundwater input would help achieve that goal. Thus, we included a size and baseflow index classification and a size and percent karst classification. Baseflow index was a variable considered to be a key secondary descriptor for stream classifications by the NAHCS report and the variable that most closely approximated groundwater input. However, because baseflow index was interpolated between stream gages and was in some cases based on altered hydrology, we also included percent karst as a variable that can approximate groundwater input to streams. In addition, a CART analysis conducted by ICPRB (Appendix E) found significant relationships between percent karst in a watershed and changes in hydrologic metrics, with 1.14 percent karst identified as the threshold above which karst can significantly impact hydrology.

Other stream classification systems that were considered included bioregion and stream size alone (both 4 levels and 2 levels). These classifications were examined for their potential to provide a simple system with few classes that still has the ability to group streams by differences in hydrology and ecology.

In a related analysis, the author examined how well the variables included in NEAHCS could predict stream classes developed by USGS using the HIP procedure for index stream gages in Pennsylvania (see Apse *et al.* 2008; Case study 2 for information on the HIP analysis for Pennsylvania stream gages). The HIP procedure is statistically-based and uses a cluster analysis on hydrologic data (151 indices) to classify streams based on similar hydrology. HIP resulted in five stream classes with distinguishing hydrologic characteristics related to flow magnitude, variability, and frequency. A drawback to using the HIP procedure to develop stream classifications is the need for unregulated hydrology to develop stream classes; thus, only sites with unregulated hydrologic data can be classified. A benefit of using NEAHCS is that the GIS-derived variables can be applied everywhere and any site of interest can be placed in a stream class. Thus, Zimmerman used a discriminant function analysis to predict HIP-derived stream classes using variables in NEAHCS to determine if similar stream classes (with similar hydrologic characteristics) could be generated by both classification methodologies and which NEAHCS variables were the most important predictors of HIP stream classes. Variables from NEAHCS could correctly classify 62 percent of sites into HIP-generated stream classes (correct classification rates varied by stream class) and the variables that had the highest correlations with stream classes (i.e., were most useful for predictions) included drainage area, baseflow index, and area weighted mean precipitation. This analysis further justified inclusion of stream size and baseflow index in a stream classification system. Area-weighted mean precipitation was not included in any system examined in this analysis, but it should be considered for examination in any future stream classification analyses.

Overall, nine classification systems were examined in the current analysis for the Susquehanna basin based on combinations of seven variables. Descriptions of the variables used in the stream classification analysis are provided in Table 1.

Modeling relationships between hydrologic alteration and ecological response

In a previous study, Apse et al. (2008; Case study 5) used cumulative water withdrawal data for sites in the Susquehanna River basin to develop an index of cumulative water use relative to the 7Q10 (the 7-day low flow that is predicted to occur every 10 years on average). These water withdrawal data (calculated as Withdrawal Index, Table 2) were paired with macroinvertebrate sampling and water quality data collected at the same locations. The project team used this database to develop statistically-based, quantitative estimates of ecological response to water withdrawals (see Table 2 for all variables used in this analysis).

Regression models were constructed to examine the relationships between water withdrawals and 13 macroinvertebrate metrics (7 monitoring metrics and 6 functional trait metrics; Table 2). Of the 13 ecological response variables, 9 were expressed as a proportion. For these nine variables, logistic regression models were examined using the events/trials syntax (PROC LOGISTIC; SAS Institute Inc. 2009). For each sample, the response variable was expressed as the number of individuals that met the criteria of the variable (events) divided by the total number of individuals in the sample (trials). For example, for percent Ephemeroptera, the response variable was expressed as the number of Ephemeroptera (events) divided by the total number of macroinvertebrates in the sample (trials), rather than as a proportion. R^2 values for logistic regression models were derived by modeling the relationship between observed data and values predicted by the model. For the remaining four ecological response variables that were not expressed as proportions, linear regression was used to examine the relationship between each response and the predictor variables (PROC GLM; SAS Institute Inc. 2009). Water quality covariates (pH and DO) were also included in all models. More detailed descriptions of variables used in the analysis, methods used in developing statistical models, and flow-ecology hypotheses that led to the selection of variables are in Apse et al. (2008).

