OFFICE OF WATER PROGRAMS

BUREAU OF CLEAN WATER

TECHNICAL DEVELOPMENT OF A THERMAL FISH INDEX

Prepared by:

Tim Wertz
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2021

ACKNOWLEDGMENTS

This project represents a culmination of extensive efforts directed towards fish-based developments over the past two decades. This project would not have been possible without the field support, expert recommendations, and technical review of biologists currently and/or formerly employed with the Pennsylvania Department of Environmental Protection and affiliated organizations. The following people have made extensive contributions to various milestones of this project through time:

Late 1990's - 2014

Rod Kime, **Andy Klinger**, **Rick Spear** and **Bill Botts** (DEP) compiled species attributes, standardized data collection efforts and conducted preliminary developmental evaluations that greatly influenced this project.

2015-2016

Frank Borsuk with the United States Environmental Protection Agency (EPA, Region 3), **Aaron Henning** with the Susquehanna River Basin Commission (SRBC), **Ben Lorson** and **Geoff Smith** with the Pennsylvania Fish and Boat Commission (PFBC) submitted reviews and comments of thermal preference rankings.

2015-2019

Lou Reynolds and **Greg Pond** (EPA, Region 3), **Matt Shank** (formerly with Susquehanna River Basin Commission, SRBC) **Ty Wagner** with the Pennsylvania Fish and Wildlife Cooperative Unit - United States Geological Survey (USGS) at The Pennsylvania State University (PSU), **Megan Schall**, **Katherine Restori** (PSU) and **Jennifer Rapp** (USGS) were pivotal in the stressor gradient developments that contributed to this project.

2008-2019

Data collection, coordination and developmental conversation was ongoing throughout this project and could not have been completed without collaborative efforts across agencies. Major contributions were made by representatives from:

PFBC – **Doug Fischer**, **Geoff Smith**, **Steve Kepler** and **Tom Shervinskie**.

SRBC – **Aaron Henning**, **Luanne Steffy** and **Matt Shank**

PSU – **Jay Stauffer**, **Megan Schall** and **Ty Wagner**

EPA – **Greg Pond**, **Frank Borsuk** and **Lou Reynolds**

Throughout this period, a tremendous amount of effort by DEP staff went into data collection, analytical support and technical guidance that made this project possible. While not specifically listed below, a significant amount of field support and data entry was conducted by numerous interns whose contributions are greatly appreciated.

Southeast Regional Office – Alan Everett, Ben Russell, Ed Filip and **Steve Unger**

Southcentral Regional Office – Andy Blascovich, Jeremy Miller, and **Kristen Bardell**

Southwest Regional Office – Dan Counahan, Gary Kenderes, Jamie Detweiler and **Rick Spear**

Northwest Regional Office – Jay Gerber, and **Joe Brancato**

Great Lakes Office – Jim Grazio

Northcentral Regional Office – Dave Rebuck, Jared Dressler and **Steve Means**

Northeast Regional Office – JR Holtsmaster, Tim Daley

Central Office – Aaron Frey, Amy Williams, Andy Klinger, Barb Lathrop, Bill Botts, Bill Brown, Brian Chalfant, Carrie Tropasso, Charlie McGarrell, Dan Bogar, Dustin Shull, Erika Bendick, Felicia Lamphere, Gary Gocek, Gary Walters, Heidi Biggs, Jeff Butt, Josh Lookenbill, Justin Lorson, Kristen Schlauderaff, Mark Brickner, Mark Hoger, Matt Kundrat, Matt Shank, Megan Bradburn, Michael Morris, Molly Pulket, Rebecca Dunlap, Rebecca Whiteash, Rick Shertzer, Rob Ryder, Rod Kime, Rod McAllister, Scott Heidel, Shawn Miller, Tom Barron, Tony Shaw and **Travis Stoe.**

PROJECT SUMMARY

This project is the result of a long history of developmental efforts spanning nearly three decades. Fish community-based assessments gained nationwide momentum in the 1980s. By the 1990s, numerous states began developmental efforts to include fish community-based monitoring and assessments. Federal guidance on developmental

strategies also began to emerge around this time. Pennsylvania initiated the first fish community-based developmental efforts in the mid-1990s when PFBC entered into a contractual agreement with DEP that funded a staff position dedicated to the task. Data collection was conducted across Pennsylvania's wadeable streams through the contract period. Once fish community data was collected and evaluated, DEP pursued assessment method development. Based on spatial coverage and sample sizes, the assessment method development effort was largely focused on regional and stream size calibrations that were specifically focused on warmwater, wadeable streams in the Ohio River basin. Throughout this time, standardized sampling protocols were developed and evaluated to achieve consistent and repeatable data collection while maintaining a reasonable level of logistical effort. Standardized sampling protocols were developed in the mid-2000s for wadeable streams. Based on the revised sampling protocols, a probabilistic sampling effort was conducted across wadeable streams throughout the state. This monumental project ("Fish IBI Project") included two years (2008-2009) of intensive data collection that was collaborative between DEP, PFBC and PSU. As data and sample sizes began to increase, developmental efforts began focusing on the eastern half of the state (Atlantic Slope). Throughout all developmental efforts, various challenges were identified that were generally related to extensive zoogeographical influences on fish distributions, sample sizes and barriers to recolonization. The challenges identified in early developmental efforts were invaluable to current efforts. As these challenges were being considered, concerns over large river system health began to emerge. The development of standardized sampling protocols (non-wadeable) followed and extensive data collection continued across both wadeable and non-wadeable waterbodies. By 2016, data was available to initiate another round of developmental effort that was able to focus on a larger spatial scale across both wadeable and non-wadeable waterbodies. The following document is the consolidated product of all these efforts.

INTRODUCTION

This document is intended to describe the purpose, applicability and development of a thermal fish index (TFI) that serves as a multidisciplinary tool for management centered around fish and their role in 25 Pa. Code § 93 for measuring water quality. Specifically, “Uses” are discussed to establish initial context for making assessments and evaluations, pursuant to Water Quality Standards, using fish. The technical development of a TFI focuses on introducing important fish-related concepts and empirical evaluations that support these concepts. General results are discussed to create an initial foundation for assessment and evaluations. Specific assessment methods stemming from this technical document are presented independently (Wertz 2021).

25 Pa. Code § 93.3 lists five categories of protected water uses: Aquatic Life, Water Supply, Recreation and Fish Consumption, Special Protection and Other. Each of these categories is specifically intended to protect and support the resource and/or the user of the resource. Section 93.3 defines four Aquatic Life Use (ALU) sub-categories, three of which – Cold Water Fishes (CWF), Warm Water Fishes (WWF) and Trout Stocking (TSF) – are narrative definitions of biological communities along a thermal regime (Table 1). These protected sub-categorical ALU definitions describe maintenance and/or propagation of certain ecological communities. Pennsylvania’s water quality standards also include water quality criteria associated with various protected uses. In this context of protected uses and water quality criteria, development of biologically based ALU assessments must be calibrated and responsive to changes in water quality and habitat as a reflection of waterbody condition and the ability of a waterbody to support relevant protected uses. In contrast to protected ALUs and associated water quality criteria, which focus on the ability of a waterbody to support certain fish and associated flora and fauna, protected recreational uses focus on human uses of waterbodies (Table 1) with associated water quality criteria designed to be protective of human health.

25 Pa. Code § 93.7 provides maximum temperature criteria for defined times of the year for three protected sub-categorical ALUs: CWF, WWF and TSF. Temperature criteria in § 93.7 are applied to heated waste sources regulated under 25 Pa. Code Chapters 92a and 96. Temperature limits apply to other sources when they are needed to protect designated and existing uses. In other words, temperature criteria are applied to specific cases and are not used for broad assessments of ALU. As indicated from the ALU definitions, an appropriate thermal evaluation includes a biological assessment (bioassessment) based on instream flora and fauna, with specific mention of fish species.

Table 1. Protected water uses for Aquatic Life and for Recreation and Fish Consumption (25 Pa. Code § 93.3)

Aquatic Life	
CWF - Cold Water Fishes	Maintenance or propagation, or both, of fish species including the family Salmonidae and additional flora and fauna which are indigenous to a cold water habitat.
WWF - Warm Water Fishes	Maintenance and propagation of fish species and additional flora and fauna which are indigenous to a warm water habitat.
MF - Migratory Fishes	Passage, maintenance and propagation of anadromous and catadromous fishes and other fishes which move to or from flowing waters to complete their life cycle in other waters.
TSF - Trout Stocking	Maintenance of stocked trout from February 15 to July 31 and maintenance and propagation of fish species and additional flora and fauna which are indigenous to a warm water habitat.
Recreation and Fish Consumption	
B – Boating	Use of the water for power boating, sail boating, canoeing and rowing for recreational purposes when surface water flow or impoundment conditions allow.
F – Fishing	Use of the water for the legal taking of fish. For recreation or consumption.
WC – Water Contact	Use of the water for swimming and related activities.
E – Esthetics	Use of the water as an esthetic setting to recreational pursuits.

Freshwater fishes are important indicators of temperature as they are obligate poikilothermic, meaning their internal body temperatures are dictated by the ambient surrounding water temperature (Wood and McDonald 1997, Beitinger et al. 2000). Thermal preference and tolerance vary among species (Wehrly et al. 2003, Yoder 2006), creating unique assemblages of fishes along a continuous gradient, upstream to downstream. These longitudinal changes in fish assemblages parallel important shifts in loading, transport and utilization of organic matter from headwaters to mouth that form the river continuum concept (RCC; Vannote et al. 1980). The thermal zonation of fishes along a longitudinal gradient has been realized for nearly a century (Carpenter and

Huxley 1928) and biological zones have been identified based on the occurrence of dominant fishes as “indicator species” (Huet 1959).

The use of indicator species may be appropriate where only presence-absence data are available, but the use of indicator species in bioassessments tends to lack responsiveness to degradation along a continuous gradient (Fausch et al. 1990). For example, when an indicator species is absent due to stress, any additional stress on the system will have no measurable effect. An alternative to using indicator species is the use of all species in an assemblage and their relative abundance based on taxonomy, traits and tolerance values to make bioassessments along a broad range of stress. The shift from indicator species to a more broad-scale, assemblage-based approach largely began in the 1970s and 1980s, and the application of these concepts were first realized by the seminal introduction of the Index of Biotic Integrity (IBI) conceptualized by Karr (1981). As assemblage-based concepts began to evolve from indicator species concepts, regulatory definitions evolved as well. Historic ALU definitions were largely dependent on using trout species (family Salmonidae) as an indicator of a cold water community, as evident from the evolution of definitions from the late 1960s (

Table ***2***) to what they are today (Table 1).

It is important to note that sub-categorical ALU definitions that make specific mention of trout (e.g., CWF, TSF), use trout as an indicator of natural thermal communities. Where trout populations are being completely supported indicates that additional flora and fauna indigenous to a cold water habitat may be supported (CWF), or not supported (WWF) by waterbody conditions, and an ecological community intermediate of CWF and WWF exists that does not fit the sub-categorical definition of TSF. While trout fishing has well-established socioeconomic value, the socioeconomic value of trout fishing is included and protected in Pennsylvania's water quality standards under the protected recreational fishing use, not under the protected ALU (Table 1).

Table 2. Historic sub-categorical aquatic life uses from Article 301 of the Sanitary Water Board Rules and Regulations, Commonwealth of Pennsylvania, Water Quality Criteria (1968).

Cold Water Fishes	Maintenance or propagation, or both, of fish species of the family Salmonidae and fish food organisms.
Warm Water Fishes	Maintenance and propagation of fish food organisms and all families of fishes except Salmonidae.
Migratory Fishes	Passage, maintenance and propagation of anadromous and catadromous fishes and other fishes which ascend to flowing waters to complete their life cycle.
Trout (Stocking Only) *	Warm water fishes and trout stocking

* Added December 20, 1967.

On the surface, sub-categorical ALU definitions could be interpreted as framework for thermal assessment with CWF, TSF and WWF considered as hierarchical along a thermal gradient. However, these ALU definitions have inherent complexities that present challenges for assessment purposes. Specifically, some waters under natural (or near-natural) conditions may not always support CWF (e.g., large streams, rivers). Additionally, the interpretation of the ALU definitions have traditionally relied heavily on the presence of fish (e.g., trout in CWF) to fulfil the “maintenance” requirements, and the presence of young-of-year or multiple age-classes of fish (e.g., trout in CWF) to fulfil the “propagation” requirements. This interpretation can be successfully applied to CWF when trout are present in high numbers but becomes more challenging as trout numbers are reduced (e.g., how many trout constitute “maintenance”?). In other words, numerical thresholds may help alleviate subjectivity while providing consistent interpretation of narrative definitions. Furthermore, species within the trout family (Salmonidae), need to clearly demonstrate a positive response to good water quality, if their use as a potential indicator of waterbody condition is to be meaningful. Preliminary investigations of trout density and abundance as an indicator of water quality suggests responses to water quality may be variable overall and species-specific (Figure 1).

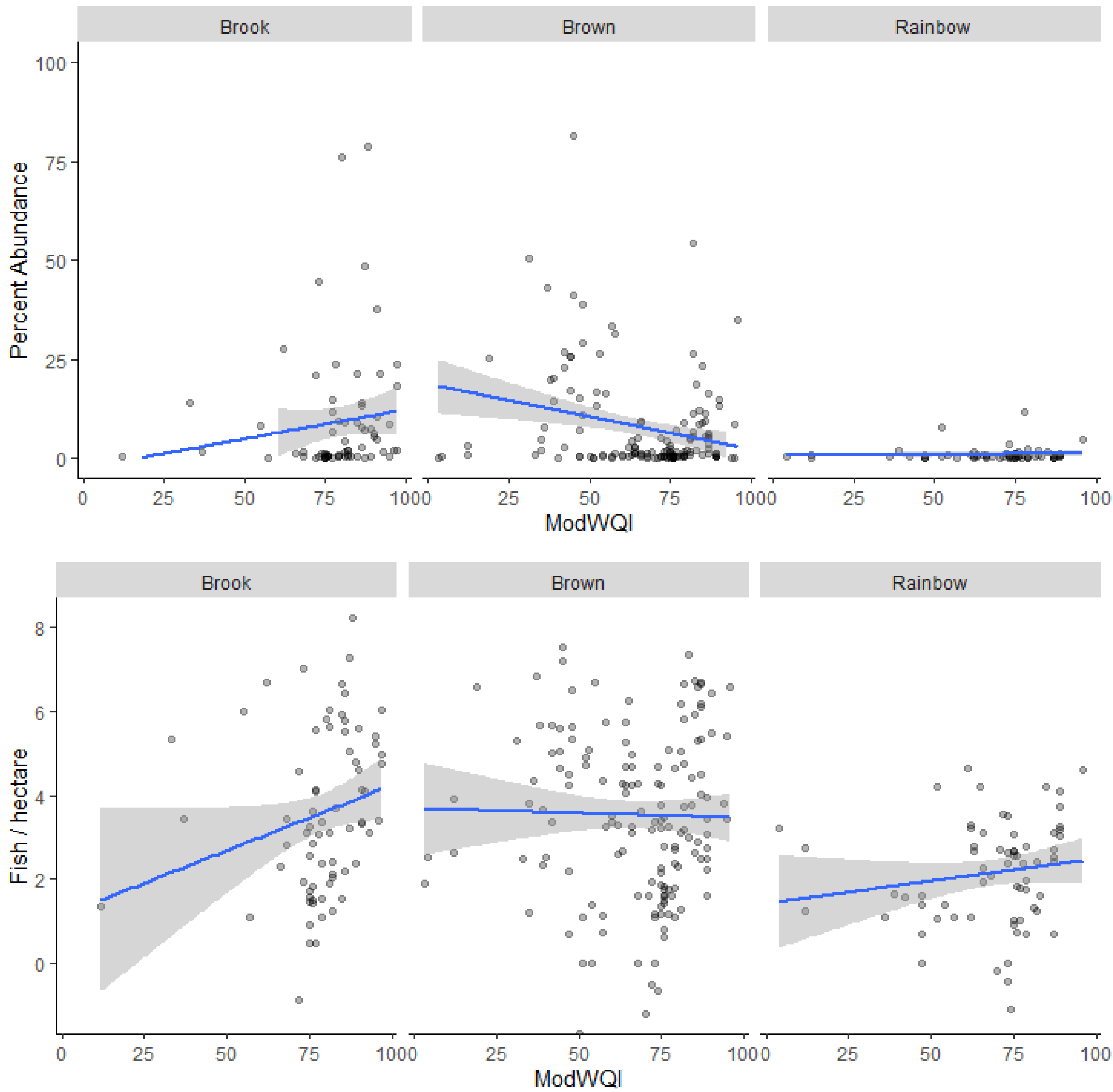

Figure 1. Three common trout species found in Pennsylvania, density (log transformed) as the number of fish per hectare and relative abundance in response to water quality, measured with the modified water quality index (ModWQI), at all sites where present. The ModWQI measures water quality stress from poor to good, on a continuous scale from 0-100, respectively.

