STREAM FISH ASSEMBLAGE ASSESSMENT METHOD

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

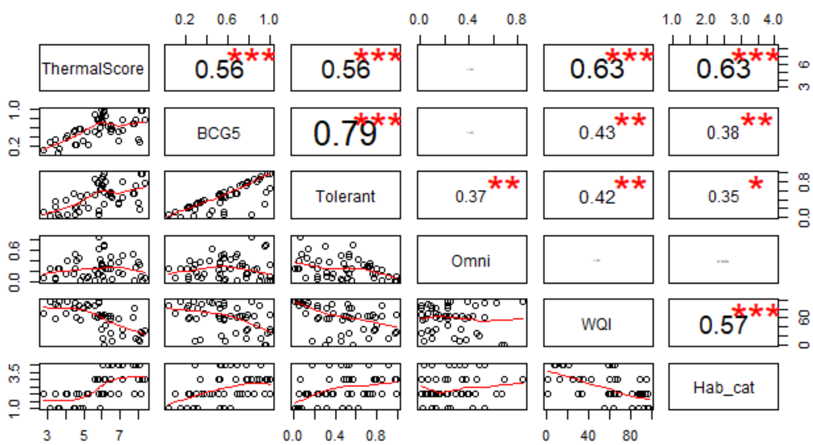
2020

INTRODUCTION

This assessment method is designed to make categorical aquatic life use (ALU) assessment determinations using fish assemblage data compiled from the standardized DEP fish collection protocol (Wertz 2017), across Pennsylvania's lotic surface waters. Using biological indicators to assess attainment of water quality standards (WQS) (more specifically; aquatic life-based WQS) is considered a "core indicator" of attainment (USEPA, 2002). Furthermore, bioassessment methods are considered "translators" of narrative criteria (USEPA, 2002); for Pennsylvania specifically;

> ***25 Pa. Code § 93.6****. General water quality criteria*
>
> *(a) Water may not contain substances attributable to point or nonpoint source discharges in concentration or amounts sufficient to be inimical or harmful to the water uses to be protected or to human, animal, plant or aquatic life.*
>
> *(b) In addition to other substances listed within or addressed by this chapter, specific substances to be controlled include, but are not limited to, floating materials, oil, grease, scum and substances that produce color, tastes, odors, turbidity or settle to form deposits.*

This assessment method is based on the development of a thermal fish index (TFI; Wertz 2019), which followed methods commonly used to develop traditional bioassessment tools (e.g., Barbour et al. 1999). The development differed, in that, classifications not only included stream types such as; freestone (FS) and limestone (LS), but also included longitudinal drainage area groups (DAGs). The resulting classification schema is similar to recent classification studies of fish distribution conducted by Olivero et al. (2015) and Troia and McManamay (2019). The DAG serves as the final classification group, wherein assessments can be made. Conceptual frameworks similar to DAGs are commonly utilized, as longitudinal (e.g., cold water vs. warm water, or headwater stream vs. large river) bioassessments have been investigated and classified for separate assessments using multi-metric approaches for both fish and macroinvertebrates (Lyons et al. 1996, Langdon 2001, Lyons et al. 2001, Hughes et al. 2004, Shull and Lookenbill 2017). Additional techniques to address longitudinal effects have been employed with traditional multi-metric indices (MMI) that scale metrics based on longitude. An example of this scaling would be maximum species richness (MSR) levels that can be used to standardize expected richness depending on stream size or zoogeography (Fausch et al. 1984). The technique of standardizing MMI scores, standardizes the measure of condition (good vs. poor) along a longitudinal gradient but in doing so, simultaneously reduces meaningful interpretation of the longitudinal effect, unless deconstructed to individual metrics. The intent here is not to discredit traditional MMIs but to provide insight into differences in strengths and weakness associated with each. As a stand-alone metric the TFI is: 1) statistically capable of communicating anthropogenic stress along a meaningful gradient of stream type and longitude, and 2) able to numerically characterize assemblages parallel to ALU definitions. However, the TFI does not discriminate or communicate other ecological factors that are typically conveyed in traditional MMIs. These factors include, but are not limited to: species richness, native status, reproductive strategies, or feeding guilds that are commonly addressed in MMIs. From a comparative perspective, a TFI-based assessment (TFI-BASS) may appear to be quite simple in design. In reality, a TFI-BASS should be viewed as a comprehensive assessment, in that: 1) all species and individuals within the assemblage are provided equal consideration based on relative abundance, 2) can be applied uniformly across the State, basins, or ecoregions, 3) has an ecologically meaningful output of assemblage thermal class (cold vs. warm; as opposed to a purely statistically-derived construct), and 4) thermal preferences exhibit collinearity with other tolerances (habitat and water quality, as evidenced herein) and traits. For example, during preliminary data exploration the TFI exhibited collinearity with traditional metrics of biological condition gradient (BCG; Davies and Jackson 2006), tolerance, and feeding guilds; while responding as well as or better than traditional metrics to abiotic stress (Figure 1).