Stream classification analysis

The original analysis of hydrologic alteration and ecological response included sites throughout the Susquehanna River watershed, without any grouping based on stream classes. All sites in the analysis were classified according to the nine different classification systems described above and outlined in Table 3 to determine if applying a stream classification to the Susquehanna River data would improve the relationships between hydrologic alteration and ecological response by accounting for natural hydrologic variability among sites. The total number of stream classes, number and percent of classes with $n>15$, number and percent of total sites included in the analysis, percent of models with significant (>0.05) p-values, and percent of models with high R^2 were compared among stream classification systems to assess the performance of each system.

Table 1. Definitions of variables included in the steam classification systems examined in this analysis.

Variable	# Classes	Descriptions	Definitions	Notes
Size	4	Headwater/creek	0 – 38.6 mi^2	Most important stream classification variable according to Olivero and Anderson (2008)
		Small river	38.6 – 200 mi^2	
		Medium river	200 – 3861 mi^2	
		Large/great river	> 3861 mi^2	
Size	2	Headwater/creek	0 – 38.6 mi^2	Collapsed version of the NEAHCS size classes
		Small/medium/large river	> 38.6 mi^2	
Gradient	4	Very low/low	< 0.1%	Second most important stream classification variable according to Olivero and Anderson (2008)
		Moderate – low	≥ 0.1 < 0.5%	
		Moderate – high	≥ 0.5 < 2.0%	
		High/very high	≥ 2.0%	
Geology	3	Acidic/low buffered	100-174 Norton Index	Scores based on total upstream geology Norton Index; Size 3, 4, 5, rivers (medium, large, great rivers) were assumed to be neutral; third most important stream classification variable according to Olivero and Anderson (2008)
		Neutral/moderately buffered	175-324	
		Calc-neutral/highly buffered	325-400	
Temperature	3	Cold	Proportion of coldwater species and habitats >50%	Categories based on species composition and habitat; fourth most important stream classification variable according to Olivero and Anderson (2008)
		Transitional cool	Increasing proportion of cool and warmwater species, decreasing coldwater habitat	
		Warm	Increasing dominance of warm species, unable to support resident coldwater species	

Table 1. Continued.

Variable	# Classes	Descriptions	Definitions	Notes
Baseflow index	7	1 2 3 4 5 6 7	0-35% 35-40% 40-45% 45-50% 50-60% 60-65% 65-100%	Calculated by TNC by interpolating baseflow index values estimated at USGS stream gages; discussed by Olivero and Anderson as being a key secondary descriptor for stream classifications; found to be a better predictor of groundwater input than geology (e.g. % calcareous)
Baseflow index	3	Low Moderate High	0-35% 35-50% 50-100%	Calculated by TNC by interpolating baseflow index values estimated at USGS stream gages; discussed by Olivero and Anderson as being a key secondary descriptor for stream classifications; found to be a better predictor of groundwater input than geology (e.g. % calcareous)
Bioregion	4	N. Appalachian Ridges Valleys Piedmont		See bioregion descriptions in Appendix C. "Northern Appalachian" combines the Northern Appalachian Plateau and Uplands and the North Central Appalachian EPA Level 3 ecoregions.
Karst	2	Low High	< 1.14% > 1.14%	Calculated as % calcareous geology, including limestone, dolomite, dolostone, and other carbonate-rich rocks, based on USGS bedrock geology maps (Olivero & Anderson 2008); Levels (1.14% threshold) are based on the preliminary CART analyses (Appendix E)

Table 2. Variables used to examine relationships between hydrologic alteration and ecological response. Species trait metrics (metrics 11-16) are from Poff et al. (2006). For a more detailed description of analyses, see case study 5 in Apse et al. (2008).