Interpretation of ALU narrative provisions for “maintenance” and/or “propagation” can be bolstered through the development of numeric thermal assemblage classes. This provides a shift from qualitative implementation to a more quantitative description of fish assemblages along a thermal gradient. Numerically describing the transition from “fish species... and additional flora and fauna which are indigenous to a cold water habitat” to “fish species and additional flora and fauna which are indigenous to a warm water habitat” is possible by quantifying the transition of assemblages dominated by cold water species to assemblages dominated by warm water species. This transition is considered continuous in nature as opposed to binary, meaning there will be assemblages dominated by warm water species that may still have cold water species present. This transitional or “cool water” assemblage appears to align with the TSF use interpretation presented, where stocked trout may be seasonally maintained within a warm water assemblage. However, important differences are noted between the TSF use and a transitional assemblage that may preclude quantification of TSF directly. Herein, a transitional assemblage is considered a segue of biological assemblages intermediate of cold and warm assemblages, based on environmental changes along a waterbody’s continuum (e.g., from headwaters downstream, slope, temperature). Subsequently, bioassessments should be directed towards quantifying a transitional assemblage (as opposed to TSF), as a measure of a waterbody’s ability to support this natural transitional assemblage. To avoid confusion between established uses in Chapter 93 and natural assemblages, the terms cold water assemblage (CWA), transitional assemblage (TSA) and warm water assemblage (WWA) will be used hereafter to describe thermal assemblages from an ALU assessment perspective. These terms are used to describe assemblage classes of fishes along a thermal (and environmental condition) gradient, and should in no way be considered redefinitions of ALUs.

As previously stated, ALU bioassessment tools are designed to evaluate a waterbody’s condition by measuring changes in biological assemblages, in response to stress. The natural thermal zonation of fishes along a longitudinal gradient can be altered by anthropogenic stressors (Caissie 2006, Stanfield and Kilgour 2013) that include but are not limited to: deforestation (Brown and Krygier 1970, Jones et. al 1999, Burcher et. al. 2008), urbanization (Brown et. al. 2005, Nelson and Palmer 2007), groundwater manipulation (Poole and Berman 2001, O’Driscoll and DeWalle 2006), impounding (Ward and Stanford 1983, Lessard and Hayes 2002), thermal effluents (Coutant 1975, Shuter et al. 1980) and global climate change (Eaton and Scheller 1996, Mohseni et al. 2003, Nelson and Palmer 2007). Thermal regimes can also be affected by natural factors that may combine to shape the fish assemblages found within a waterbody. Common natural effects that influence the thermal regime include effects related to latitude, elevation, slope, velocity, groundwater and canopy cover, among others. Less common effects may include effects associated with turbidity, basin orientation or substrate characteristics. Effects of anthropogenic and/or natural factors are usually spatiotemporally stochastic and are both responsible (at varying degrees), for the formation of modern-day fish assemblages. This theory forms a physical habitat template and suggests that recovery from disturbances and the response of fish assemblages may vary accordingly (Southwood 1977, Poff and Ward 1990). In other words, as anthropogenic stress increases in a waterbody, the natural thermal assemblage may adjust accordingly (Figure 2). Alternatively, as anthropogenic stress is mitigated (naturally or through management), the thermal assemblage may adjust accordingly. Therefore, it is important to note that the thermal response of fish assemblages is not exclusively limited to changes in temperature.

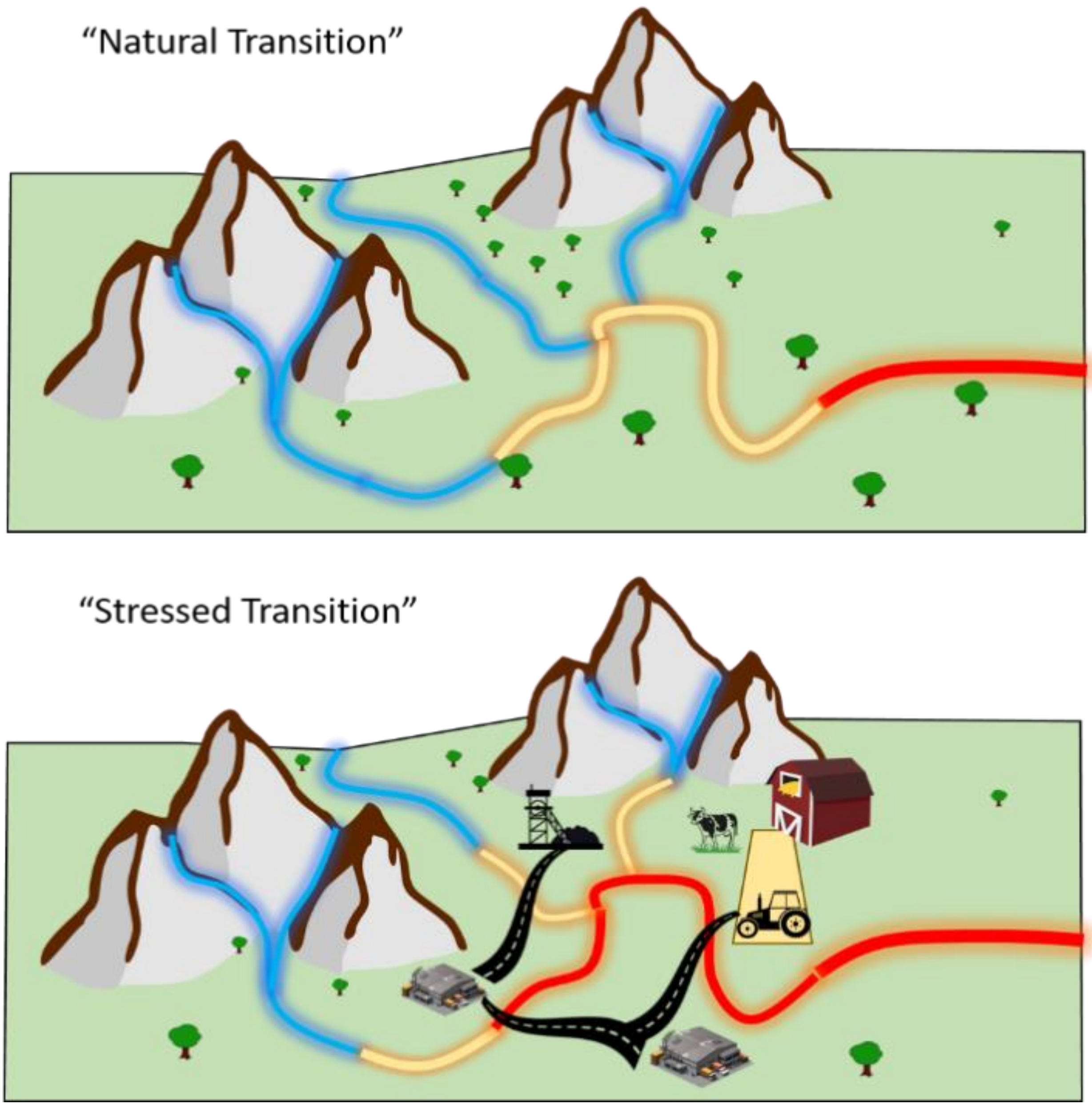

Figure 2. Theoretical example of natural longitudinal transition areas versus stress induced fish assemblage transitions. With applied stress to a cold water assemblage (CWA; blue), the CWA reduces, the transitional assemblage (TSA; yellow) is shifted upstream and the warm water assemblage (WWA; red) is expanded.

Historically, DEP has relied on bioassessments based on macroinvertebrate assemblages to make categorical ALU assessment determinations. Fish-based bioassessment tools offer a suite of benefits that compliment macroinvertebrate-based assessments. Benefits include: (1) fish life cycles are longer than macroinvertebrates, providing insight into acute and chronic exposure through time, (2) fish often respond to stress at different landscape scales than macroinvertebrates (Lammert and Allan 1999), (3) fish life history and tolerance information is widely available, (4) fish are relatively easy to identify, (5) fish have well-established socioeconomic value and (6) fish provide important evidence of sub-categorical ALU narrative descriptions. Fish-based bioassessment tools are typically considered more complex than macroinvertebrate-based tools, in that: (1) fish can be highly mobile within dendritic freshwater systems, (2) species distributions are based on zoogeographic factors that can make inter-basin comparisons challenging and (3) barriers (physical and/or chemical) to recolonization efforts may delay recovery. Pennsylvania has had extensive zoogeographic influences that have shaped six major drainage basins and nearly 200 fish species, represented by 28 families, have been recorded from Pennsylvania's non-tidal waters (Stauffer et al. 2016). To overcome distributional challenges and have a fish-based bioassessment tool that can be broadly applied throughout Pennsylvania, focus should be shifted away from taxonomic assessments (e.g., species-level assessments) and directed towards tolerance/preferences at the assemblage level. Additionally, focus should be directed towards relative abundance changes as a way to help mitigate potential effects of barriers to recolonization. For example, if a species is excluded from an area due to recolonization barriers, the species that are present may still respond to improving waterbody conditions that facilitate reproductive success and relative abundance. Specifically, this assessment method development is directed towards thermal tolerances (or preferences) of fish assemblages to make categorical ALU assessments, while numerically aligning assemblages with the intent of sub-categorical ALU definitions, to the extent possible.

Although there is a great deal of literature concerning the thermal response of fishes, there is little information regarding the quantification of entire fish assemblages along a thermal gradient (but see Zorn et al. 2002). The following represents the introduction of a metric, the TFI, that quantifies the thermal preference of entire assemblages as a numerical description of how "cold" or "warm" a fish assemblage is, based on a unitless scale. The TFI ranks assemblages from coldest to warmest along a 2.0 to 10.0 scoring gradient, respectively.

METHODS

Index Calculation

Fish species were designated within a thermal class as determined from thermal studies compiled by Eaton and Scheller (1996) and Lyons et al. (2009). Eaton and Scheller (1996) ranked each species on a three-tiered classification of Cold, Cool and Warm from streams across the continental United States, whereas Lyons et al. (2009) utilized an additional fourth tier by splitting Cool into Cool-transitional and Warm-transitional in Wisconsin and Michigan streams. Tiered delineations were converted to five-tiers with associated numerical values: 1-Cold (Cd), 2-Cold-Cool (CdCl), 3-Cool (Cl), 4-Cool-Warm (ClWm) and 5-Warm (Wm), similar to Coker et al. (2001), to normalize any disagreement between delineations. The list of Pennsylvania fish taxa and their thermal delineations were then independently sent to regional experts familiar with fishes of the Northeastern and Mid-Atlantic United States – including representatives of PFBC, the Susquehanna River Basin Commission (SRBC), EPA Region 3 and DEP – to delineate taxa not directly addressed by Eaton and Scheller (1996) and Lyons et al. (2009). Final delineations from regional experts were chosen based on modal values (Appendix A). Where modal values were not achieved, the delineations were made by using arithmetic mean rounded up or down by considering latitudinal distributions and habitat preferences for each species, similar to Coker et al. (2001).

To calculate the TFI, the number of individuals within each thermal class, as a proportion (e.g., 20% cold water individuals = 0.2), was calculated. A weighted average was obtained by multiplying the numeric value for the thermal class by the proportion of individuals, summed across classes. The final value is then multiplied by two to expand and standardize the range from two to ten, coldest to warmest, respectively (

Table ***3***). Calculation of the TFI follows:

$$TFI = \left(\sum_{1}^{5} NP_i\right) 2$$

where, N is the numeric value for the thermal class and P is the proportion of individuals at the i^{th} thermal class.

Table 3. Example of proportional abundance shifts of individuals within a thermal class, across the five thermal classes, and the resulting thermal fish index (TFI) score for six example assemblages.

Example Assemblage	Cold 1	Cold-Cool 2	Cool 3	Cool-Warm 4	Warm 5	TFI Score
Assemblage 1	1.00					2
Assemblage 2	0.60	0.30	0.10			3
Assemblage 3		0.60	0.30	0.10		5
Assemblage 4			0.60	0.30	0.10	7
Assemblage 5			0.10	0.30	0.60	9
Assemblage 6					1.00	10

Reference Condition and Stressor Gradient

A least-disturbed (LD) approach was used to develop a reference condition, or the “best available” condition, based on Stoddard et al. (2006). The criteria for establishing the LD condition was determined *a priori* and applied consistently across streams of all sizes to allow for assemblage characterization along a longitudinal gradient. Three major stress categories were identified: stressed (S), moderately stressed (M) and LD (Figure 3). Two major stressors on aquatic environments were used to delineate stress categories: water quality and habitat. Water quality stress was measured using a modified version of DEP’s water quality index (WQI), originally described by Wertz and Shank (2019). The original WQI used 21 parameters to inform stress condition along a land-use-similarity index (range = 0-100, S to LD, respectively). The modified WQI (modWQI) was reduced to 18 parameters (Table 4), which increased the number of fish sites available with paired water quality. Instream habitat measures were conducted following a modified version of the U.S. Environmental Protection Agency’s (EPA’s) Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers (RBP III) (Plafkin et al. 1989, Barbour et al.1999) associated with DEP and SRBC collection methods at each fish site (Shull and Lookenbill 2018, Shank et al. 2016). Since habitat measures varied across sampling methods (i.e., wadeable vs. nonwadeable), habitat measures were standardized into a habitat category score (Habcat; range = 1-4, LD to S, respectively) based on available measures of sedimentation, embeddedness, sand, silt and detritus. Finally, a dam proximity criterion was added to ensure fish sampling sites were not close to habitat-modified systems, or barriers to migration, that could potentially influence the fish assemblage (Table 5).

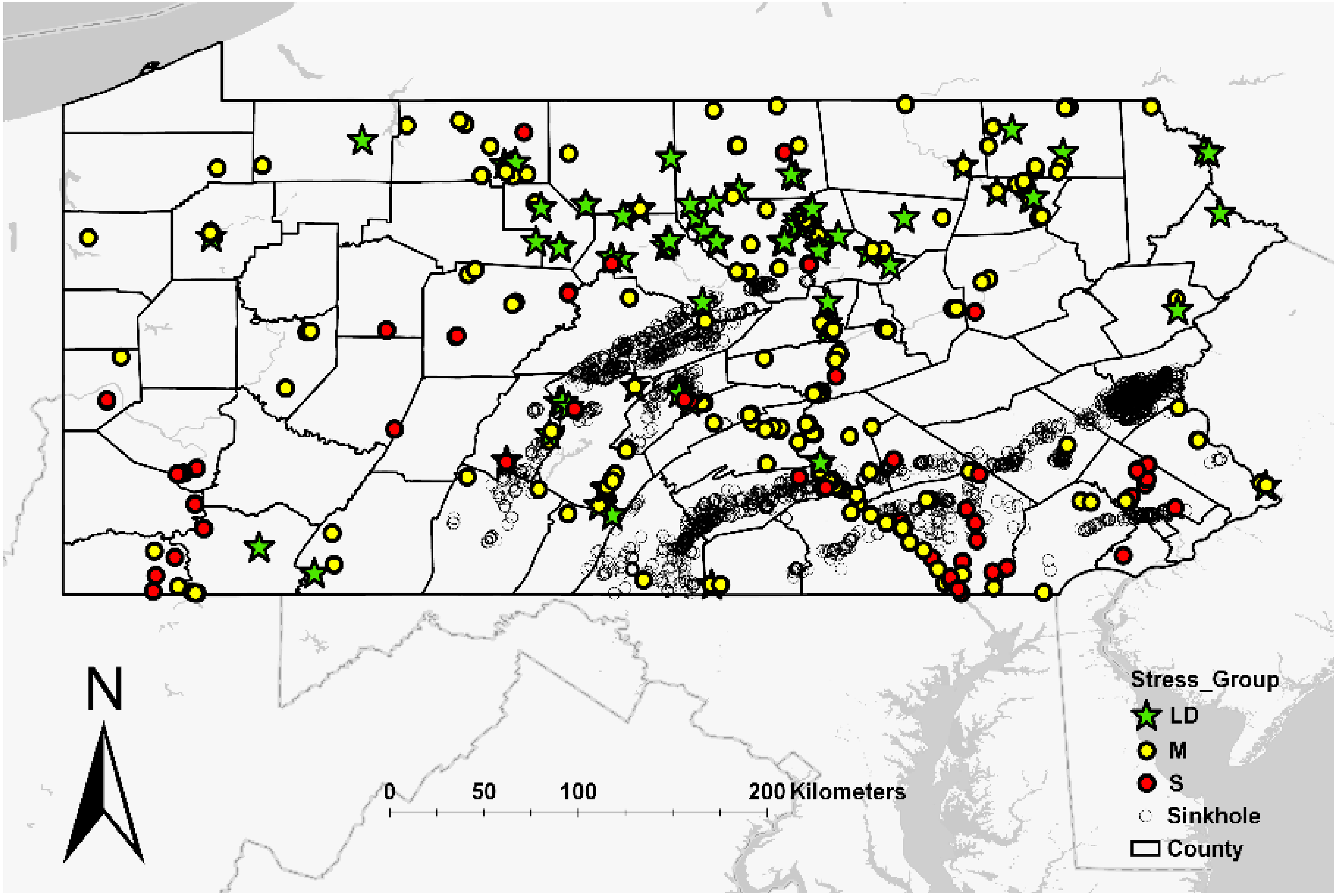

Figure 3. Map of sites considered least disturbed (LD), moderately stressed (M) and stressed (S) across a gradient of water quality and habitat conditions. Open circles represent sinkhole locations and relative density.