P=*< 0.05, **<0.01, ***<0.001

Figure 1. 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, percent omnivorous (Omni), water quality index (WQI) and habitat (Habcat) in FS<40 streams.

The underlying concepts of the TFI provide valuable insight into the use of traditional metrics, notably; species richness. The TFI-BASS can be viewed as an observed vs. expected assessment, where the observed thermal assemblage is compared to the expected for each DAG. The shift from cold water assemblage (CWA) to transitional assemblage (TSA) to warm water assemblage (WWA) generally includes the initial displacement and eventual replacement of cold water species (Dunham et al. 2002, Troia et al. 2015). The natural transition from CWA (low species richness) to TSA (increasing species richness) is considered ecologically important. Conversely, the stress-induced transition from CWA to TSA may exhibit similar species richness to a natural TSA, however, in doing so spatially condenses the CWA (Figure 2; Wertz 2019).

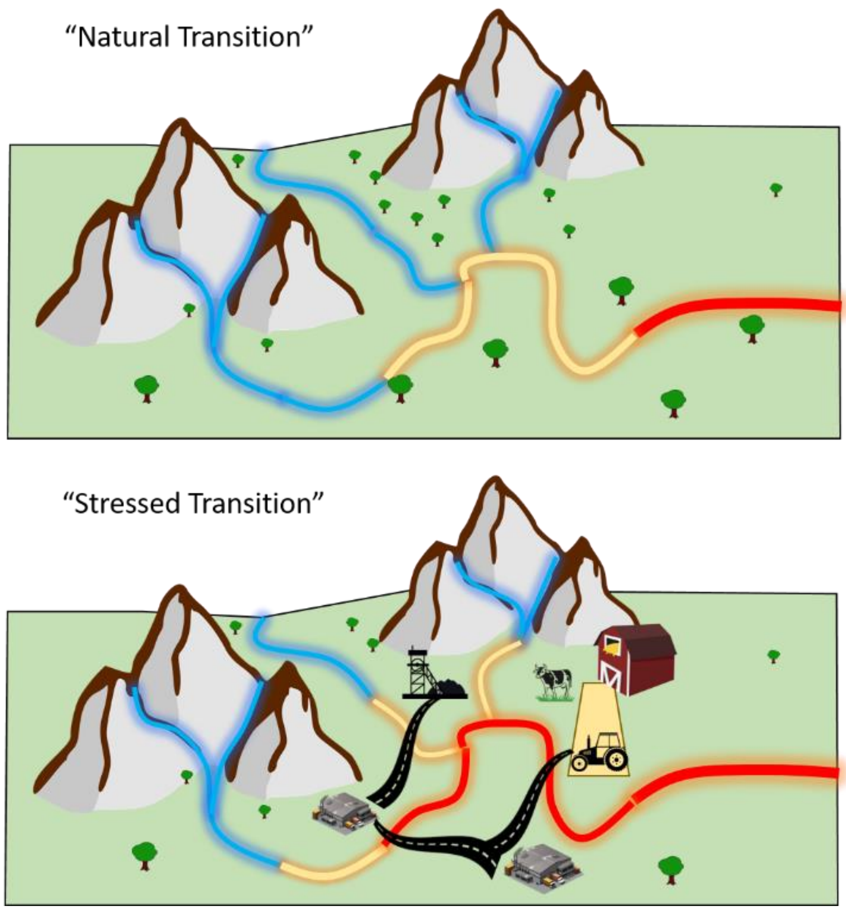


Figure 2. Theoretical example of natural longitudinal transition areas versus stress induced fish assemblage transitions. With applied stress to a cold water assemblage (CWA; blue), the CWA reduces, the transitional assemblage (TSA; yellow) is shifted upstream and the warm water assemblage (WWA; red)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

In this example, a lower species richness in the natural CWA is desired over increased species richness of stress-induced TSA. Therefore, blanket statements of condition based solely on species richness without an "expected" condition, should be avoided.