Metric	Type	Description
Withdrawal Index	Predictor	TotalWaterUse/TotalSurfaceWater. Total Water Use (gallons per day) is an estimate of the total water withdrawn upstream of each pour point based on registered water withdrawals. Withdrawals were adjusted to reflect variable monthly withdrawals. Total Surface Water (gallons per day) is a modeled estimate of the 7Q10 at each pour point based on the regression models developed by USGS and described by Stuckey (2006). Withdrawal Index is a simple ratio between the water withdrawn and the 7Q10 to show the proportion of 7Q10 that is withdrawn for use. This is the withdrawal index used by Freeman and Marcinek (2006) to develop a relationship between water use and fish assemblages.
pH	Covariate	water quality metric measuring pH
DO	Covariate	water quality metric measuring dissolved oxygen
Richness	Response	The total number of taxa present in the sample at each site. Number decreases with increasing stress.
Modified Hilsenhoff Biotic Index (HBI)	Response	A measure of the organic pollution tolerance of a benthic macroinvertebrate community. Index value increases with increasing stress.
Ephemeroptera count	Response	The total number of Ephemeroptera in the sample at each site. Count decreases with increasing stress.
Dominant Taxa	Response	Percentage of the taxon with the largest number of individuals out of the total number of macroinvertebrates in the sample. Percentage increases with increasing stress.
EPT Index	Response	The total number of Ephemeroptera (mayfly), Plecoptera (stonefly), and Trichoptera (caddisfly) taxa present in the sample at each site. Number decreases with increasing stress.
Chironomidae count	Response	The total number of Chironomidae in the sample at each site. Count increases with increasing stress.
Shannon-Wiener Diversity Index	Response	A measure of biological community complexity based on the number of equally or nearly equally abundant taxa in the community. Index value decreases with increasing stress.
Drift3 count	Response	The total number of individuals that are abundant in the drift, based on species traits. Count increases with increasing disturbance.
Size1 count	Response	The total number of individuals that are small sized at maturity, based on species traits. Count increases with increasing disturbance.

Table 2. Continued.

Metric	Type	Description
Volt3 count	Response	The total number of individuals that are bi- or multi-voltine, based on species traits. Count increases with increasing disturbance.
Trop1 count	Response	The total number of individuals that are collector-gatherers, based on species traits. Count decreases with increasing disturbance.
Ther1 count	Response	The total number of individuals that are cold stenothermal or cool eurythermal, based on species traits. Count decreases with increasing disturbance.
Rheo1 count	Response	The total number of individuals that are obligate in depositional habitats, based on species traits. Count increases with increasing disturbance.

Table 3. Summary statistics for stream classification systems included in the analysis, including variables comprising each system, the total number of stream classes, the number of stream classes included in the analysis (classes with n>15, with percent of total classes in parentheses), and the number of sites with biological sites included in the analysis (sites included in stream classes with n>15, percent of total sites in parentheses). Also included are the percentage of models examined for each classification system that had significant associations between hydrologic alteration and ecological response and the percent of those significant models that also had an R^2 value greater than 0.30 (indicating that greater than 30 percent of the variation in the ecological response data was explained by hydrologic alteration).

Classification system	Variables included	Total # classes	Classes with n > 15	Sites used in analysis	% models with p<0.05	% models with R^2>0.30
(1) CLSIMP4433	Stream size (4) Gradient (4) Geology (3) Temperature (3)	30	7 (23%)	224 (78%)	55%	24%
(2) CLSIMP433	Stream size (4) Geology (3) Temperature (3)	12	5 (42%)	254 (89%)	68%	16%
(3) BSFLCL_7_S	Baseflow index (7) Stream size (4)	17	15 (53%)	252 (88%)	49%	14%
(4) BSFLCL_3_S	Baseflow index (3) Stream size (4)	7	5 (71%)	278 (97%)	54%	23%
(5) BIOREGIO_5	Bioregion (4)	4	4 (100%)	287 (100%)	50%	4%
(6) SIZE_4/KARST	Stream size (4) Karst (2)	6	5 (83%)	278 (97%)	60%	18%
(7) SIZE_2/KARST	Stream size (2) Karst (2)	4	4 (100%)	287(100%)	69%	3%
(8) SIZE ONLY_4	Stream size (4)	3	3 (100%)	287 (100%)	62%	13%
(9) SIZE ONLY_2	Stream size (2)	2	2 (100%)	287(100%)	69%	0%