Table 4. Water quality parameters (n =18) used to create a modified water quality index (modWQI). See Wertz and Shank (2019) for further details.

PARAMETER
ALKALINITY, TOTAL
ALUMINUM, TOTAL
AMMONIA TOTAL AS NITROGEN
BROMIDE, TOTAL
CALCIUM, TOTAL
CHLORIDE, TOTAL
DISSOLVED SOLIDS, TOTAL
HARDNESS, TOTAL
IRON, TOTAL
MAGNESIUM, TOTAL
MANGANESE, TOTAL
NICKEL, TOTAL
pH
PHOSPHOROUS, TOTAL
SPECIFIC CONDUCTIVITY @ 25.0 C
SULFATE
SUSPENDED SOLIDS, TOTAL
ZINC, TOTAL

Table 5. Least-disturbed and stressed criteria based on water quality using a modified water quality index (modWQI), habitat from a categorical measure (Habcat) and dam proximity.

	Freestone		Limestone	
Description	Least Disturbed Criteria	Stressed Criteria	Least Disturbed Criteria	Stressed Criteria
modWQI score	> 60	< 40	> 40	< 20
Habcat score	1	3 or 4	1 or 2	3 or 4
Proximity to dam or impoundment	> 1.5 km		> 1.5 km	

Datasets

Prior to development of the TFI, water quality and habitat data were spatiotemporally paired with fish assemblage data; the full dataset was then inspected for potential outliers and anomalies. Three potential issues were considered at this stage and addressed accordingly. First, if an assemblage had >10% of the individuals not identified to the species-level (e.g., not having a thermal class), the TFI score was considered not representative of the assemblage and the sample was removed. Second, samples with less than 50 individuals were investigated for potential cause and representativeness. Potential causes for low sample sizes investigated were: 1) appropriate application of collection protocols (e.g., electrofishing settings, survey distance and time), 2) potentially toxic water quality conditions and 3) near-sterile conditions (e.g., extremely low productivity). Only two of these causes were identified in the dataset. Toxic conditions were observed at sites with extreme acid mine drainage and near-sterile conditions were observed in extreme headwaters. In both cases all samples were still considered to be representative of the overall site conditions and were retained. Lastly, sites were coded based on spatiotemporal representation of at least one water chemistry sample to fish sample location and time. Sites were coded from zero to three, best to worst expected representation, respectively, where: 1- represented same site and year, 2- represented a separation of year or distance but no evidence suggested a significant change in space or time and 3- represented a separation of year or distance with evidence to suggest the spatiotemporal gap may not be representative. All 3s (those fish samples where the spatiotemporally closest water chemistry sample was most likely to be unrepresentative) were removed from the dataset.

The full dataset was divided into three subsets; 1) precision dataset, 2) calibration dataset and 3) validation dataset. The precision dataset was first partitioned from the full dataset to reduce any pseudo-replication or spatial-autocorrelation issues (Sokal and Oden 1978, Hurlbert 1984). Since the stress measures used have a strong temporal component (i.e., based on instream measures instead of land use) sample independence was defined in an attempt to retain repeated samples from the same site that have measurable, spatiotemporal change in water quality or habitat. Independent samples were identified, randomly across samples at the same site, as having either a five-point modWQI score change or a one-point change in Habcat score. These respective temporal changes to water quality or habitat are biologically meaningful as they can occur as a result of anthropogenic activities, which would be expected to influence fish assemblages through time. Samples that didn’t meet this definition were regarded as repeat measures and were used in the precision dataset (samples removed; n = 193, 36%). The temporal strength of this method negated the need for temporal precision estimates as it treated all replicated sites, given similar water quality and habitat conditions, the same. In other words, repeated sites were considered standardized by habitat and water quality, to ensure variability was associated with natural conditions (i.e., seasonal, sampling, processing). This result was desired as sampling for quality assurance and precision estimates at fish community sites presents theoretical challenges. For example, sampling the exact same reach twice on the same day is not advised since fish were removed on the first sampling event. Over time the first sampled reach will go back pre-sampled conditions as fish are returned to the reach after processing and recolonizing occurs. The exact period of time for recolonization is unknown as it varies with reach conditions. Conversely, sampling a nearby reach (i.e., replicating) on the same day that is representative of the first is advisable but must be evaluated thoroughly to ensure habitat and water quality were similar to the first. Standardizing by habitat and water quality at replicated sites allows same-time replicates to be compared to across-time replicates.

The full development dataset (i.e., the dataset after splitting out the precision dataset from the full dataset) included sites from all stress groups (i.e., LD, M, S), and was randomly split into a calibration dataset and a validation dataset (80/20, respectively; *n* = 360/90). The LD sites within the calibration dataset were used for site classification purposes. The calibration dataset was used for development of the TFI and the validation dataset served as an independent basis for measuring the performance of the TFI.

Landscape Variables

Landscape variables were compiled at the local (stream segment) and watershed (total upstream catchment) scale. Local variables were obtained from the Appalachian Landscape Conservation Cooperative (AppLCC) stream classification system. The AppLCC contained data across six major variable types: size, gradient, temperature, hydrology, buffering capacity and confinement (Olivero et al. 2015). Of the six AppLCC variable groups available, temperature was the only variable group not used in development of the TFI as it was based on fish assemblage data and was considered redundant. Catchment data was obtained by delineating upstream drainage areas for each site and measuring area (km^2), density of sinkholes ($\#/km^2$) and the percent of limestone geology within the catchment. Drainage area was considered a longitudinal variable as catchment size increases from headwater to mouth. Sinkholes and limestone geology were specifically chosen to address potential relationships identified from previous studies relating to limestone and karst systems, and their effect on fish assemblages (Steffy and Kilham 2006, Carline et al. 2011, Kollaus and Bonner 2012). All landscape variables were compiled using ArcGIS Pro version 2.2 (ESRI 2018).

Site Classification

Boosted regression trees (BRTs) and recursive partitioning were used to classify LD and calibration datasets. Regression trees are a form of classification tree that utilize machine learning, where “boosting” generally improves performance from traditional regression trees by fitting multiple “simple” models with an error term to avoid overfitting. Boosted regression trees, used in a continuous regression situation, use recursive partitioning to split data into homogenous groups and sub-groups based on between-group sum-of-squares, similar to analysis of variance (ANOVA), (Qian 2016, Elith et al. 2008). Regression trees or recursive partitioning trees, (with or without boosting) have been utilized extensively for environmental modeling (Prasad et al. 2006, Breiman et al. 1984, Cutler et al. 2007, De'ath and Fabricius 2000, De’ath 2007) and groundwater studies (Trauth and Xanthopoulos 1997, Naghibi et al. 2016). All statistics were performed using R (R core team, 2016). Boosted regression trees were performed using R package ‘gbm’ (Ridgeway, 2006). Recursive partitioning trees were performed using R package ‘rpart’ (Therneau and Atkinson, 2019), method = ANOVA for continuous response variable. Least disturbed sites were first modeled using BRTs to determine appropriate classification groups from natural variables using a minimum of 1,000 trees and adjusting the learning rates (lr) and tree complexity (tc) following Elith et.al. (2008). The variable importance output from the BRTs were used to indicate which variables should be used for classification. Recursive partitioning trees (single tree) were then investigated to determine where to best split the important variables. Results were analyzed for ecological relevance and minimal cross-validation error (Qian 2016). Similarly, BRTs were investigated in the calibration dataset to explore potential effects that may be problematic for analysis. Potential problematic issues may arise from stressors or site classification groups not represented in the LD dataset. The results of the classification schema were then applied to the calibration dataset and adjusted as needed to obtain final classification groups based on ecologically relevant concepts (i.e., RCC). For example, if the result of the recursive partitioning trees indicated a classification group split was evident around a drainage area of 170 km^2, the area could be adjusted to 150 km^2 with the understanding that the extra 20 km^2 is ecologically arbitrary.

During preliminary data exploration investigations within the calibration dataset, using recursive partitioning and BRTs, the effect of karst geology became apparent (using sinkhole density within upstream catchment as an indicator metric). This finding was in concordance with previous studies conducted in watersheds dominated by limestone geology. Generally, limestone streams (i.e., streams in karst-dominated geology) have a unique ability to maintain cold water assemblages at increased stress levels, relative to their freestone counterparts (Steffy and Kilham 2006, Carline et al. 2011). This phenomenon was apparent in the calibration dataset, where sites with increased sinkhole densities were observed to have reduced modWQI scores while still maintaining lower TFI scores, when compared to the rest of the dataset. Sites with ≥ 0.03 sinkholes/km^2 were classified from sites with < 0.03 sinkholes/km^2, hereafter referred to as limestone (LS) and freestone (FS) stream types, respectively. It is important to note that no LS streams met LD criteria for water quality as LS streams typically are found in wide, fertile valleys that tend to be dominated by agricultural practices. Habitat quality was also reduced in the limestone group, as many of these streams are typically lower gradient with moderate sand and gravel substrates. Additionally, the effect of habitat quality on the TFI score was apparent within the LS group. Since the LS group of streams did not meet LD criteria for the FS group, the LD criteria for LS streams was adjusted from the LD criteria for the FS streams (**Error! Reference source not found.**5) after investigating the range of water quality across the sites in the LS streams. It is important to note that LS stress criteria were only adjusted due to lower (colder) TFI scores than FS streams under similar stress conditions. This is hereafter referred to as the “karst effect”. In this dataset, the karst effect began to dissipate for LS streams after reaching significant size (~1,000 km^2) where TFI scores began to resemble that of similar-sized FS streams. To compensate for the karst effect, streams with sinkhole densities >0.03/km^2 that were in catchments >1,000 km^2 were considered FS streams.

Data Analysis

Thermal fish index scores were investigated within the FS and LS datasets independently to determine thresholds that best align with ALU definitions, based on trout responses. To this end, TFI scores were rounded down to the next lowest integer (e.g., a score of 2.9 was rounded down to 2) and the proportion of samples with trout in each TFI integer group was calculated, referred to here as percent occurrence (PO). It should be noted that this was conducted at the sample level, as opposed to site level. At some sites where trout are present fleetingly and/or in low numbers, they may not be captured in some samples from that site. As such, trout PO may be lower calculated at the sample level than at the site level. Furthermore, the percent abundance (PA) of trout averaged across samples in each TFI integer group was calculated for comparison. This approach produced two measures of trout response (PO and PA) which were analyzed along the TFI gradient for each of the two stream groups (FS and LS). To establish TFI thresholds representing quantifiable transitions from CWA to WWA, plots of trout PO and PA by TFI integer group were visually investigated to find inflection points for LS and FS streams. The PO and PA measures were chosen to demonstrate the drastic difference across measures of occurrence and abundance. The occurrence measure (PO) is very different than abundance (PA), as trout can be present throughout a wide range of stream types at low abundance (e.g., one individual). Subsequently, the strict presence-of-trout measure reduces ecological meaningfulness without associated abundance measures.

Applying the site classification schema derived from the BRT and recursive partitioning analysis of the LD dataset to the calibration dataset allowed for analysis of TFI response to those classifications alongside the stress categories. TFI scores were plotted and regressed to test for responsiveness to longitudinal gradient and stress levels. The datasets were tested for within-group normalcy and homogeneity of variance by inspecting residual distributions from linear models and Shapiro-Wilk tests. Final group sample sizes were relatively small and non-normal distributions were not all successfully transformed to meet parametric assumptions. Subsequently, Kruskal-Wallis chi-squared tests were used to measure among-group longitudinal differences of the final classification groups, using LD sites, followed by Dunn's test of multiple comparison, post hoc ($\alpha = 0.05$) adjusted using Bonferroni correction. Effect size (ETA^2) and magnitude were calculated from results using 0.01 to < 0.06 (small effect), 0.06 to < 0.14 (moderate effect) and >= 0.14 (large effect; Tomczak and Tomczak, 2014). Least disturbed sites were used to test for longitudinal response, minimizing effects from potential stressors. This procedure was repeated within stress level groups to measure significant differences in stress effect. Discrimination efficiency (DE) between LD and S sites was calculated to measure the TFI's ability to characterize stress (i.e., how much overlap exists between the LD and S TFI scores) (Barbour et al. 1999, Gerritsen et al. 2000). To measure DE, the percentage of S sites under the 75th percentile for LD sites was calculated by:

$$\%DE = \left(\frac{A}{B}\right) * 100$$

where, *A* is the number of S sites scoring below the 75th percentile for LD range and *B* is the total number of S sites.

After calculating DE, the 95th percentile of the LD sites within each group was used to establish impairment thresholds for assessment decisions. The 95th percentile is considered a high threshold for impairment which has two important considerations on assessments: 1) confidence in impairing a stressed site is increased, 2) confidence in not impairing a stressed site is reduced. For example, if a stressed site is below the 95th percentile of the LD TFI range, it would be considered attaining. The decision to use the 95th percentile of LD sites is based on two reasons: 1) the modWQI is continuous in nature and allows for comparisons of stress response along a robust gradient of water quality across all stream classes; and 2) using Habcat scores (1-4) there is more confidence that LD sites are characterized as a 1 and less confidence that moderately stressed sites are characterized as a 4. For example, a moderately affected site with sedimentation issues is more likely to be classified as a 3 (stressed) or 4 (very stressed) than a 1 (not stressed); however, a site with severe habitat modifications (impoundment) may also be classified as a 3 or 4. As such, more confidence is placed on LD sites being accurately characterized, and less confidence on stressed sites being accurately characterized.

Once TFI impairment thresholds were established, validation was conducted to measure the TFI assessments ability to classify sites not used in development of the TFI. Classification efficiency (CE) was calculated to measure the percentage of sites correctly classified based on the exceedance of established TFI attainment thresholds. The validation dataset was used to classify both impaired and attaining samples based on exceedance of TFI impairment thresholds by measuring the percent correctly reclassified (i.e., the percentage of stressed sites being reclassified as impaired and the percentage of unstressed sites being reclassified as attaining).

The TFI was considered novel in both concept and application. A comparative analysis that demonstrated how the TFI compares to traditional metrics was needed to enhance understanding of metric function, both in ecological relevance and performance. Traditional metrics were calculated for the biological condition gradient (BCG; Davies and Jackson 2006) level five (BCG5), percent tolerant individuals and percent omnivorous individuals. The BCG5 attribute is generally based on relative tolerance value of a species but also includes native/non-native status. A pairwise comparison using Spearman’s rank correlation coefficient was conducted across metrics as well as the modWQI and Habcat to compare metrics responses to stress.

RESULTS

Thermal Assemblage Classes

The inflection point for trout PO was between TFI scores 6-7 in both FS and LS streams (Figure 4). Trout PA sharply decreased with TFI scores > 4 and the range of inflection was strongly noted between TFI scores 4-7 (Figure 4). Overall, the range of TFI scores from 5.0-7.0 indicates a strong transition in assemblages based on both trout abundance and occurrence. Upper thresholds were established to numerically define thermal assemblage classes that best represent the transition from an assemblage dominated by cold water species (TFI <= 5.0), to dominated by cool water species (TFI = 5.1-7.0) and dominated by warm water species (TFI > 7.0) (Figure 4).