COLLECTION METHODS

When using the TFI for ALU assessments, the strict use of DEP collection protocols is necessary to making assessment determinations. Fish assemblage data not collected using DEP collection protocol are readily accepted by DEP but may be qualified based on the level of quality assurance and representativeness; following methods for Outside Data outlined in Chapter 5 of the assessment book (Walters and Pulket 2017). Herein, the site selection process relating to representativeness, directly addressed in the data collection protocol for fishes (Wertz 2017), is considered critical as site selection will influence the TFI score. Additionally, knowledge of the assessment method, stream types and appropriate DAGs should be considered in the site selection process; a desktop and field reconnaissance should be conducted prior to sampling to determine the appropriate DAG. To determine the appropriate DAG two pieces of information are needed; stream type (LS or FS) and drainage area. Stream type is determined by the density of sinkholes in the upstream catchment area, where ≥ 0.03 sinkholes/km^2 is used as the inclusive criterion for limestone (karst) streams. This density is achieved by first creating a polygon of the upstream catchment area of the sample site. The second piece of data needed is the sinkhole locations, specifically the “Digital data set of mapped karst features in south-central and southeastern Pennsylvania” (DCNR 2007). Finally, the number of sinkholes within the catchment (km^2) polygon can be summarized as a density, following:

$$Sinkhole\ Density = \frac{\#\ Sinkholes}{Catchment\ area}$$

Desktop reconnaissance should be followed by a field reconnaissance to confirm DAG. The field reconnaissance may reveal additional information that may be necessary to correctly classify DAG. For example, it is likely that small catchments within a larger karst system may not have any sinkholes measurable in the upstream catchment. This occurs by having an unmeasurable (without undertaking complex tracer studies) underground springshed that is larger than the measurable surface watershed. In this situation it would be most appropriate to classify the stream as LS. Additionally, proximal tributaries near the sample site should be evaluated for their influence on not only the representativeness of the site but also the DAG. For example, the fish collection method (Wertz 2017) describes the potential for unrepresentative samples as a result of being too close to a proximal tributary or mouth. This effect becomes even more important when making TFI-BASS determinations as the DAG may change drastically just downstream of a nearby tributary by increasing the upstream catchment area. Simultaneously, the increase in catchment area may be aggravated if there are distinct temperature influences from the tributary, lowering the TFI score while increasing the DAG (and impairment threshold). This example reinforces the importance of proper site selection and representativeness for not only conducting fish assemblage surveys but also for making assessment determinations.

Prior knowledge of stream type and drainage size should also reduce complexities that may arise from a continuous metric (TFI) being nested within a hierarchical construct (DAG). For example, a freestone stream with a 39km^2 drainage area (DAG=FS<40) has an impairment threshold of 4.8. However, a freestone stream with only 2km^2 larger (41km^2) drainage area (DAG=FS<150) has an impairment threshold of 6.0 (Figure 3). It would be unrealistic to expect the fish assemblage to make such a radical transition in such a short distance. Therefore, a 10% buffer is placed on each DAG transition to minimize the effect of drastic transitions, hereafter referred to as the “grey zone” (Figure 4).