Results

Nine stream classification systems were examined for a) the total number of stream classes and sample size of biological/hydrologic data points in each stream class for the Susquehanna basin and b) improvements in the explanatory power of relationships between hydrologic alteration (measured by Withdrawal Index) and ecological response for the Susquehanna River. The nine classification systems were numbered for discussion purposes:

(1) Stream size (4 levels), gradient (4 levels), geology (3 levels) and temperature (3 levels)
(2) Stream size (4 levels), geology (3 levels), and temperature (3 levels)
(3) Baseflow index (7 levels) and stream size (4 levels)
(4) Baseflow index (3 levels) and stream size (4 levels)
(5) Bioregion (4 levels)
(6) Stream size (4 levels) and karst (2 levels)
(7) Stream size (2 levels) and karst (2 levels)
(8) Stream size (4 levels)
(9) Stream size (2 levels)

Classification systems 1, 2, and 3 yielded high numbers of stream classes for the Susquehanna River with 30 classes for system 1; 12 classes for system 2, and 17 classes for system 3 and a corresponding low number of classes with a sufficient sample size for statistical analyses (Range: 23-53 percent). Classification systems 4 through 9 each had a total number of stream classes of less than 10 (range: 2 – 7) and a higher proportion of classes with $n>15$ (range: 71-100 percent). All nine classification systems had similar proportions of models with significant p-values (<0.05) for Withdrawal Index (range: 49-69 percent). However, there was greater variation in the proportion of total models with a high R^2 ($R^2>0.30$; range 0-24 percent), indicating models with stronger relationships between Withdrawal Index and biological response variables and higher predictive ability. Classification system 1 had the highest proportion of models with high R^2 values (24 percent), followed by classification system 4 (23 percent) and 6 (18 percent). Classification system 1 had been eliminated from consideration due to the high number of stream classes and classes with limited sample size, leaving classification systems 4 (Baseflow index (3 levels), stream size (4 levels)) and 6 (Stream size (4 levels), karst (2 levels)) as the two systems under consideration for application to the Potomac River.

The models with high R^2 values in stream classifications 4 and 6 were not evenly distributed across stream classes, but rather concentrated in the class of medium sized rivers with low to moderate baseflow contribution or karst presence (Tables 4, 5). For the baseflow, stream size classification system (system 4), models of the relationships between Withdrawal Index and taxa richness, HBI, Ephemeroptera count, percent Dominant taxon, EPT, and Shannon Wiener diversity index all indicated high explanatory power (R^2 range: 0.36 – 0.68) and highly significant p-values for Withdrawal Index ($p<0.01$) in the medium size, moderate baseflow stream class. For the stream size, karst classification system (system 6), models of the relationships between Withdrawal Index and the same biological variables (taxa richness, HBI, Ephemeroptera count, percent Dominant taxon, EPT, and Shannon Wiener diversity index) also indicated high explanatory power (R^2 range: 0.35 – 0.75) and highly significant p-values for Withdrawal Index ($p<0.01$) in the medium size, low karst stream class. Collapsing stream size into two groups (<38.6 mi^2 and >38.6 mi^2) and creating a stream classification system of two size levels and two karst levels (system 7; 4 classes total) did not improve or maintain the explanatory ability of the models compared with system 6 (Table 6).

Table 4. Modeling results for the BSFLCL_3_S (baseflow, 3 classes and stream size, 4 classes) stream classification system. Stream classes listed are only those that had n>15. See Table 1 for definitions of response variables. Models in bold had significant associations between hydrologic alteration and ecological response and high R^2 values (>0.30).