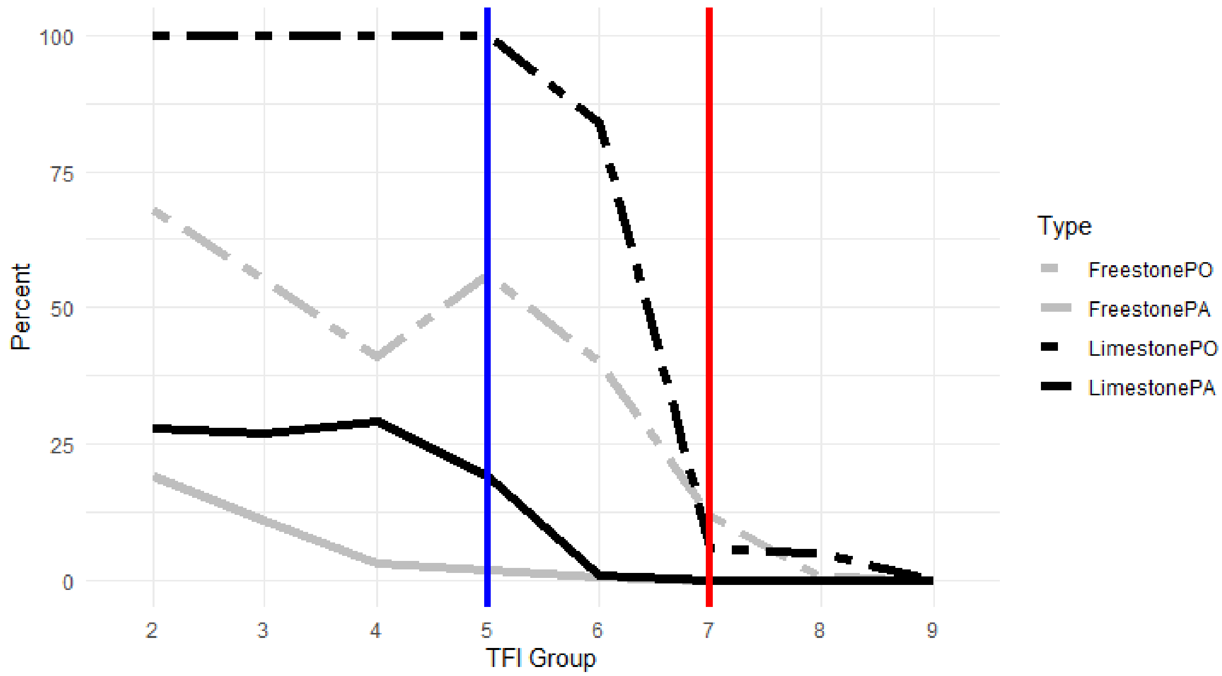

Figure 4. Percent abundance (PA) and percent occurrence (PO) of trout by thermal fish index (TFI) group in both freestone and limestone streams. Dotted lines represent PO and solid lines represent PA. The transitions from cold water assemblage (CWA) to transitional assemblage (TSA) and from TSA to warm water assemblage (WWA) are represented by blue and red vertical lines, respectively.

Modeled Results

Results from the BRTs using LD sites in FS streams (*n* trees = 1,550, lr = 0.005, tc = 5) indicated a strong longitudinal and slope effect, with minimal ecoregional effect. Variable importance was partitioned relating to stream size (76%), slope (10%), water quality (10%) and ecoregion (4%). Recursive partitioning trees indicated five classifications based on stream size, with minimal ecoregional effects in small streams (Figure 5). Boosted regression trees in the LS dataset were conducted across all stress groups, as sample sizes from the LD sites precluded analysis. Boosted regression trees in LS streams (*n* trees = 2,800, lr = 0.005, tc = 2) indicated habitat, longitudinal and water quality effects, with additional karst effects. Variable importance was partitioned relating to habitat quality (35%), stream size (33%), water quality (17%) and karst and limestone geology (sum = 15%). Recursive partitioning trees indicated a single split based on stream size (Figure 6).

The specific catchment-size ranges were modified slightly from recursive partitioning tree output to maintain sufficient sample sizes in each group and to ensure ecological relevance of the longitudinal effect of mean TFI distribution (i.e., the mean TFI increased as drainage area increase), maintaining ecological conformance with RCC.

The ecoregional effect in small FS streams was reinvestigated for a smaller catchment group to maintain a consistent classification scheme. Six final stream type/longitudinal classifications were determined by stream type and upper range of catchment area (km^2) as: LS<1000, FS<40, FS<150, FS<550, FS<6000 and FS>6000; hereafter, these classifications are referred to as drainage area groups (DAGs; Figure 7).

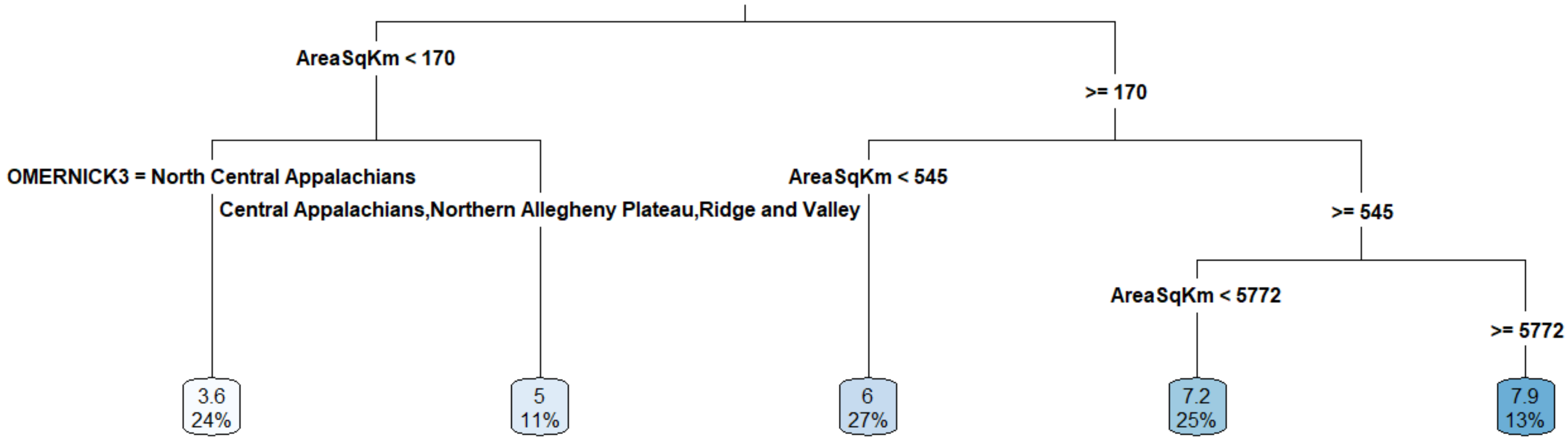

Figure 5. Recursive partitioning tree model of least disturbed (LD) sites showing important variables to classify freestone streams (FS) are generally related to catchment size and ecoregion. The bottom "leaflets" correspond to the mean thermal fish index (TFI) and the percentage of the dataset within each group.

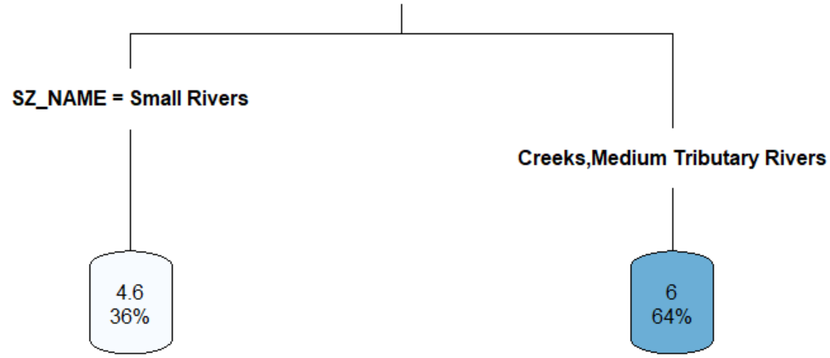

Figure 6. Recursive partitioning tree model showing the most important variable to classify limestone streams (LS) is stream size. The size classifications here (i.e., small rivers, creeks, medium tributary rivers) are from Olivero et al. 2015. All stress groups within the dataset were used as the sample size using only least disturbed (LD) sites precluded analysis.

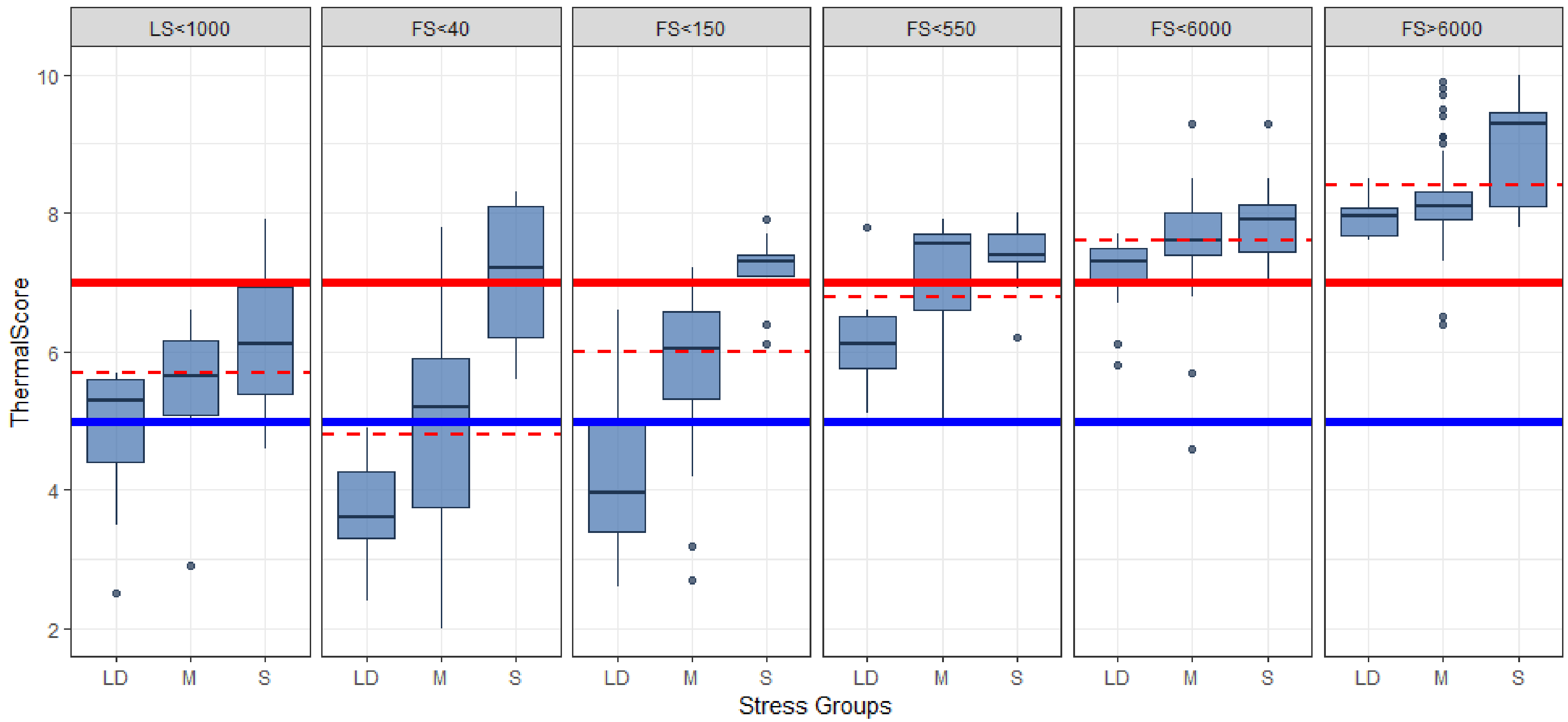

Figure 7. Boxplot of thermal fish index (TFI) scores (ThermalScore) for the final limestone (LS) and freestone (FS) drainage area groups (DAGs) (upper km^2 range). Stress groups are denoted as: Least Disturbed (LD), Moderate (M) and Stressed (S). Dotted red lines represents the 95th percentile of LD sites, signifying the impairment threshold for each DAG. The solid blue line (TFI = 7.0) represents the upper limit for cold water assemblage (CWA) and the solid red line (TFI = 5.0) represents the lower limit for warm water assemblage (WWA) with the transitional assemblage (TSA) range in between.

Between-DAG comparisons of LD sites in FS streams using regression and Kruskall Wallis showed a significant increase in TFI score along a longitudinal gradient (adjusted $R^2 = 0.76$, P= < 0.001, chi-squared = 53.6, $ETA^2 = 0.60$, Magnitude = Large). All mean and 95th percentile TFI estimates were increasing as DAGs increased, suggesting ecological relevance of the TFI. The 95th percentile for the LD sites within the LS DAG was 5.7. The 95th percentile for the LD sites within each longitudinally progressing FS DAG are as follows: 4.8, 6.0, 6.8, 7.6 and 8.4. Discrimination efficiencies were > 80% within each DAG, with the exception of the LS<1000 group, where the DE was 70%; the average DE across all DAGs was 88% (Table 6).

Table 6. Between-group and within-group results describing thermal scores across drainage area groups (DAGs) and stress categories, respectively. Shared superscripted letters within the DAG column (i.e., ^{abc}) designate non-significant differences of the least disturbed (LD) groups between DAGs (Dunn's test, p < 0.05). Sample sizes for each stress group and shared superscripted letters within the same cell designate non-significant differences across stress categories within each DAG (Dunn's test, p < 0.05). Kruskal-Wallis test chi-squared values in bold represent significant results (p < 0.05); values in bold italic represent significant results (p < 0.01).

DAG	n = LD, M, S	chi-squared	ETA^2	Magnitude	DE
LS<1000^{abc}	9^{a}, 6^{ab}, 10^{b}	4.92	0.13	Moderate	70%
FS<40^{ab}	6^{a}, 31^{a}, 15^{b}	***24.93***	0.46	Large	100%
FS<150^{ab}	16^{a}, 30^{b}, 9^{c}	***27.65***	0.48	Large	100%
FS<550^{ac}	18^{a}, 30^{bc}, 9^{c}	***18.89***	0.31	Large	89%
FS<6000^{cd}	15^{a}, 30^{bc}, 8^{c}	**8.48**	0.13	Moderate	88%
FS>6000^{d}	6^{a}, 82^{ab}, 11^{c}	**8.57**	0.06	Moderate	82%
Average		15.57			88%

TFI precision estimates measured with coefficient of variation (CV) and standard deviation (SD) across all sites averaged 4.3% (TFI score ± 0.3) and 0.25, respectively. The highest CV and SD was observed in the FS<150 DAG, averaging 8.8% (TFI score ± 0.7) and 0.4, respectively (Table 7). Classification efficiency, calculated to validate the calibration dataset and averaged across DAG groups, was 95% for LD sites and 87% for S sites (
Table 8).

Table 7. Precision estimates using standard deviation (SD) and coefficient of variation (CV), expressed in percentage, for repeated sites within each drainage area group (DAG), regardless of stress level.

DAG	SD	CV %	n
LS<1000	0.2	4.0	16
FS<40	0.1	1.8	11
FS<150	0.4	8.8	59
FS<550	0.2	3.2	61
FS<6000	0.3	4.5	39
FS>6000	0.3	3.3	178

Table 8. Validation classification efficiency: the percent stressed above impairment threshold and percent least disturbed under impairment threshold from the validation dataset. Values in parentheses denote sample size.

	LS<1000	FS<40	FS<150	FS<550	FS<6000	FS>6000	Avg.
Least Disturbed	100% (2)	100% (2)	100% (4)	100% (4)	100% (6)	67% (3)	95%
Stressed	50% (4)	100% (1)	100% (3)	67% (3)	100% (1)	100% (3)	87%
Avg.	67%	100%	100%	86%	100%	83%	91%

Pairwise comparisons of the TFI to traditional metrics demonstrated numerous, significant correlations that were generally considered weak to moderate relationships (Figures 8-13). Strong and significant correlations were observed between the BCG5 and percent tolerant metrics throughout all DAGs. The omnivore metric tended to correlate with other metrics in larger streams but was considered highly variable. Relationships between the TFI and water quality were noted in all DAGs, except in the FS>6000 DAG, where the relationship was reduced. Across all DAGs the TFI consistently outperformed traditional metrics in response to water quality and habitat based on Spearman’s correlation coefficients.