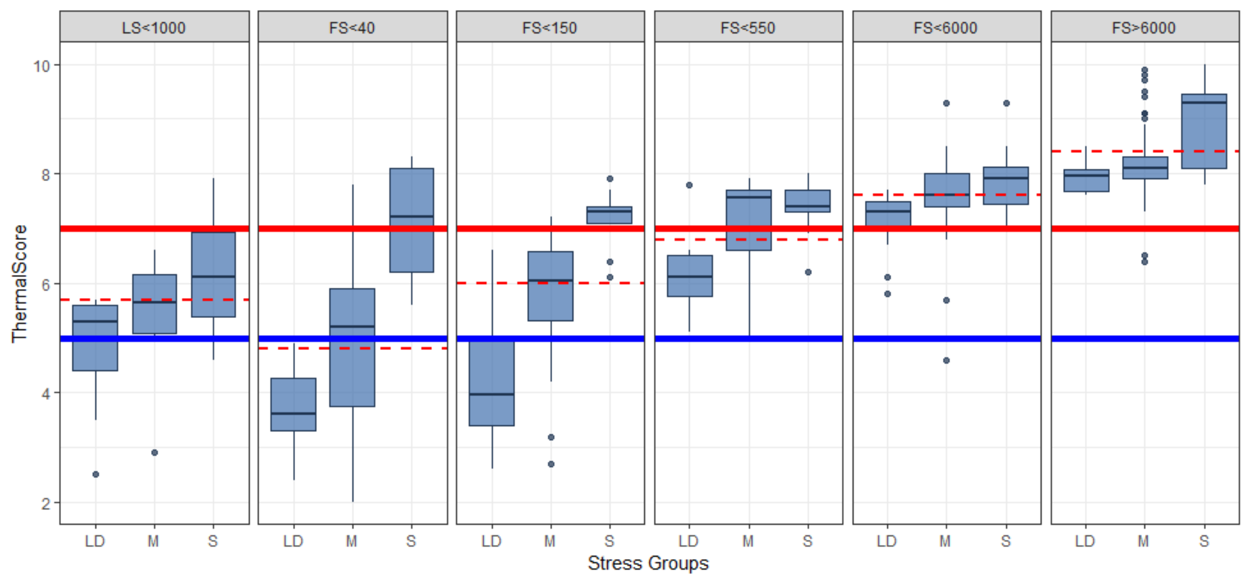


Figure 3. Boxplot of the final limestone (LS) and freestone (FS) drainage area groups (DAGs) (upper km^2 range). Stress groups are denoted as; Least Disturbed (LD), Moderate (M) and Stressed (S). Dotted red lines represents the 95th percentile of least disturbed sites signifying the impairment threshold. The solid blue line represents the upper limit for cold water assemblage and the solid red line represents the lower limit for warm water assemblage, transitional assemblage range is between.

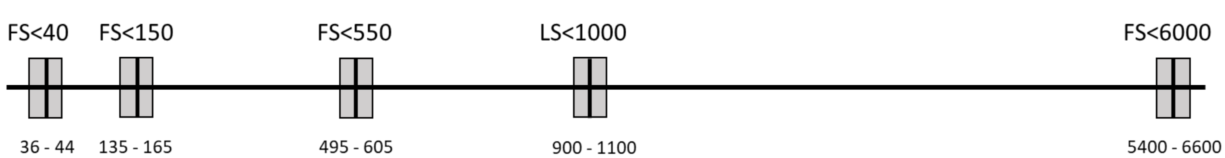


Figure 4. Drainage area groups and corresponding grey zone drainage area range (km^2); illustration not to scale.

If a site is in a grey zone, two options are available: 1) move the site either upstream or downstream (outside of grey zone), or 2) add a second site in either direction (outside of grey zone) to serve as supplementary evidence. Of the two choices, the latter is preferred if feasible. For example, Loyalhanna Creek (a FS stream) had two fish surveys; one upstream and one downstream of Latrobe, PA. The upstream site had a drainage area of 503km^2 (FS<550), a TFI score of 6.7, and was considered attaining. The downstream site had a drainage area of 572km^2 (FS<6000), a TFI score of 7.2 (TFI increase of 0.5 compared to the upstream site), and was also considered attaining. In this example when the 10% buffer rule is applied, both sites should be either moved or supplemented. In certain cases, there may be minimal flexibility on moving a site based on access or representativeness, and a supplemental site may be desired. In this specific example the upstream site could be moved or supplemented to a site upstream in an area ≤495km^2. The downstream site could be moved or supplemented to a site further downstream in an area 605-900km^2. In all grey zone cases, supplemental sites are preferred over moving to reduce bias and longitudinal data gaps. It should be noted that while ALU assessments may be conducted on streams that inherently bracket influences (e.g., changes in land use or land cover) they should not be confused with cause and effect surveys. Herein, cause and effect surveys are used to evaluate local scale impacts and are generally not considered representative of overall waterbody condition; the intent of categorical ALU assessments.