Stream class	**Response variable**	**Sample size**	**WI p-value**	**R^2**
Headwaters, moderate baseflow	Richness	108	0.98	0.2
	HBI	108	0.41	0.24
	Ephemeroptera count	108	<0.01	0.06
	Dominant taxon	108	0.14	0.08
	EPT	108	0.16	0.27
	Chironomidae count	108	<0.01	0.07
	Shannon Wiener	108	0.82	0.18
	Drift3 count	108	0.03	0.09
	Size1 count	108	<0.01	0.07
	Volt3 count	108	0.05	0.11
	Trop1 count	108	<0.01	0.23
	Ther1 count	108	0.04	0.14
	Rheo1 count	108	0.04	0.16
Small rivers, moderate baseflow	Richness	93	0.21	0.02
	HBI	93	0.22	0.19
	Ephemeroptera count	93	<0.01	0.06
	Dominant taxon	93	<0.01	0.01
	EPT	93	0.09	0.09
	Chironomidae count	93	<0.01	0.07
	Shannon Wiener	93	0.32	0.04
	Drift3 count	93	0.82	0.03
	Size1 count	93	<0.01	0.21
	Volt3 count	93	<0.01	0.16
	Trop1 count	93	0.81	0.25
	Ther1 count	93	<0.01	0.04
	Rheo1 count	93	<0.01	0.09
Medium rivers, moderate baseflow	**Richness**	**22**	**0.02**	**0.63**
	HBI	**22**	**0.01**	**0.38**
	Ephemeroptera count	**22**	**<0.01**	**0.36**
	Dominant taxon	**22**	**<0.01**	**0.44**
	EPT	**22**	**<0.01**	**0.54**
	Chironomidae count	22	<0.01	0.07
	Shannon Wiener	**22**	**<0.01**	**0.68**
	Drift3 count	22	0.54	0.07
	Size1 count	22	0.17	0.11
	Volt3 count	22	<0.01	0.05
	Trop1 count	22	0.74	0.17
	Ther1 count	22	0.33	0.30
	Rheo1 count	22	<0.01	0.09

Table 4. Continued

Stream class	Response variable	Sample size	WI p-value	R^2
Headwaters, high baseflow	Richness	23	0.46	0.23
	HBI	23	0.07	0.63
	Ephemeroptera count	23	0.64	0.05
	Dominant taxon	23	<0.01	0.03
	EPT	23	0.22	0.46
	Chironomidae count	**23**	**<0.01**	**0.39**
	Shannon Wiener	23	0.88	0.23
	Drift3 count	23	<0.01	0.06
	Size1 count	23	<0.01	0.07
	Volt3 count	23	<0.01	0.19
	Trop1 count	23	<0.01	0.24
	Ther1 count	23	<0.01	0.11
	Rheo1 count	23	0.48	0.05
Small rivers, high baseflow	Richness	32	0.15	0.16
	HBI	32	0.02	0.27
	Ephemeroptera count	32	<0.01	0.13
	Dominant taxon	32	<0.01	0.15
	EPT	32	0.31	0.18
	Chironomidae count	32	0.27	0.02
	Shannon Wiener	32	0.26	0.16
	Drift3 count	32	0.03	0.12
	Size1 count	32	<0.01	0.12
	Volt3 count	32	0.10	0.04
	Trop1 count	32	0.17	0.19
	Ther1 count	**32**	**<0.01**	**0.34**
	Rheo1 count	32	0.22	0.03

Table 5. Modeling results for the Size_4/Karst stream classification system. Stream classes listed are only those that had n>15. See Table 1 for definitions of response variables. Models in bold had significant associations between hydrologic alteration and ecological response and high R^2 values (>0.30).