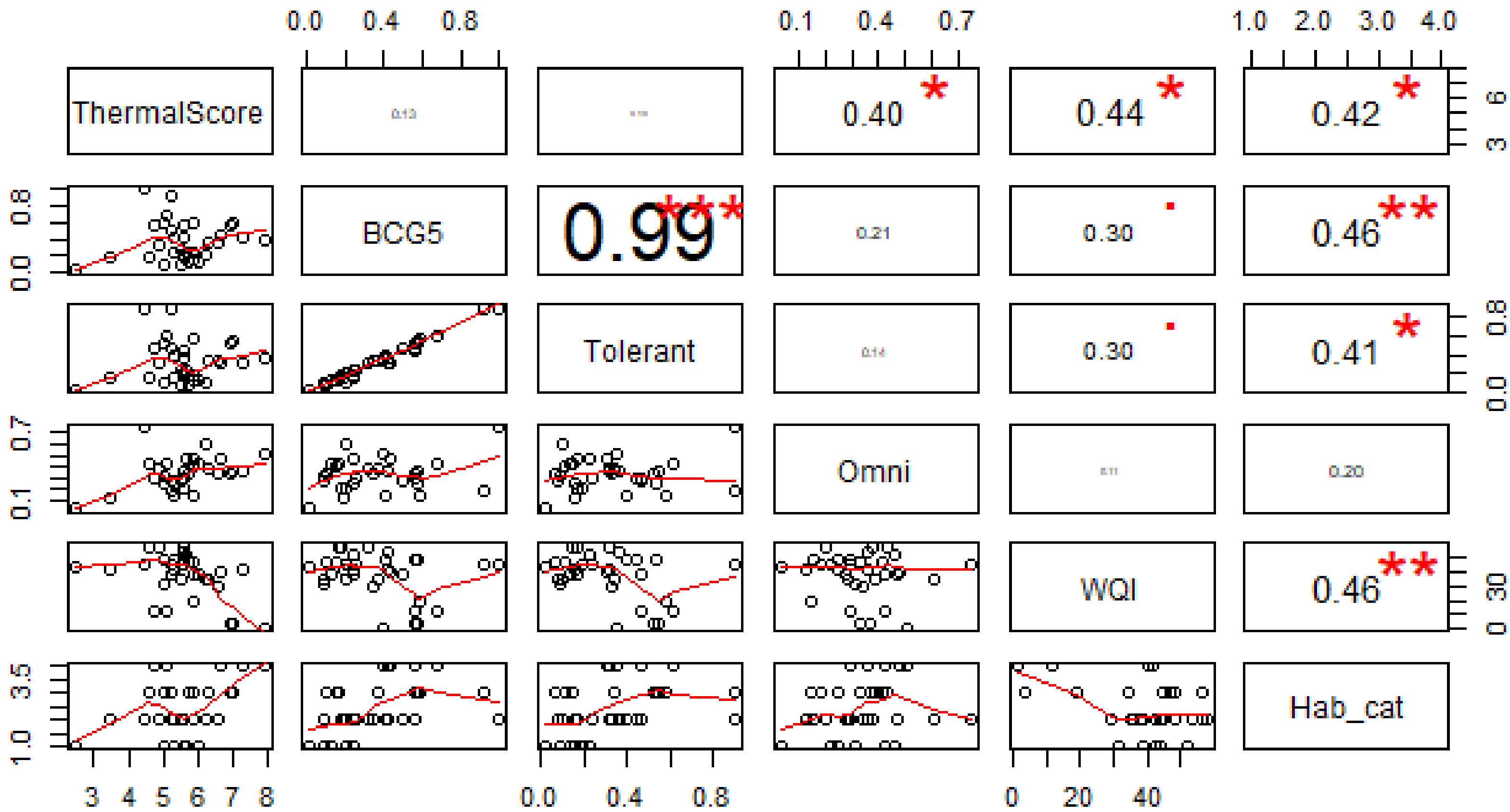

Figure 8. Pairwise comparison, using Spearman's correlation coefficient, of the thermal fish index score (ThermalScore) to traditional metrics: Biological Condition Gradient category 5 (BCG5), percent tolerant (Tolerant), percent omnivorous (Omni), water quality index (WQI) and habitat (Habcat) in LS<1000 streams. Fitted red lines are LOESS smoothed. *** (P<0.001), ** (P<0.01), * (P<0.05)

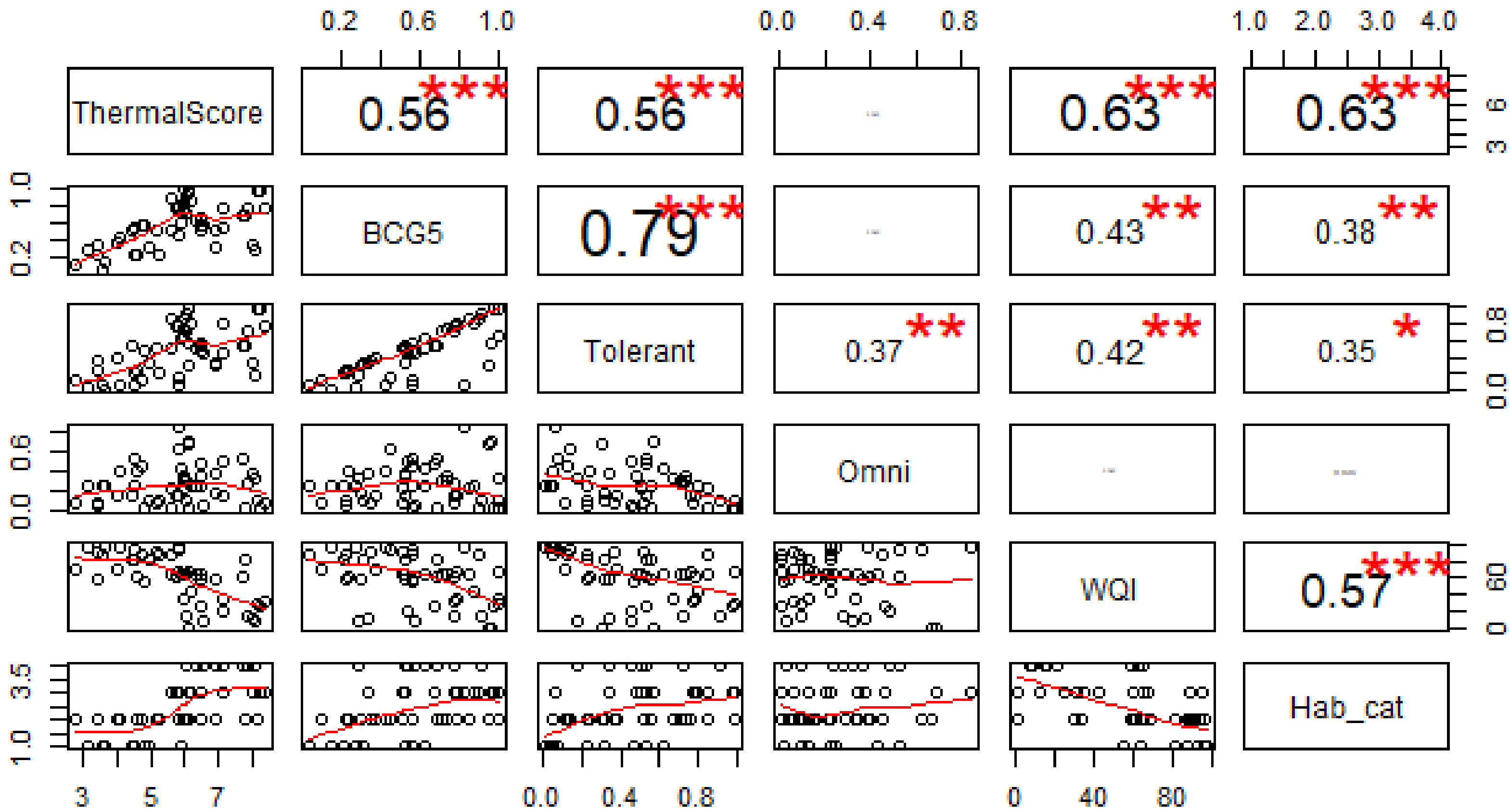

Figure 9. Pairwise comparison, using Spearman's correlation coefficient, of the thermal fish index score (ThermalScore) to traditional metrics: Biological Condition Gradient category 5 (BCG5), percent tolerant (Tolerant), percent omnivorous (Omni), water quality index (WQI) and habitat (Habcat) in FS<40 streams. Fitted red lines are LOESS smoothed. *** (P<0.001), ** (P<0.01), * (P<0.05)

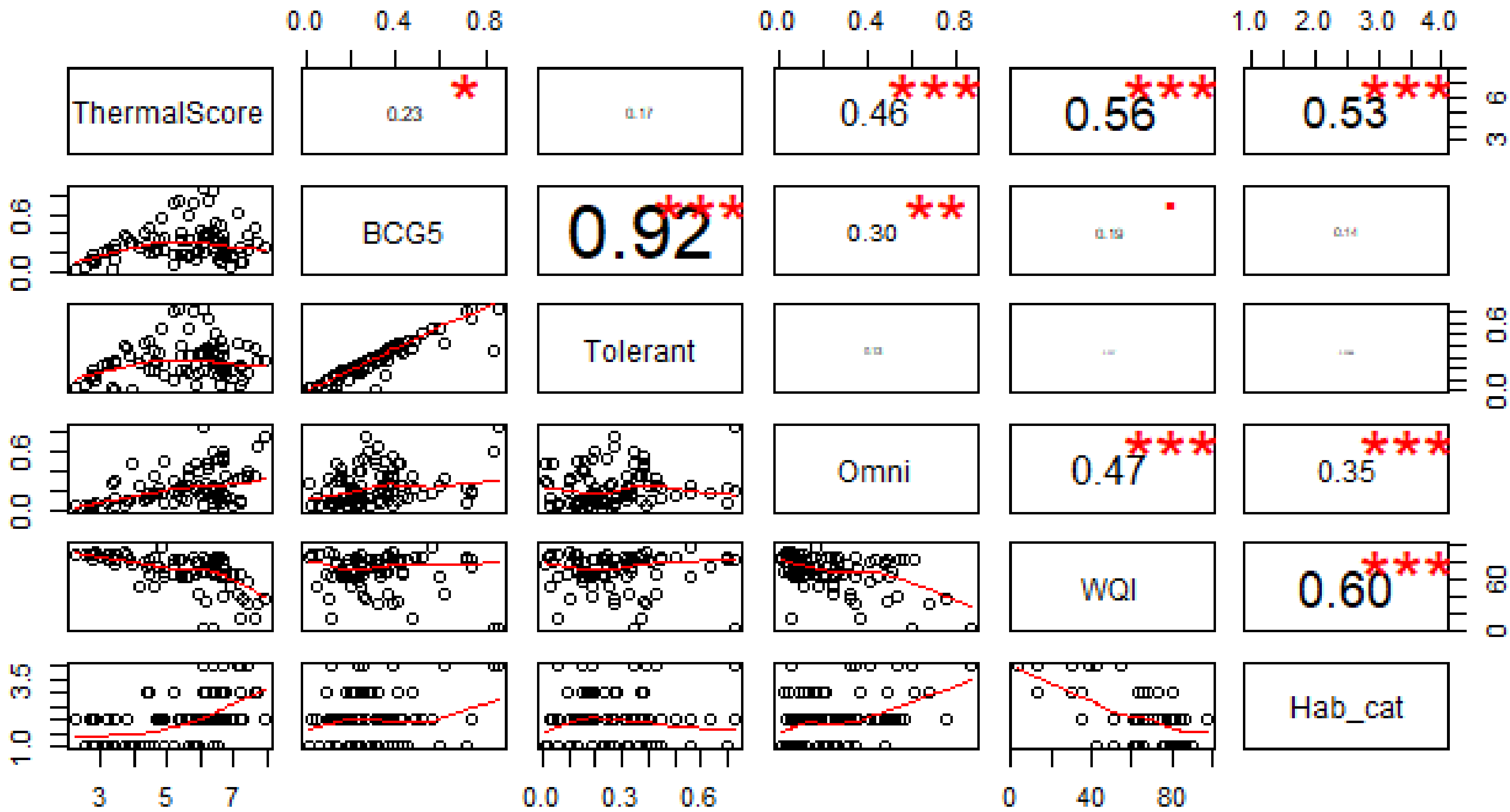

Figure 10. Pairwise comparison, using Spearman's correlation coefficient, of the thermal fish index score (ThermalScore) to traditional metrics: Biological Condition Gradient category 5 (BCG5), percent tolerant (Tolerant), percent omnivorous (Omni), water quality index (WQI) and habitat (Habcat) in FS<150 streams. Fitted red lines are LOESS smoothed. *** (P<0.001), ** (P<0.01), * (P<0.05)

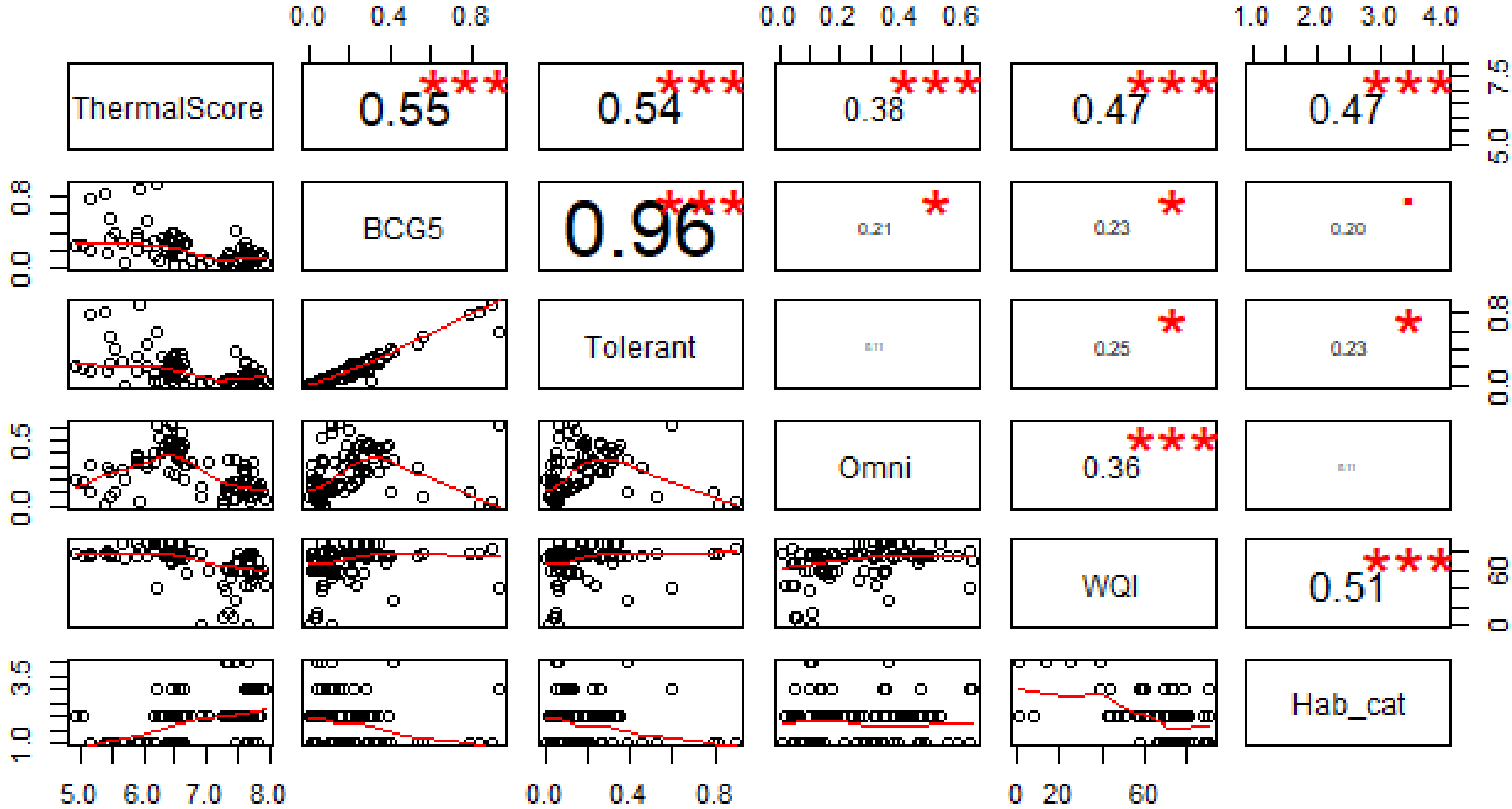

Figure 11. Pairwise comparison, using Spearman's correlation coefficient, of the thermal fish index score (ThermalScore) to traditional metrics: Biological Condition Gradient category 5 (BCG5), percent tolerant (Tolerant), percent omnivorous (Omni), water quality index (WQI) and habitat (Habcat) in FS<550 streams. Fitted red lines are LOESS smoothed. *** (P<0.001), ** (P<0.01), * (P<0.05)