HABITAT CONSIDERATIONS

The TFI responded significantly to habitat alterations as habitat stress was an integral part of the developmental stressor gradient (Wertz 2019; Figure 5). This response to habitat needs to be further discussed, as assessment determinations may also be affected. Herein, the TFI is likely not capable of discriminating between anthropogenically modified habitat and naturally unsuitable habitat conditions. For example, it is to be expected that samples collected in very low gradient areas, or in and around natural marshes and wetlands may have higher TFI scores than similarly-sized high gradient, cobble-dominated streams. This concept was evident in the results of slope being an important secondary factor (a reduced sample size precluded its analysis) from boosted regression tree outputs during development (Wertz 2019). To this end, if a sample exceeds the appropriate threshold and unique habitat conditions are suspected, further investigations should be conducted before assessments are made. Furthermore, if habitat conditions are thought to be preventing or precluding attainment, an evaluation of these conditions should be made. As site-specific or general habitat conditions are identified, their inclusion into future calibration events will likely add precision and accuracy. For example, some streams will naturally flow subsurface for some distance throughout all/portions of the year. These general conditions may prevent the attainment of categorical ALU upstream by reducing fish migration, and once evaluated across numerous streams, could serve as a blanket template for assessments across streams of similar description or classification.

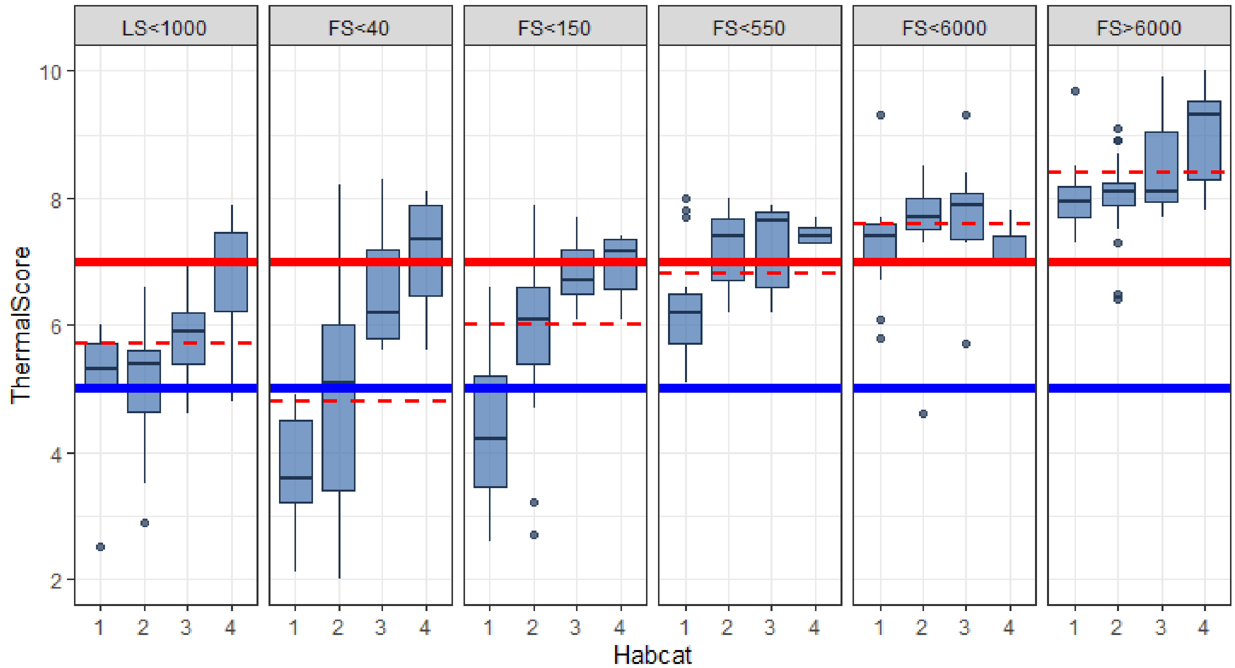


Figure 5. Boxplot of the final limestone (LS) and freestone (FS) drainage area groups (DAGs) (upper km^2 range). Habcat groups 1-4 are on a gradient of good to poor respectively. Dotted red lines represents the 95th percentile of least disturbed sites signifying the impairment threshold. The solid blue line represents the upper limit for cold water assemblage and the solid red line represents the lower limit for warm water assemblage, transitional assemblage range is between these lines.