Stream class	Response variable	Sample size	WI p-value	R^2
Headwaters, low karst	Richness	95	0.17	0.12
	HBI	95	0.04	0.14
	Ephemeroptera count	95	<0.01	0.07
	Dominant taxon	95	0.01	0.02
	EPT	95	0.99	0.19
	Chironomidae count	95	<0.01	0.09
	Shannon Wiener	95	0.18	0.05
	Drift3 count	95	<0.01	0.01
	Size1 count	95	<0.01	0.03
	Volt3 count	95	<0.01	0.11
	Trop1 count	95	<0.01	0.09
	Ther1 count	95	0.22	0.08
	Rheo1 count	95	0.13	0.31
Headwaters, high karst	Richness	36	0.51	0.03
	HBI	36	0.77	0.27
	Ephemeroptera count	36	<0.01	0.15
	Dominant taxon	36	<0.01	0.02
	EPT	36	0.53	0.15
	Chironomidae count	36	0.01	0.26
	Shannon Wiener	36	0.82	0.02
	Drift3 count	36	0.08	0.05
	Size1 count	36	0.23	0.20
	Volt3 count	36	<0.01	0.23
	Trop1 count	**36**	**<0.01**	**0.30**
	Ther1 count	36	0.2	<0.01
	Rheo1 count	36	0.44	0.64
Small rivers, low karst	Richness	79	0.25	0.02
	HBI	79	0.09	0.17
	Ephemeroptera count	79	<0.01	0.12
	Dominant taxon	79	<0.01	0.02
	EPT	79	0.44	0.1
	Chironomidae count	79	<0.01	0.07
	Shannon Wiener	79	0.22	0.05
	Drift3 count	79	<0.01	0.05
	Size1 count	79	<0.01	0.25
	Volt3 count	79	<0.01	0.15
	Trop1 count	79	<0.01	0.23
	Ther1 count	79	<0.01	0.03
	Rheo1 count	79	<0.01	0.09

Table 5. Continued.

Stream class	**Response variable**	**Sample size**	**WI p-value**	**R^2**
Small rivers, high karst	Richness	46	0.14	0.08
	HBI	46	0.18	0.01
	Ephemeroptera count	46	0.01	0.07
	Dominant taxon	46	0.03	0.03
	EPT	46	0.04	0.14
	Chironomidae count	46	<0.01	0.05
	Shannon Wiener	46	0.30	0.10
	Drift3 count	46	<0.01	0.09
	Size1 count	46	0.70	0.15
	Volt3 count	46	<0.01	0.12
	Trop1 count	46	<0.01	0.19
	Ther1 count	46	0.13	<0.01
	Rheo1 count	46	0.21	0.07
Medium rivers, low karst	**Richness**	**22**	**<0.01**	**0.71**
	HBI	**22**	**<0.01**	**0.58**
	Ephemeroptera count	**22**	**<0.01**	**0.35**
	Dominant taxon	**22**	**<0.01**	**0.58**
	EPT	**22**	**<0.01**	**0.60**
	Chironomidae count	22	<0.01	0.16
	Shannon Wiener	**22**	**<0.01**	**0.75**
	Drift3 count	22	<0.01	<0.01
	Size1 count	22	0.49	0.09
	Volt3 count	22	<0.01	0.22
	Trop1 count	22	<0.01	0.16
	Ther1 count	22	0.23	0.36
	Rheo1 count	22	0.47	0.07

Table 6. Modeling results for the Size_2/Karst stream classification system. See Table 1 for definitions of response variables. No models had significant associations between hydrologic alteration and ecological response and high R^2 values (>0.30).