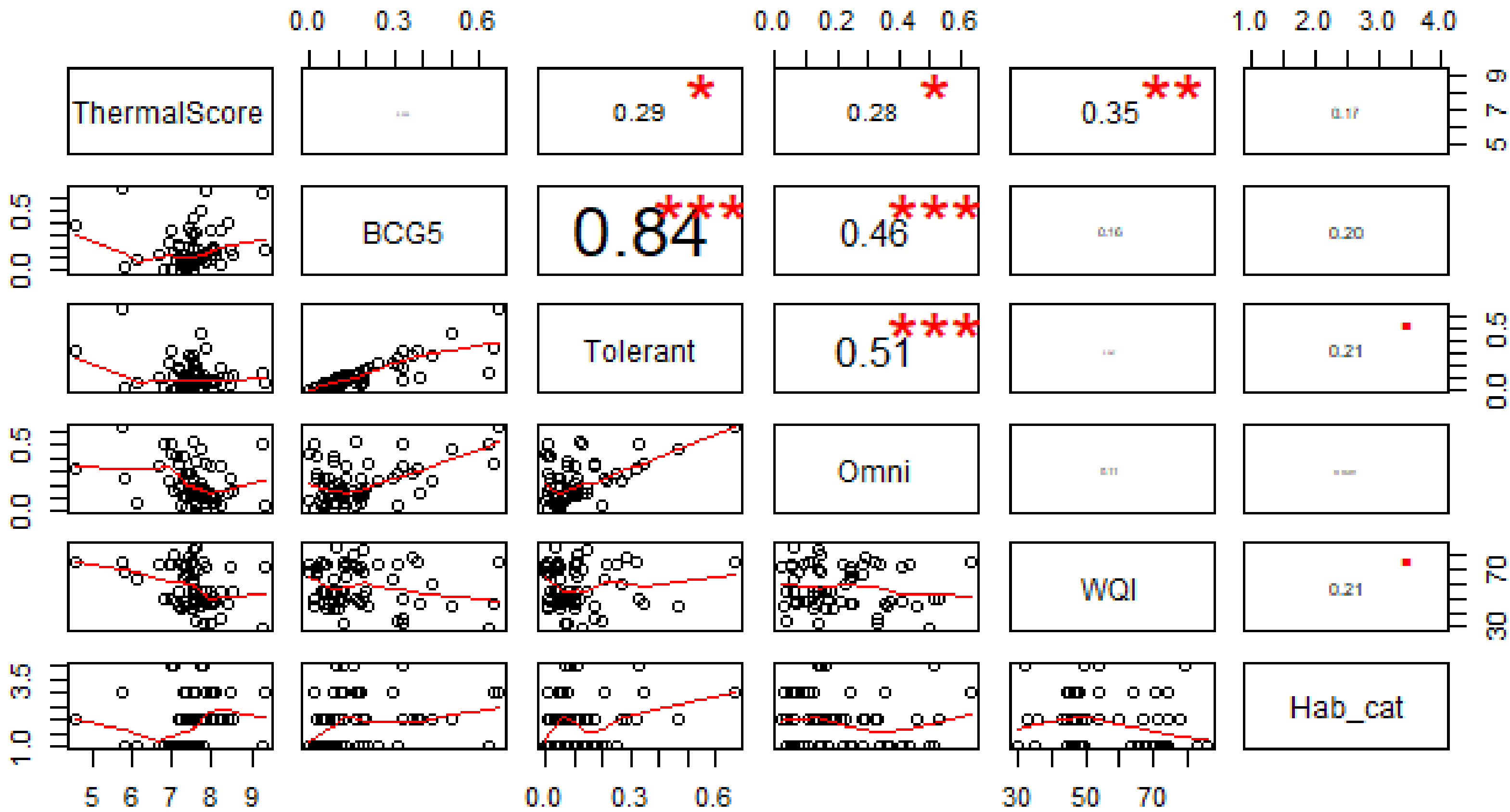

Figure 12. Pairwise comparison, using Spearman's correlation coefficient, of the thermal fish index score (ThermalScore) to traditional metrics: Biological Condition Gradient category 5 (BCG5), percent tolerant (Tolerant), percent omnivorous (Omni), water quality index (WQI) and habitat (Habcat) in FS<6000 streams. Fitted red lines are LOESS smoothed. *** (P<0.001), ** (P<0.01), * (P<0.05)

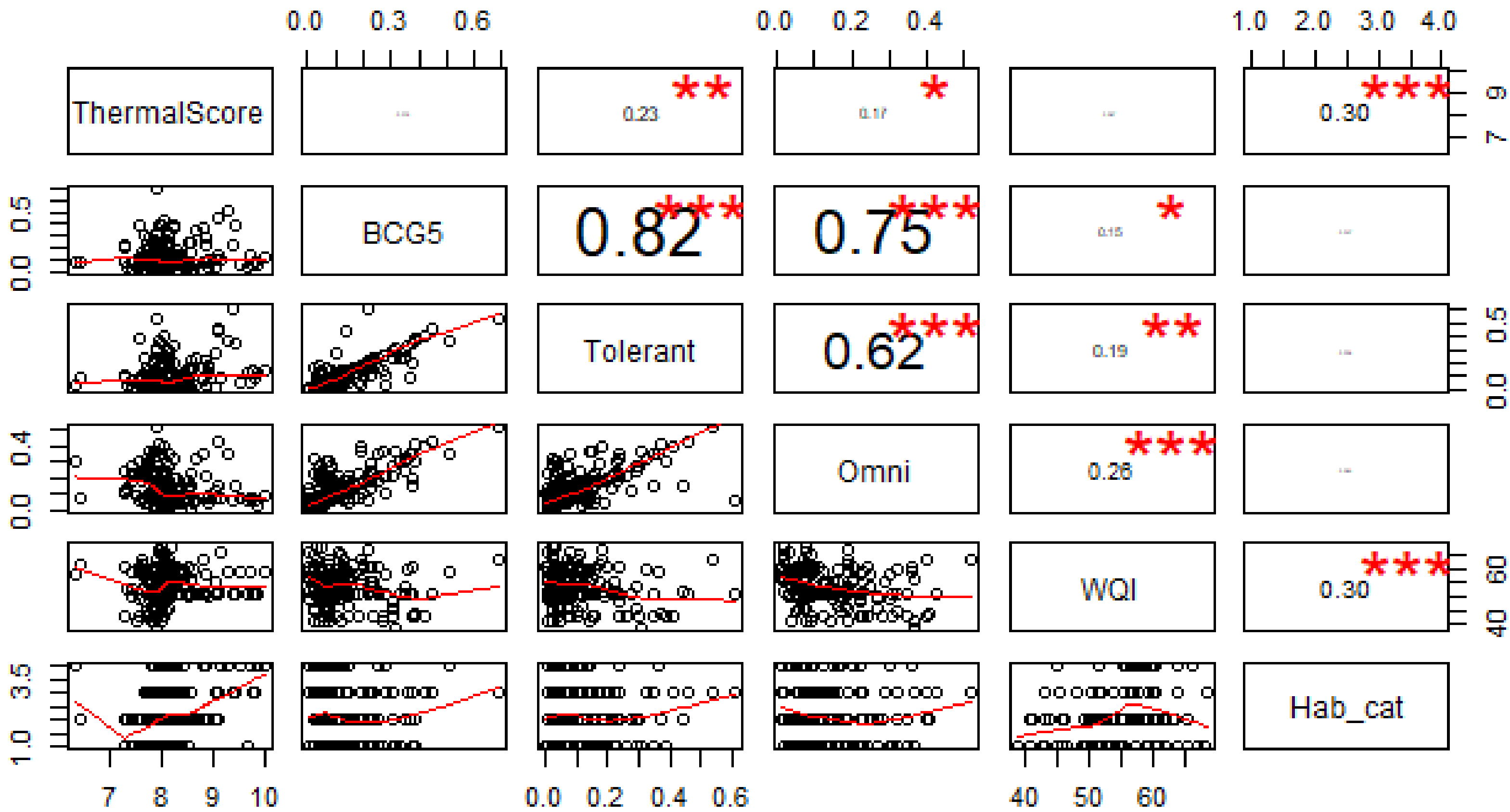

Figure 13. Pairwise comparison, using Spearman's correlation coefficient, of the thermal fish index score (ThermalScore) to traditional metrics: Biological Condition Gradient category 5 (BCG5), percent tolerant (Tolerant), percent omnivorous (Omni), water quality index (WQI) and habitat (Habcat) in FS>6000 streams. Fitted red lines are LOESS smoothed. *** (P<0.001), ** (P<0.01), * (P<0.05)

DISCUSSION

In preliminary classification analyses during TFI development, limestone streams were identified as a distinct group by their unique ability to support CWAs in larger stream sizes than FS counterparts. Catchment area was determined to be the strongest natural predictor of the TFI in both FS and LS systems, with slope being an important (albeit secondary) predictor. This result is beneficial as it provides a template for the transition of fish assemblages along a longitudinal gradient for both stream types. By taking into account key factors like stream size and stream type, the TFI provides an ecologically relevant, numerical indicator of thermal fish assemblages based on environmental characteristics (e.g., DAGs). The longitudinal regression response of the TFI using the LD sites in this analysis was strong and significant (adjusted $R^2 = 0.76$, $p = < 0.001$) and meets expectations for general ecological relevance, based on the RCC (Vannote et al. 1980).

On the surface, the LS<1000 DAG appears to be unique with respect to the TFI when compared to FS DAGs. This group had the lowest DE (70%) and lowest CE (in the S group, 50%) recorded. However, the reason for this apparent discrepancy is attributed to one major factor, trout-stocking. Herein, all (100%) LS streams that were considered S yet had a TFI score below the impairment threshold are streams regularly stocked with trout, thereby lowering the TFI score. This concept identifies a small degree of complexity that may be present in all fish-based bioassessments, where intentional (or unintentional) stocking co-occurs. Herein, a tradeoff exists between enhancing valuable recreational opportunities through stocking and measuring the response of fish assemblages that may not be solely driven by waterbody conditions. Subsequently, the effect of stocking should be realized and treated as inherent, yet subtle “noise” that will likely be present in many fish-based bioassessments.

The lowest DE in the FS DAGs was noted in FS>6000, with a DE of 82%. This is attributed to: 1) a stress effect-size change and 2) using a four-tiered habitat category (Habcat) as a measure of stress. Generally, as stream size increases the range of water quality decreases, where large rivers tend to occupy a more narrow and centralized distribution, as an effect of dilution (see generally, Nilssan 2008); this phenomenon was observed with the modWQI in the FS>6000 DAG (Figure 14). Alternatively, small streams are more susceptible to the extreme ends of the water quality range, being “very good” in heavily forested headwaters to “very poor” in effluent-dominated headwaters. This important concept suggests as streams increase in size the effect of water quality stress may be mitigated, to some degree, and the effect of habitat quality may become more important.

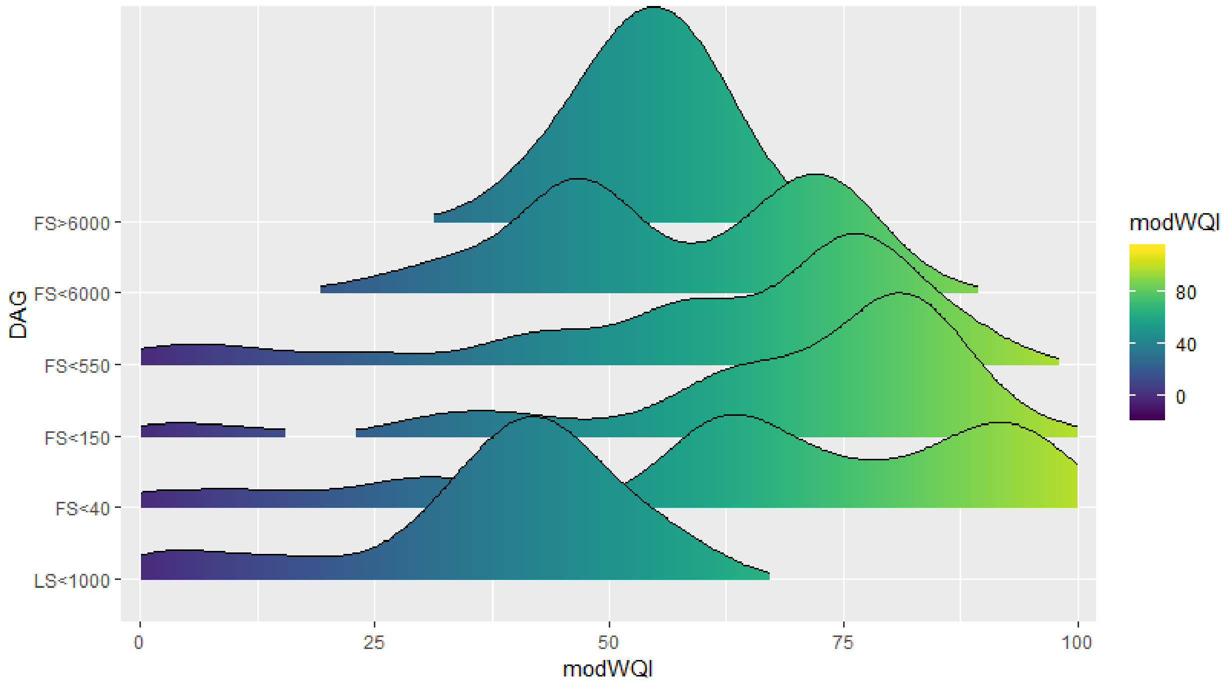

Figure 14. The shift in water quality range distribution across drainage area groups (DAG), as measured from the modified water quality index (modWQI).

Using a categorical habitat metric (Habcat) as a measure of habitat quality successfully standardized habitat stress across all sampling methodologies employed (wadeable vs. nonwadeable), but was not without consequence for TFI development. The Habcat metric is generally, but not always, comparable across all sites within the same stress level. For example, a river characterized as a 4 for sediment deposition may not be the same stress as a river that is impounded for miles, also characterized as a 4. This concept was the major driver of the reduced DE in FS>6000, where naturally occurring increases in sedimentation caused a site to fall in the S group. This reaffirms aforementioned confidence in correctly identified LD sites and reduced confidence in correctly identifying S sites.

The TFI was responsive to changes along a longitudinal gradient, temperature and stress (both habitat and water quality). The effect of water quality stress on the TFI was reduced longitudinally, as larger DAGs tended to occupy a narrow and more-central range of the modWQI (Figure 14). The effect of habitat on the TFI was important across both FS and LS groups (and DAGs) and tended to increase dramatically with increased sedimentation and impounding (Figure 15). These observations are important as multiple stressors have synergistic, antagonistic or additive effects on the TFI. For example, as water quality is reduced by agricultural activities and loss of riparian areas, changes to instream habitat and temperature will likely parallel the reductions in water quality, having a dramatic combined effect on the TFI. Alternatively, a stream with mining influences may have reduced water quality, without drastic changes in habitat and temperatures, which may have a smaller effect on the TFI. In other words, as the number of stressors and/or intensity of stressors increases, increases in the TFI are expected. This is a desired outcome from a management perspective, as measured improvements in individual stressors may result in measurable recovery. For example, best management practices applied to small reaches of a larger watershed may have localized, measurable biological effects.