SAMPLE CONSIDERATIONS

Once the sample has been processed in accordance with collection methods (Wertz 2017) the final assemblage data needs to be reviewed for representativeness before making assessment determinations. The sample is first evaluated for the number of individuals and species present. If fewer than 50 individuals are in the sample the site should be evaluated for: 1) representative sampling as outlined in the collection methods (Wertz 2017), 2) toxic conditions, and 3) near-sterile conditions. If collection methods are suspected to be the cause of the low numbers, the site should be resampled as confirmation. If toxic conditions are suspected, and supporting water quality evidence is present, a biological impairment is justified. If near-sterile conditions are expected the TFI should generally be considered representative of site conditions and assessments can be pursued. The sample should then be evaluated for the number of species present. A sample represented by only one species (or family for salmonids) is generally considered abnormal even in naturally-depauperate headwater streams. In many headwater streams, trout will be the only species (or family) present in a sample. This effect will often be an indicator of acidified conditions as many of the species found in cold water environments have a lower tolerance to acidity (and acid-effects) than trout (Johnson et al. 1987, Baker et al. 1996). This salmonid-dominated CWA has been an indicator of acidified conditions not only in the Northeastern U.S. (Baker et al. 1996) but across multiple continents (Schofield 1976). Consequently, when salmonids (one or more species) represent the entire assemblage, the acid precipitation source and cause determination method (Friday and Shull 2017) should be investigated; and, the sample could be considered a biological impairment. If the sample is represented by only one species (non-salmonid) and the number of individuals is ≥50 the TFI should generally be considered representative of site conditions and assessments can be pursued.

When reporting the TFI scores, if identification of individuals is not at the species-level (e.g., hybrids, family or genus level) or contains a species without an associated thermal preference score, the percentage of individuals not used in the TFI should be noted. Furthermore, when making an assessment with these data, this percentage should be taken into consideration. In the developmental analysis, 10% was used as a criterion (Wertz 2019) which can be applied as a general rule-of-thumb hereafter. For example, if 11% of the sample is not accounted for in the TFI, the decision to make an assessment may still be applicable but should be thoroughly evaluated and reported. An example of a thorough evaluation could include but is not limited to: 1) hybrid individuals represented by suspected parental species, 2) unidentified juveniles represented by suspected adults, and 3) TFI scores similar across a higher taxonomic level. For example, some species of sculpins may be difficult to identify to species-level increasing the percentage of individuals without a TFI score. However, since all sculpins (family/genus level) found in Pennsylvania prefer cold (or cold-cool for Potomac Sculpin, *Cottus girardi*) habitats the TFI score with >10% missing, could still be considered representative. Alternatively, minnows range from cold-cool to warm habitat preference; where the 10% rule is violated by “unidentified minnows” the TFI is likely not representative.

Once the sample and corresponding TFI score are evaluated and considered representative of the waterbody (number of individuals/species in the sample, habitat representative of general stream conditions, TFI calculated from ≥ 90% of assemblage) the appropriate DAG determinations can be pursued. Two pieces of information are necessary to determine the appropriate DAG; stream type and drainage area. The stream type is identified by having evaluated, through a desktop analysis, sinkhole densities in the upstream catchment, if not directly obvious. Once stream type is established, the appropriate DAGs and corresponding impairment threshold is used to measure attainment of categorical ALU, for fishes. Herein, impairment thresholds per DAG are considered numerical interpretations of the narrative criteria at 25 Pa. Code §

93.6(a) for making categorical ALU assessment determinations (USEPA 1990). Impairment thresholds should not be confused with thermal assemblage classification thresholds (i.e. CWA=≤5.0, TSA= ≤7.0, WWA=>7.0). Assessment determinations can be aided using a flow chart (Figure 6). This final TFI-BASS provides the first fish-based assessment tool for assessing aquatic life use across lotic waterbodies in Pennsylvania.