Stream class	Response variable	Sample size	WI p-value	R^2
Headwaters, low karst	Richness	95	0.17	0.12
	HBI	95	0.04	0.14
	Ephemeroptera count	95	<0.01	0.07
	Dominant taxon	95	0.01	0.02
	EPT	95	0.99	0.19
	Chironomidae count	95	<0.01	0.09
	Shannon Wiener	95	0.18	0.05
	Drift3 count	95	<0.01	0.01
	Size1 count	95	<0.01	0.03
	Volt3 count	95	<0.01	0.11
	Trop1 count	95	<0.01	0.09
	Ther1 count	95	0.22	0.08
	Rheo1 count	95	0.13	0.31
Headwaters, high karst	Richness	36	0.51	0.03
	HBI	36	0.77	0.27
	Ephemeroptera count	36	<0.01	0.15
	Dominant taxon	36	<0.01	0.02
	EPT	36	0.53	0.15
	Chironomidae count	36	0.01	0.26
	Shannon Wiener	36	0.82	0.02
	Drift3 count	36	0.08	0.05
	Size1 count	36	0.23	0.20
	Volt3 count	36	<0.01	0.23
	Trop1 count	**36**	**<0.01**	**0.30**
	Ther1 count	36	0.2	<0.01
	Rheo1 count	36	0.44	0.64
Small/medium rivers, low karst	Richness	101	0.02	0.09
	HBI	101	<0.01	0.18
	Ephemeroptera count	101	<0.01	0.13
	Dominant taxon	101	<0.01	0.07
	EPT	101	0.02	0.11
	Chironomidae count	101	<0.01	0.05
	Shannon Wiener	101	<0.01	0.14
	Drift3 count	101	<0.01	0.03
	Size1 count	101	<0.01	0.18
	Volt3 count	101	<0.01	0.13
	Trop1 count	101	<0.01	0.18
	Ther1 count	101	<0.01	0.05
	Rheo1 count	101	<0.01	0.07

Table 6. Continued.

Stream class	Response variable	Sample size	WI p-value	R^2
Small/medium rivers, high karst	Richness	55	0.02	0.12
	HBI	55	0.03	0.17
	Ephemeroptera count	55	<0.01	0.06
	Dominant taxon	55	<0.01	0.06
	EPT	55	<0.01	0.19
	Chironomidae count	55	<0.01	0.08
	Shannon Wiener	55	0.30	0.12
	Drift3 count	55	<0.01	0.10
	Size1 count	55	0.34	0.16
	Volt3 count	55	<0.01	0.13
	Trop1 count	55	<0.01	0.18
	Ther1 count	55	<0.01	<0.01
	Rheo1 count	55	0.32	0.05

Conclusions from Susquehanna River analyses

These analyses suggest that a stream classification system can reduce the natural hydrologic and/or ecological variability within each resulting stream class and improve the explanatory ability of models of hydrologic alteration and ecological response. Of the variables included in the stream classification systems tested, those that represent flow volume (stream size) and groundwater input (baseflow index and karst) were best at accounting for natural variability among classes, measured by increased R^2 values of models relating hydrologic alteration and ecological response. Baseflow index and presence of karst seem to describe similar stream processes, and the analyses suggest that either variable may be used to represent groundwater input to streams. One other classification system, the system that included stream size, gradient, geology, and temperature, also had a high proportion of models with high R^2 compared with the other classification systems we examined. However, this stream classification system resulted in 30 stream classes for the Susquehanna basin and only seven of those stream classes had a sample size sufficient for statistical analyses (n>15). We had decided *a priori* that we would give priority to stream classification systems that yielded less than ten stream classes for the Susquehanna and Potomac basins, unless the stream classes were distributed so that most of the classes had a sufficient sample size for analyses (n>15 for the Susquehanna; n>30 for the Potomac).

Although the stream classification systems that included stream size and groundwater input improved the R^2 values for some of the models relating hydrologic alteration to ecological response, this result was only observed for one stream class in each classification system: medium rivers with moderate baseflow (in the stream size, baseflow index classification) and medium rivers with low karst (in the stream size, karst classification). These classification systems did not improve the R^2 values of models for classes with smaller stream sizes (headwaters or small rivers) or high groundwater input (high baseflow index or high percent karst). We did not have the data to examine models in the large stream size classes (with any degree of groundwater input) or medium rivers with low baseflow index.