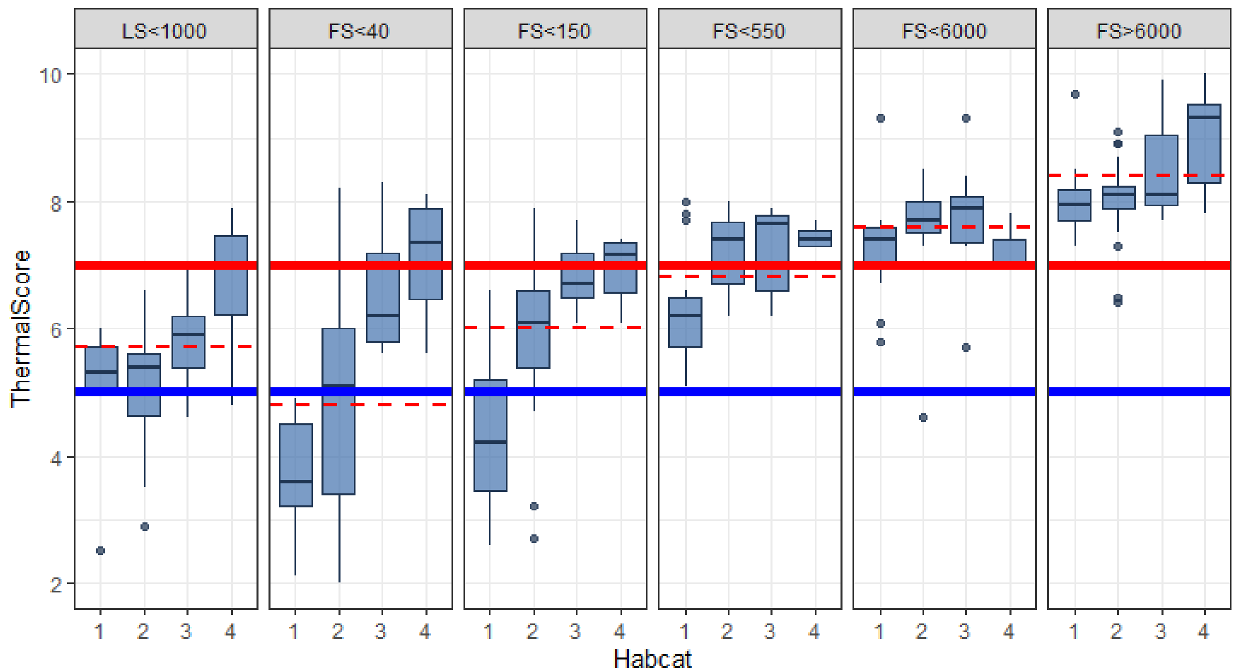

Figure 15. Boxplot of thermal fish index (TFI) scores (ThermalScore) for the final limestone (LS) and freestone (FS) drainage area groups (DAGs) (upper km^2 range). Habitat category (Habcat) groups 1-4 are on a gradient of good to poor, respectively. Dotted red lines represents the 95th percentile of least disturbed (LD) sites signifying the impairment threshold for each DAG. The solid blue line (TFI = 7.0) represents the upper limit for cold water assemblage (CWA) and the solid red line (TFI = 5.0) represents the lower limit for warm water assemblage (WWA), with the transitional assemblage (TSA) range in between.

From a comparative perspective, the TFI may appear to be quite simple in design. In reality, the TFI should be viewed as a comprehensive metric, in that: 1) all species and individuals within the assemblage are provided equal consideration based on relative abundance; 2) the TFI can be applied uniformly across the State, basins, or ecoregions; 3) the TFI has an ecologically meaningful output of assemblage thermal class (cold vs. warm) as opposed to a purely statistically-derived construct; and 4) the TFI exhibits fairly strong correlation with some other common bioassessment metrics and with indicators of water quality and habitat quality. The TFI performed as well as or better than the three traditional metrics (BCG5, percent tolerant, percent omnivorous) in response to water quality and habitat conditions (Figures 8-13). Traditional metrics convey important ecological and biological concepts that can complement TFI scores in fish-based assessments of water quality. However, interpreting and conveying the results of traditional metrics can be challenging at times without identifying a comparative baseline to serve as the relevant “reference condition”. For traditional metrics, the relevant reference condition may change based on the pool of available species, not only within a DAG but also across major drainage basins. For example, species richness in an FS<150 stream in the Ohio River basin may be twice the species richness of a comparable stream in the Susquehanna River basin. Additionally, some traditional metrics may have a bimodal response across DAGs, increasing in response to increasing stress in some DAGs, but decreasing in response to increasing stress in other DAGs. For example, species richness tends to increase in small watersheds in response to increasing stress (e.g., cold/cool species displaced by warmer species), but this metric tends to decrease in response to increasing stress in larger watersheds. Similar patterns can be observed in the traditional metric comparisons made within and across DAGs (Figures 8-15). Finally, it is essential that comparisons of metric responses be relative to assessment determinations (i.e., attaining vs. impaired). Herein, comparisons of traditional metric responses to a relevant reference condition should be conducted relative to: 1) basin, 2) DAG and 3) assessment determination. The relevant reference condition approach provides a systematic method for making comparisons to one or more reference (least disturbed) sites for further insight into important ecological or biological processes that contribute to a “balanced indigenous community”. For purposes of this document, the definition of a “balanced indigenous community” (or population) is as defined in 40 CFR § 125.71(c) (see EPA 2010):

> *“a biotic community typically characterized by diversity, the capacity to sustain itself through cyclic seasonal changes, presence of necessary food chain species and by lack of domination by pollution tolerant species. Such a community may include historically non-native species introduced in connection with a program of wildlife management and species whose presence or abundance results from substantial irreversible environmental modifications. Normally however, such a community will not include species whose presence or abundance is attributable to the introduction of pollutants that will be eliminated by compliance by all sources with section 301(b)(2) of the [Clean Water] Act; and may not include species whose presence or abundance is attributable to alternative effluent limitations imposed pursuant to section 316(a).”*

Two of the three requirements for defining a relevant reference condition establish a baseline community, and its diversity and structure, that is indigenous to a set of environmental stream conditions: basin and DAG. The third requirement for defining the relevant reference condition relies on the assessment determination to convey the response of species (and their abundance) to stressors – water quality, habitat or temperature – as described herein.

Numerous unique samples were noted within the dataset that warrant further discussion. Unique samples within each DAG were apparent in both directions. The assemblages with lower TFI scores than the rest of the distribution were generally caused by: 1) hydrologic alterations in the form of augmented bottom-releases from upstream impoundments; 2) unique natural features such as increased groundwater volume or canopy cover; and 3) unrepresentative sample locations influenced strongly by proximal tributaries. Individual streams or stream segments that have a natural ability to maintain colder fish assemblages can be viewed as unique and important from an ecological and/or recreational perspective. For example, the Delaware River near Balls Eddy is in the FS<6000 DAG and has achieved TFI scores as low as 4.6. This exceptionally low score for such a large DAG is the result of flow management and cold water releases from upstream reservoirs. This portion of the Delaware River remains an important recreational destination for trout fishing. Conversely, while flow management and cold water augmentation scenarios may initially be portrayed as an improvement, these practices are not without consequences. An example of these consequences is apparent in Clarks Creek near Harrisburg, a small tributary to the Susquehanna River. This stream is impounded by a drinking water reservoir. Fish surveys were conducted at two sites on Clarks Creek, bracketing the reservoir. The site downstream of the reservoir had a catchment area of 62 km^2 (FS<150) with a TFI score of 5.4. This site is augmented by both cold water releases from the reservoir and trout stocking. The TFI score of 5.4 is below the impairment threshold of 6.0 based on the site's DAG; water quality is supportive of trout stocking and the assemblage is characterized as a TSA. The site upstream of the reservoir had a much smaller catchment area of 34 km^2 (FS<40) with a TFI score of 7.8. The TFI score for this DAG is well above the 95th percentile of reference for the FS<40, set at 4.8. The upper site had excellent water quality and habitat but was located only 500 meters upstream of the impounded portion of the reservoir. Herein, the upstream site was influenced by fishes migrating upstream from the reservoir and was dominated by the family Centrarchidae. In other words, fishes indigenous to a cold water habitat were being replaced by fishes indigenous to a warm water habitat; conceptually, a "thermally invasive species". This effect is therefore considered a "biological pollution" as a result of significant habitat alterations within proximity (Pringle 1997, Elliott 2003). Consequently, bolstering of CWAs downstream of cold water releases from impoundments was observed but a reduction of CWAs within and upstream of impounded areas was also observed. Additionally, this tradeoff is likely to be in both directions of top release "spillover" impoundments (warmer assemblages upstream and downstream).

Overall, the TFI responded significantly to changes in stream size and stress in both freestone and limestone waterbodies across Pennsylvania. The discrimination and classification efficiencies were within acceptable ranges, averaging 88% and 91% across all groups, respectively. Precision estimates measured from coefficient of variation were within (below) recommended threshold ranges of 10-15% (Stribling et al. 2008) and averaged 4.3%, with maxima still within acceptable limits, topping out at 8.8%. The TFI correlated with, and often outperformed, traditional metrics in comparative analysis. These factors combined with added benefits of a large spatial application and ecological relevance solidify the TFI as a tool for assessing and evaluating fish assemblages across Pennsylvania's lotic waterbodies.

LITERATURE CITED

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. 2nd edition. EPA 841-B-99-002. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

Beitinger, T. L., W. A. Bennett, and R. W. McCauley. 2000. Temperature tolerances of North American freshwater fishes exposed to dynamic changes in temperature. Environmental Biology of Fishes 58(3):237–275.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 2017. Classification and regression trees. Routledge, Boca Raton. DOI: https://doi.org/10.1201/9781315139470.

Brown, G. W., and J. T. Krygier. 1970. Effects of clear-cutting on stream temperature. Water Resources Research 6(4):1133–1139.

Brown, L. R., R. H. Gray, R. M. Hughes, and M. R. Meador. 2005. Introduction to effects of urbanization on stream ecosystems. American Fisheries Society Symposium 47:1–8.

Burcher, C. L., M. E. McTammany, E. F. Benfield, and G. S. Helfman. 2008. Fish assemblage responses to forest cover. Environmental Management 41(3):336–346.

Caissie, D. 2006. The thermal regime of rivers: A review. Freshwater Biology 51(8):1389–1406.

Carline, R. F., R. L. Dunlap, J. E. Detar, and B. A. Hollender. 2011. The fishery of Spring Creek, a watershed under siege. Technical Report Number 1. Pennsylvania Fish and Boat Commission, Harrisburg, Pennsylvania.

Carpenter, K. E., and J. S. Huxley. 1928. Life in inland waters with especial reference to animals. 1st edition. Sidgwick and Jackson, Ltd.

Coker, G. A., C. B. Portt, and C. K. Minns. 2001. Morphological and ecological characteristics of Canadian freshwater fishes. Fisheries and Oceans Canada, Burlington, Ontario.

Coutant, C. C. 1975. Temperature selection by fish – A factor in power-plant impact assessments. Environmental effects of cooling systems at nuclear power plants. International Atomic Energy Agency 575–597.

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random forests for classification in ecology. Ecology 88(11):2783–2792.

Davies, S. P., and S. K. Jackson. 2006. The biological condition gradient: A descriptive model for interpreting change in aquatic ecosystems. Ecological Applications 16(4):1251–1266.

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.

De'ath, G. 2007. Boosted trees for ecological modeling and prediction. Ecology 88(1):243–251.

Eaton, J. G., and R. M. Scheller. 1996. Effects of climate warming on fish thermal habitat in streams of the United States. Limnology and Oceanography 41(5):1109–1115.

Elith, J., J. R. Leathwick, and T Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77(4):802–813.

Elliott, M. 2003. Biological pollutants and biological pollution–An increasing cause for concern. Marine Pollution Bulletin 46(3):275–280.

ESRI. 2018. ArcGIS Pro. Environmental Systems Research Institute, Redlands, California.

Fausch, K. D., J. O. H. N. Lyons, J. R. Karr, and P. L. Angermeier. 1990. Fish communities as indicators of environmental degradation. American Fisheries Society Symposium 8(1):123–144.

Gerritsen, J., J. Burton, and M. T. Barbour. 2000. A stream condition index for West Virginia wadeable streams. Prepared for U.S. Environmental Protection Agency, Region 3 Environmental Services Division, and U.S. Environmental Protection Agency, Office of Science and Technology, Office of Water, Washington, D.C. (Available online at: https://dep.wv.gov/WWE/watershed/bio_fish/Documents/WVSCI.pdf.)

Huet, M. 1959. Profiles and biology of western European streams as related to fish management. Transactions of the American Fisheries Society 88(3):155–163.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54(2):187–211.

Jones, E. B., G. S. Helfman, J. O. Harper, and P. V. Bolstad. 1999. Effects of riparian forest removal on fish assemblages in southern Appalachian streams. Conservation Biology 13(6):1454–1465.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6(6):21–27.

Kollaus, K. A., and T. H. Bonner. 2012. Habitat associations of a semi-arid fish community in a karst spring-fed stream. Journal of Arid Environments 76:72–79.

Lammert, M., and J. D. Allan. 1999. Assessing biotic integrity of streams: Effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. Environmental Management 23(2):0257–0270.

Lessard, J. L., and D. B. Hayes. 2003. Effects of elevated water temperature on fish and macroinvertebrate communities below small dams. River Research and Applications 19(7):721–732.

Lyons, J., T. Zorn, J. Stewart, P. Seelbach, K. Wehrly, and L. Wang. 2009. Defining and characterizing coolwater streams and their fish assemblages in Michigan and Wisconsin, USA. North American Journal of Fisheries Management 29(4):1130–1151.

Mohseni, O., H. G. Stefan, and J. G. Eaton. 2003. Global warming and potential changes in fish habitat in US streams. Climatic Change 59(3):389–409.

Naghibi, S. A., H. R. Pourghasemi, and B. Dixon. 2016. GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. Environmental Monitoring and Assessment 188(1):44.

Nelson, K. C., and M. A. Palmer. 2007. Stream temperature surges under urbanization and climate change: Data, models, and responses. Journal of the American Water Resources Association 43(2):440–452.

Nilsson, C., and B. Malm Renöfält. 2008. Linking flow regime and water quality in rivers: A challenge to adaptive catchment management. Ecology and Society 13(2):18.

O’Driscoll, M. A., and D. R. DeWalle. 2006. Stream-air temperature relations to classify stream-ground water interactions in a karst setting, central Pennsylvania, USA. Journal of Hydrology 329(1):140–153.

Olivero, S. A., A. Barnett, and M. G. Anderson. 2015. A stream classification for the Appalachian Region. The Nature Conservancy, Eastern Conservation Science, Eastern Regional Office, Boston, Massachusetts.

Plafkin, J. L., M. T. Barbour, K. D. Porter, S. K. Gross, and R. M. Hughes.1989. Rapid bioassessment protocols for use in streams and rivers: Benthic macroinvertebrates and fish. EPA 440-4-89-001. U.S. Environmental Protection Agency, Office of Water Regulations and Standards, Washington, D.C.

Poff, N. L., and J. V. Ward. 1990. Physical habitat template of lotic systems: Recovery in the context of historical pattern of spatiotemporal heterogeneity. Environmental Management 14(5):629.

Poole, G. C., and C. H. Berman. 2001. An ecological perspective on in-stream temperature: Natural heat dynamics and mechanisms of human-caused thermal degradation. Environmental Management 27(6):787–802.

Prasad, A. M., L. R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. Ecosystems 9(2):181–199.

Pringle, C. M. 1997. Exploring how disturbance is transmitted upstream: Going against the flow. Journal of the North American Benthological Society 16(2):425–438.

Qian, S. S. 2016. Environmental and ecological statistics with R. Chapman and Hall/CRC Applied Environmental Statistics, Boca Raton, Florida.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (Available online at: https://www.R-project.org/.)

Ridgeway, G. 2006. Generalized boosted regression models: A guide to the gbm Package. Documentation on the R Package 'gbm' version 1.5-7. (Available online at: https://rdocumentation.org/packages/gbm/versions/1.5-7.)

Shank, M. K., A. M. Henning, and A. S. Leakey. 2016. Examination of a single-unit, multiple-pass electrofishing protocol to reliably estimate fish assemblage composition in wadeable streams of the Mid-Atlantic Region of the USA. North American Journal of Fisheries Management 36(3):497–505.

Shuter, B. J., J. A. MacLean, F. E. J. Fry, and H. A. Regier. 1980. Stochastic simulation of temperature effects on first-year survival of Smallmouth Bass. Transactions of the American Fisheries Society 109(1):1–34.

Sokal, R. R., and N. L. Oden. 1978. Spatial autocorrelation in biology: 1. Methodology. Biological Journal of the Linnean Society 10(2):199–228.

Southwood, T. R. 1977. Habitat, the templet for ecological strategies?. Journal of Animal Ecology 46(2):337–365.

Steffy, L. Y., and S. S. Kilham. 2006. Effects of urbanization and land use on fish communities in Valley Creek watershed, Chester County, Pennsylvania. Urban Ecosystems 9(2):119–133.

Stanfield, L. W., and B. W. Kilgour. 2013. How proximity of land use affects stream fish and habitat. River Research and Applications 29(7):891–905.

Stauffer, J. R., R. W. Criswell, and D. P. Fischer. 2016. The fishes of Pennsylvania. Cichlid Press, El Paso, TX.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectations for the ecological condition of streams: The concept of reference condition. Ecological Applications 16(4):1267–1276.

Stribling, J. B., B. K. Jessup, and D. L. Feldman. 2008. Precision of benthic macroinvertebrate indicators of stream condition in Montana. Journal of the North American Benthological Society 27(1):58–67.

Trauth, R., and Xanthopoulos, C. 1997. Non-point pollution of groundwater in urban areas. Water Research 31(11):2711–2718.

Therneau, T., and B. Atkinson. 2019. Rpart: Recursive partitioning and regression trees. R package ‘rpart’ version 4.1–15. (Available online at: https://CRAN.R-project.org/package=rpart.)

Tomczak, M., and E. Tomczak. 2014. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. Trends in Sport Sciences 21(1).