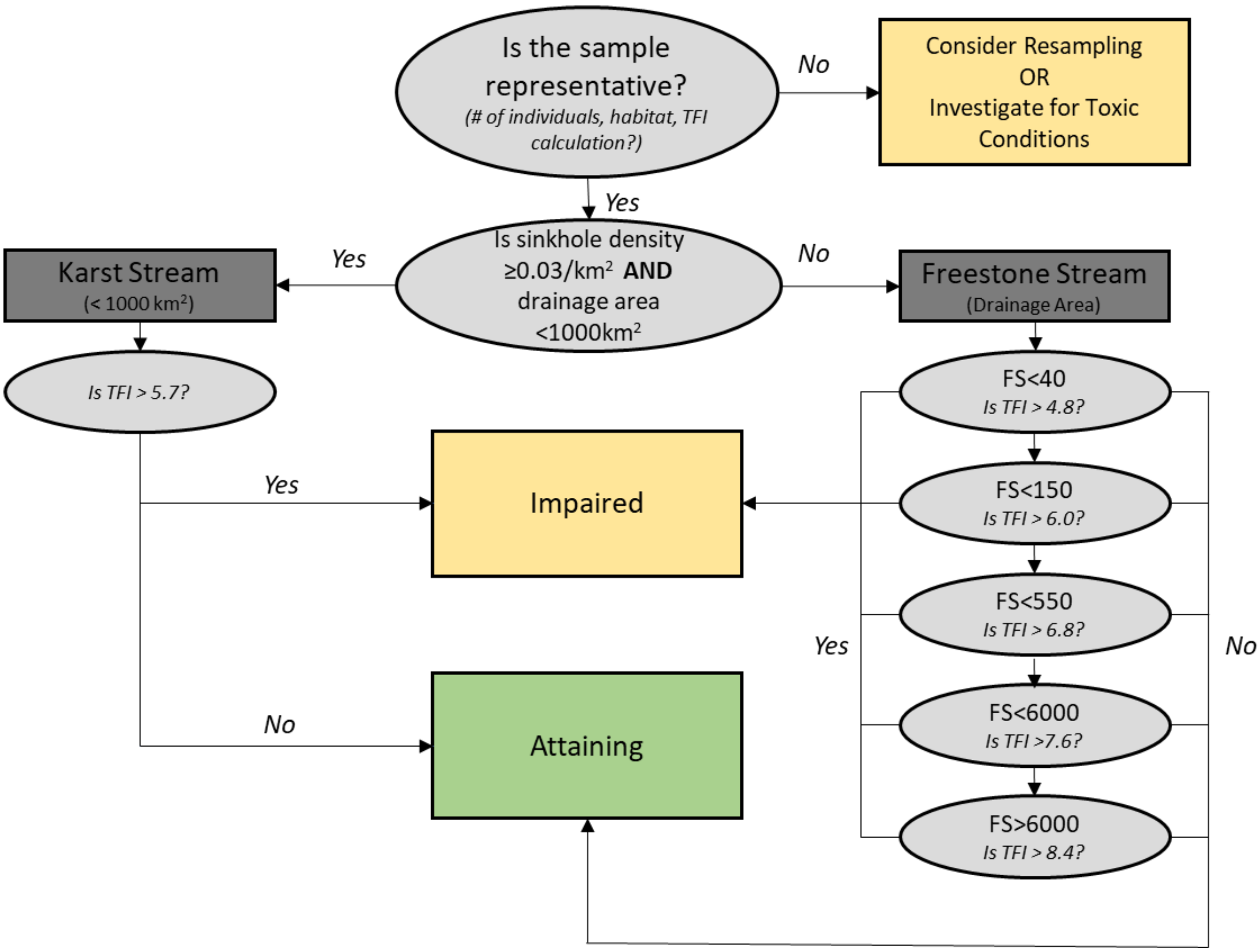


Figure 6. Flowchart to aide in aquatic life use assessments based on Thermal Fish Index (TFI) scores.

SOURCE/CAUSE DISCUSSION

The TFI-BASS is a fish-assemblage-based biological assessment that was developed from, and is considered calibrated to: water quality, habitat and temperature stress. Therefore, it is important to note that source/cause investigations resulting from TFI-BASS impairments should not be limited to temperature but should also extend to water quality and habitat. By identifying stressors that elicit significant metric (TFI) responses, causal inferences can be better focused. Herein, a general weight-of-evidence approach following Walters (2017a) should be conducted while focusing specifically on the three-key variables (Figure 7). As described in Wertz (2019);

"…the effects of multiple stressors will be synergistic, antagonistic or additive to the TFI scores. For example, as water quality is reduced by agricultural activities and loss of riparian areas, changes to instream habitat and temperature will likely parallel, having a dramatic 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."