There are several possible explanations why we may have seen stronger relationships between hydrologic alteration and ecological response for larger (medium-sized) rivers. First, although the same macroinvertebrate sampling methods were used at all sites, it is likely that entire riffles were sampled in small streams due to the ease of wading across the channel. In larger rivers, it is possible that samples were concentrated at stream margins rather than the middle of the channel, sites that would be more susceptible to alterations in hydrology. Second, most of the larger river sites had larger values for Withdrawal Index and also likely had poor water quality compared with smaller streams; thus, multiple stressors at larger river sites may have resulted in poorer ecological condition. Finally, water withdrawals in small streams are often from groundwater pumping or a mixture of groundwater and surface water withdrawals, whereas water withdrawals in large rivers are almost exclusively from surface water. Thus, macroinvertebrates may respond more directly to surface water withdrawals, which represent a greater proportion of total flow at larger river sites. As for why we observed stronger relationships between flow alteration and ecology for medium-sized rivers with low to moderate groundwater input compared with the same size rivers with high groundwater input, it is possible that groundwater input mitigates some of the effects of water withdrawals in these rivers, providing more water and cooler temperatures during critical low-flow periods and decreasing the ecological effects of the withdrawal.

Overall, our recommendation is to examine the application of the stream classification system based on stream size (4 levels) and groundwater input (either using baseflow index or karst) to the Potomac basin, by examining the total number of stream classes and the sample size of sites with both hydrologic data and biological data in each class. Using karst may be preferable to baseflow index as the metric related to groundwater input in the final classification system. The team creating NEAHCS found that baseflow index was a better predictor of stream temperature compared with geology data (7 bedrock and 2 surficial local and cumulative statistics) in CART analyses (Arlene Olivero, personal communication). However, baseflow index data is a coarse-scale 1 km grid dataset that was created by interpolating baseflow index values estimated at USGS stream gages, not correcting for any existing hydrologic alterations at those gages (Olivero & Anderson 2008).

Finally, although we found that application of a stream classification system improved some of our modeled relationships between hydrologic alteration and ecological response, we expect to use improved input data in our analyses for the Potomac River which will hopefully lead to less variability inherent in the data and stronger statistical relationships. The Water Analysis Screening Tool (WAST) data that we used for the Susquehanna analyses are based on permitted water withdrawals as a proportion of the estimated 7Q10 statistic for specific pour points. Data from WAST are excellent for calculating water withdrawal indices consistently for many sites basin- or state-wide. However, WAST does not account for hydrologic alterations associated with reservoir operations. Second, because all upstream withdrawals (and discharges) are aggregated for each pour point, potential local impact of particularly large withdrawals (or discharges) may be missed. Third, water withdrawals are based on permitted withdrawals (or discharges) as well as estimated water use, and may not accurately reflect actual use at a pour point. Finally, WAST does not give seasonal estimates of withdrawals, and larger withdrawals during typical low flow periods would likely have stronger ecological effects than withdrawals at other times of the year. The hydrologic data for the Potomac River will incorporate permitted water withdrawals as well as reservoir operations and land use change in modeling of unregulated and regulated flow data. In addition, we propose exploring additional statistical techniques (beyond linear and logistic regression) to model the relationships between hydrologic alterations and ecological response for the Potomac basin.

Literature Cited

Apse, C., DePhilip, M., Zimmerman, J., and Smith, M. P. 2008. Developing instream flow criteria to support ecologically sustainable water resource planning and management. Final report to the Pennsylvania Instream Flow Technical Advisory Committee. The Nature Conservancy, Harrisburg, PA.

Freeman, M. C., and P. A. Marcinek. 2006. Fish assemblage responses to water withdrawals and water supply reservoirs in Piedmont streams. Environmental Management **38**:435-450.

Olivero, A. P. and Anderson, M. G. 2008. Northeast aquatic habitat classification system. The Nature Conservancy, in collaboration with the Northeast Association of Fish and Wildlife Agencies, Boston, MA.

Poff, N. L., J. D. Olden, N. K. M. Vieira, D. S. Finn, M. P. Simmons, and B. C. Kondratieff. 2006. Functional trait niches of North American lotic insects: traits-based ecological applications in light of phylogenetic relationships. Journal of the North American Benthological Society **25**:730-755.

SAS Institute Inc. 2009. SAS OnlineDoc®, version 9.2. SAS Institute Inc., Cary, NC.

Stuckey, M. H. 2006. Low-flow, base-flow, and mean-flow regression equations for Pennsylvania Streams. U.S. Geological Survey, Scientific Investigations Report 2006-5130.