USEPA. 2010. Criteria and standards for the National Pollutant Discharge Elimination System. Title 40 Code of Federal Regulations. Part 125, Subpart H, §125.71. (Available online at: https://www.ecfr.gov/current/title-40/chapter-I/subchapter-D/part-125/subpart-H/section-125.71.)

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37(1):130–137.

Ward, J. V., and J. A. Stanford. 1983. The serial discontinuity concept of lotic ecosystems. Dynamics of Lotic Ecosystems 10:29–42.

Wehrly, K. E., M. J. Wiley, and P. W. Seelbach. 2003. Classifying regional variation in thermal regime based on stream fish community patterns. Transactions of the American Fisheries Society 132:18–38.

Wertz, T. A. 2021. Fish data collection protocol. Chapter 3.7, pages 44–63 *in* M. J. Lookenbill, and R. Whiteash (editors). Water quality monitoring protocols for streams and rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

Wertz, T. A., and M. K. Shank. 2019. Land use from water quality: Development of a water quality index across Pennsylvania streams. Ecosphere 10(11).

Wood, C. M., and D. G. McDonald (editors). 1997. Global warming: Implications for freshwater and marine fish. Cambridge University Press, Cambridge, United Kingdom.

Yoder, C. O., E. T. Rankin, and B. J. Armitage. 2006. Re-evaluation of the technical justification for existing Ohio River mainstem temperature criteria. Technical Report MBI/05-05-2. Report to Ohio River Valley Water Sanitation Commission, Columbus, Ohio.

Zorn, T. G., P. W. Seelbach, and M. J. Wiley. 2002. Distribution of stream fishes and their relationship to stream size and hydrology in Michigan's Lower Peninsula. Transactions of the American Fisheries Society 131(1):70–85.

APPENDIX A: THERMAL PREFERENCE BY SPECIES

Family	Common Name	Scientific Name	Numeric Value	Preference
Petromyzontidae	Ohio Lamprey	*Ichthyomyzon bdellium*	3	Cool
Petromyzontidae	Northern Brook Lamprey	*Ichthyomyzon fossor*	3	Cool
Petromyzontidae	Silver Lamprey	*Ichthyomyzon unicuspis*	3	Cool
Petromyzontidae	American Brook Lamprey	*Lampetra appendix*	3	Cool
Petromyzontidae	Sea Lamprey	*Petromyzon marinus*	3	Cool
Polydontidae	Paddlefish	*Polyodon spathula*	5	Warm
Lepisosteidae	Spotted Gar	*Lepisosteus oculatus*	5	Warm
Lepisosteidae	Longnose Gar	*Lepisosteus osseus*	5	Warm
Lepisosteidae	Shortnose Gar	*Lepisosteus platostomus*	5	Warm
Amiidae	Bowfin	*Amia calva*	5	Warm
Hiodontidae	Goldeye	*Hiodon alosoides*	5	Warm
Hiodontidae	Mooneye	*Hiodon tergisus*	5	Warm
Anguillidae	American Eel	*Anguilla rostrate*	3	Cool
Clupeidae	Blueback Herring	*Alosa aestivalis*	5	Warm
Clupeidae	Skipjack Herring	*Alosa chrysochloris*	5	Warm
Clupeidae	Hickory Shad	*Alosa mediocris*	5	Warm
Clupeidae	Alewife	*Alosa pseudoharengus*	4	Cool-Warm
Clupeidae	American Shad	*Alosa sapidissima*	5	Warm
Clupeidae	Gizzard Shad	*Dorosoma cepedianum*	5	Warm
Cyprinidae	Central Stoneroller	*Campostoma anomalum*	4	Cool-Warm
Cyprinidae	Goldfish	*Carassius auratus*	5	Warm
Cyprinidae	Northern Redbelly Dace	*Chrosomus eos*	4	Cool-Warm
Cyprinidae	Finescale Dace	*Chrosomus neogaeus*	3	Cool
Cyprinidae	Redside Dace	*Clinostomus elongatus*	3	Cool
Cyprinidae	Rosyside Dace	*Clinostomus funduloides*	2	Cold-Cool
Cyprinidae	Grass Carp	*Ctenopharyngodon idella*	5	Warm
Cyprinidae	Satinfin Shiner	*Cyprinella analostana*	5	Warm
Cyprinidae	Spotfin Shiner	*Cyprinella spiloptera*	5	Warm
Cyprinidae	Common Carp	*Cyprinus carpio*	5	Warm
Cyprinidae	Streamline Chub	*Erimystax dissimilis*	4	Cool-Warm
Cyprinidae	Gravel Chub	*Erimystax x-punctatus*	3	Cool
Cyprinidae	Tonguetied Minnow	*Exoglossum laurae*	3	Cool
Cyprinidae	Cutlip Minnow	*Exoglossum maxillingua*	3	Cool
Cyprinidae	Eastern Silvery Minnow	*Hybognathus regius*	5	Warm

Family	Common Name	Scientific Name	Numeric Value	Preference
Cyprinidae	Striped Shiner	*Luxilus chrysocephalus*	4	Cool-Warm
Cyprinidae	Common Shiner	*Luxilus cornutus*	4	Cool-Warm
Cyprinidae	Redfin Shiner	*Lythrurus umbratilis*	4	Cool-Warm
Cyprinidae	Silver Chub	*Macrhybopsis storeriana*	4	Cool-Warm
Cyprinidae	Pearl Dace	*Margariscus margarita*	2	Cold-Cool
Cyprinidae	Hornyhead Chub	*Nocomis biguttatus*	3	Cool
Cyprinidae	River Chub	*Nocomis micropogon*	3	Cool
Cyprinidae	Golden Shiner	*Notemigonus crysoleucas*	5	Warm
Cyprinidae	Comely Shiner	*Notropis amoenus*	5	Warm
Cyprinidae	Emerald Shiner	*Notropis atherinoides*	5	Warm
Cyprinidae	Silverjaw Minnow	*Notropis buccatus*	3	Cool
Cyprinidae	Blackchin Shiner	*Notropis heterodon*	4	Cool-Warm
Cyprinidae	Blacknose Shiner	*Notropis heterolepis*	4	Cool-Warm
Cyprinidae	Spottail Shiner	*Notropis hudsonius*	4	Cool-Warm
Cyprinidae	Silver Shiner	*Notropis photogenis*	4	Cool-Warm
Cyprinidae	Swallowtail Shiner	*Notropis procne*	3	Cool
Cyprinidae	Rosyface Shiner	*Notropis rubellus*	4	Cool-Warm
Cyprinidae	Sand Shiner	*Notropis stramineus*	4	Cool-Warm
Cyprinidae	Mimic Shiner	*Notropis volucellus*	4	Cool-Warm
Cyprinidae	Bluntnose Minnow	*Pimephales notatus*	4	Cool-Warm
Cyprinidae	Fathead Minnow	*Pimephales promelas*	4	Cool-Warm
Cyprinidae	Eastern Blacknose Dace	*Rhinichthys atratulus*	3	Cool
Cyprinidae	Longnose Dace	*Rhinichthys cataractae*	3	Cool
Cyprinidae	Western Blacknose Dace	*Rhinichthys obtusus*	3	Cool
Cyprinidae	Creek Chub	*Semotilus atromaculatus*	3	Cool
Cyprinidae	Fallfish	*Semotilus corporalis*	4	Cool-Warm
Catostomidae	River Carpsucker	*Carpiodes carpio*	5	Warm
Catostomidae	Quillback	*Carpiodes cyprinus*	5	Warm
Catostomidae	Highfin Carpsucker	*Carpiodes velifer*	5	Warm
Catostomidae	Longnose Sucker	*Catostomus catostomus*	2	Cold-Cool
Catostomidae	White Sucker	*Catostomus commersonii*	3	Cool
Catostomidae	Creek Chubsucker	*Erimyzon oblongus*	4	Cool-Warm
Catostomidae	Northern Hog Sucker	*Hypentelium nigricans*	3	Cool
Catostomidae	Smallmouth Buffalo	*Ictiobus bubalus*	5	Warm

Family	Common Name	Scientific Name	Numeric Value	Preference
Catostomidae	Bigmouth Buffalo	*Ictiobus cyprinellus*	5	Warm
Catostomidae	Silver Redhorse	*Moxostoma anisurum*	4	Cool-Warm
Catostomidae	Smallmouth Redhorse	*Moxostoma breviceps*	4	Cool-Warm
Catostomidae	River Redhorse	*Moxostoma carinatum*	4	Cool-Warm
Catostomidae	Black Redhorse	*Moxostoma duquesnei*	4	Cool-Warm
Catostomidae	Golden Redhorse	*Moxostoma erythrurum*	4	Cool-Warm
Catostomidae	Shorthead Redhorse	*Moxostoma macrolepidotum*	4	Cool-Warm
Ictaluridae	Black Bullhead	*Ameiurus melas*	5	Warm
Ictaluridae	Yellow Bullhead	*Ameiurus natalis*	4	Cool-Warm
Ictaluridae	Brown Bullhead	*Ameiurus nebulosus*	4	Cool-Warm
Ictaluridae	Channel Catfish	*Ictalurus punctatus*	5	Warm
Ictaluridae	Stonecat	*Noturus flavus*	4	Cool-Warm
Ictaluridae	Tadpole Madtom	*Noturus gyrinus*	5	Warm
Ictaluridae	Margined Madtom	*Noturus insignis*	4	Cool-Warm
Ictaluridae	Brindled Madtom	*Noturus miurus*	4	Cool-Warm
Ictaluridae	Flathead Catfish	*Pylodictis olivaris*	5	Warm
Osmeridae	Rainbow Smelt	*Osmerus mordax*	1	Cold
Salmonidae	Cisco	*Coregonus artedi*	1	Cold
Salmonidae	Lake Whitefish	*Coregonus clupeaformis*	1	Cold
Salmonidae	Pink Salmon	*Oncorhynchus gorbuscha*	1	Cold
Salmonidae	Coho Salmon	*Oncorhynchus kisutch*	1	Cold
Salmonidae	Hybrid Golden Trout	*Oncorhynchus mykiss (hybrid)*	1	Cold
Salmonidae	Rainbow Trout	*Oncorhynchus mykiss*	1	Cold
Salmonidae	Steelhead	*Oncorhynchus mykiss(steelhead)*	1	Cold
Salmonidae	Chinook Salmon	*Oncorhynchus tshawytscha*	1	Cold
Salmonidae	Brown Trout	*Salmo trutta*	2	Cold-Cool
Salmonidae	Hybrid Tiger Trout	*Salvelinus fontinalis x Salmo trutta*	1	Cold
Salmonidae	Brook Trout	*Salvelinus fontinalis*	1	Cold
Salmonidae	Lake Trout	*Salvelinus namaycush*	1	Cold
Esocidae	Redfin Pickerel	*Esox americanus americanus*	4	Cool-Warm

Family	Common Name	Scientific Name	Numeric Value	Preference
Esocidae	Grass Pickerel	*Esox americanus vermiculatus*	4	Cool-Warm
Esocidae	Northern Pike	*Esox Lucius*	4	Cool-Warm
Esocidae	Muskellunge	*Esox masquinongy*	4	Cool-Warm
Esocidae	Chain Pickerel	*Esox niger*	4	Cool-Warm
Umbridae	Central Mudminnow	*Umbra limi*	4	Cool-Warm
Percopsidae	Trout Perch	*Percopsis omiscomaycus*	1	Cold
Gadidae	Burbot	*Lota lota*	2	Cold-Cool
Atherinidae	Brook Silverside	*Labidesthes sicculus*	5	Warm
Fundulidae	Eastern Banded Killifish	*Fundulus diaphanus diaphanous*	5	Warm
Fundulidae	Western Banded Killifish	*Fundulus diaphanus menoma*	5	Warm
Fundulidae	Mummichog	*Fundulus heteroclitus*	5	Warm
Poeciliidae	Eastern Mosquitofish	*Gambusia holbrooki*	5	Warm
Belonidae	Atlantic Needlefish	*Strongylura marina*	5	Warm
Gasterosteidae	Fourspine Stickleback	*Apeltes quadracus*	1	Cold
Gasterosteidae	Brook Stickleback	*Culaea inconstans*	3	Cool
Gasterosteidae	Threespine Stickleback	*Gasterosteus aculeatus*	1	Cold
Gasterosteidae	Blackspotted Stickleback	*Gasterosteus wheatlandi*	1	Cold
Gasterosteidae	Ninespine Stickleback	*Pungitius pungitius*	1	Cold
Cottidae	Mottled Sculpin	*Cottus bairdii*	1	Cold
Cottidae	Blue Ridge Sculpin	*Cottus caeruleomentum*	1	Cold
Cottidae	Slimy Sculpin	*Cottus cognatus*	1	Cold
Cottidae	Potomac Sculpin	*Cottus girardi*	2	Cold-Cool
Cottidae	Spoonhead Sculpin	*Cottus ricei*	1	Cold
Cottidae	Deepwater Sculpin	*Myoxocephalus thompsoni*	1	Cold
Cottidae	Unidentified sculpin	*Unidentified Cottus*	1	Cold
Moronidae	White Perch	*Morone Americana*	5	Warm
Moronidae	White Bass	*Morone chrysops*	5	Warm
Moronidae	White x Striped bass	*Morone chrysops x saxatalis*	4	Cool-Warm
Moronidae	Striped Bass	*Morone saxatilis*	4	Cool-Warm
Centrarchidae	Rock Bass	*Ambloplites rupestris*	4	Cool-Warm

Family	Common Name	Scientific Name	Numeric Value	Preference
Centrarchidae	Redbreast Sunfish	*Lepomis auratus*	4	Cool-Warm
Centrarchidae	Green Sunfish	*Lepomis cyanellus*	5	Warm
Centrarchidae	Pumpkinseed	*Lepomis gibbosus*	4	Cool-Warm
Centrarchidae	Warmouth	*Lepomis gulosus*	5	Warm
Centrarchidae	Orangespotted Sunfish	*Lepomis humilis*	5	Warm
Centrarchidae	Bluegill	*Lepomis macrochirus*	5	Warm
Centrarchidae	Longear Sunfish	*Lepomis megalotis*	4	Cool-Warm
Centrarchidae	Redear Sunfish	*Lepomis microlophus*	5	Warm
Centrarchidae	Smallmouth Bass	*Micropterus dolomieu*	4	Cool-Warm
Centrarchidae	Spotted Bass	*Micropterus punctulatus*	4	Cool-Warm
Centrarchidae	Largemouth Bass	*Micropterus salmoides*	5	Warm
Centrarchidae	White Crappie	*Pomoxis annularis*	5	Warm
Centrarchidae	Black Crappie	*Pomoxis nigromaculatus*	5	Warm
Percidae	Greenside Darter	*Etheostoma blennioides*	4	Cool-Warm
Percidae	Rainbow Darter	*Etheostoma caeruleum*	4	Cool-Warm
Percidae	Iowa Darter	*Etheostoma exile*	3	Cool
Percidae	Fantail Darter	*Etheostoma flabellare*	3	Cool
Percidae	Johnny Darter	*Etheostoma nigrum*	4	Cool-Warm
Percidae	Tessellated Darter	*Etheostoma olmstedi*	3	Cool
Percidae	Tippecanoe Darter	*Etheostoma tippecanoe*	3	Cool
Percidae	Variegate Darter	*Etheostoma variatum*	3	Cool
Percidae	Banded Darter	*Etheostoma zonale*	4	Cool-Warm
Percidae	Ruffe	*Gymnocephalus cernuus*	3	Cool
Percidae	Yellow Perch	*Perca flavescens*	3	Cool
Percidae	Cheseapeake Logperch	*Percina bimaculata*	4	Cool-Warm
Percidae	Logperch	*Percina caprodes*	4	Cool-Warm
Percidae	Channel Darter	*Percina copelandi*	4	Cool-Warm
Percidae	Gilt Darter	*Percina evides*	3	Cool
Percidae	Longhead Darter	*Percina macrocephala*	4	Cool-Warm
Percidae	Blackside Darter	*Percina maculata*	3	Cool
Percidae	Shield Darter	*Percina peltata*	3	Cool
Percidae	River Darter	*Percina shumardi*	4	Cool-Warm
Percidae	Sauger	*Sander canadensis*	3	Cool
Percidae	Saugeye	*Sander canadensis x vitreus*	3	Cool

Family	Common Name	Scientific Name	Numeric Value	Preference
Percidae	Walleye	*Sander vitreus*	3	Cool
Sciaenidae	Freshwater Drum	*Aplodinotus grunniens*	5	Warm
Gobiidae	Round Goby	*Neogobius melanostomus*	3	Cool