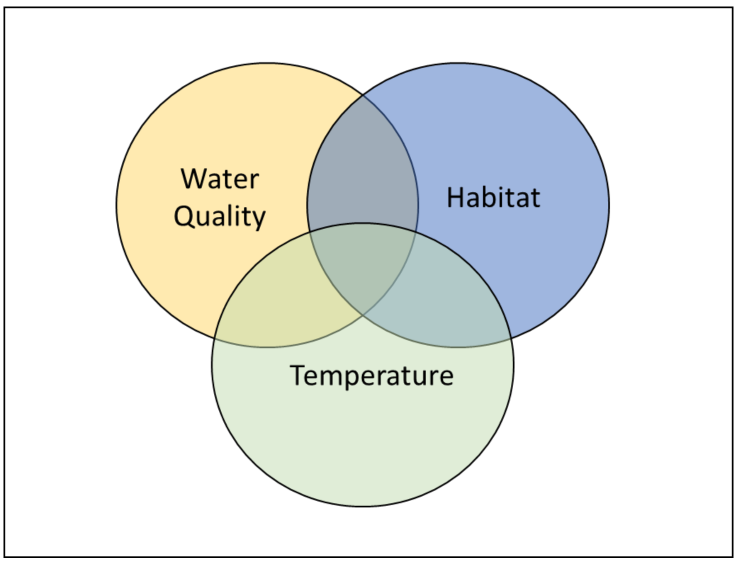


Figure 7. Venn diagram illustrating the three-key variables used in developing and calibrating the thermal fish index-based assessment.

The source/cause investigation should be completed according to general source and cause determination methods (Walters 2017a) and begin with the most-obvious stressor(s) group of the three based on prior knowledge gained during site reconnaissance. Habitat stress will generally be the most obvious from a field investigation and subsequent stream habitat data collection (Shaw 2002) and physical habitat assessments (Walters 2017b) that may characterize a habitat modification cause of impairment. Generally, increases in sedimentation, siltation, embeddedness or impounding should be conveyed, as these were identified as important stress variables during development (Wertz 2019). Water quality and temperature data may not be obvious from a field investigation and may require additional data collection and evaluation.

To appropriately investigate and document a thermal modification cause, two major components are needed; a biological response and temperature data. The biological response is measured with a TFI-BASS impairment. The temperature data needed to support a thermal modification cause should be collected as a comparative study that includes both impacted and unimpacted areas. By demonstrating a significant difference in temperature between ambient conditions and impacted areas the thermal modification cause is supported. Ambient temperature conditions should be measured in areas that spatially and temporally represent the overall waterbody, meaning they should not be taken downstream of nearby tributaries that may influence the "reference condition" and are measured at times when differences are likely to occur. A thorough investigation of ambient conditions includes an average temperature of multiple samples that spatially represent the unimpacted area. Fish data should also be collected in the unaffected areas to bracket temperature stressors, but do not necessarily need to demonstrate attainment to apply a thermal modification cause to the impacted reach. For example, if a thermal discharge exists in a stream already stressed by siltation the "temperature reference" site may be impaired by siltation and the temperature impacted site may be impaired by siltation and thermal modification. Due to the highly-mobile nature of fishes, communities in close proximity to thermally modified areas will likely be affected. For example, thermal plumes exclude "cooler" fishes and provide areas suitable for reproduction of "warmer" fishes that will inevitably radiate to areas outside of the direct plume. This situation is dynamic and becomes challenging to directly quantify. To address this situation, if the "temperature reference" sites are impaired by the TFI-BASS with no evidence of cause, aside from proximity to temperature impacts, "Thermal Modifications" cause (Appendix A; Shull and Pulket 2018) is justified.

To appropriately investigate and document a water quality or pollutant cause, two major components are needed; a biological response and pollutant data. The biological response is measured with a TFI-BASS impairment. The pollutant data needed to support a specific pollutant cause must be evaluated from a spatiotemporal perspective. In other words, the causal pollutant may not be directly obvious from a review of current data alone and should be augmented by historical data. Multiple spatial and temporal environmental variables must be considered, as recovery of fish assemblages is dependent upon multiscalar recolonization potential (Poff 1997). For example, where small scale disturbance occurs and localized refugia (from a pollutant) is present, recovery may occur rapidly. Where large scale disturbance occurs and no refugia is present, recovery may occur over years or decades (Detenbeck et al.1992). Where spatiotemporal pollutant data is available, and values are elevated, causal determinations will be fairly obvious. Where spatiotemporal pollutant data is not readily available or no links between pollutant and impairment are made, the cause of "Cause Unknown" (Appendix A; Shull and Pulket 2018) is justified.

NATURAL VARIATION

As sites are resampled through time, TFI scores will likely change as well. Understanding whether these changes in TFI scores are actually measuring degradation (increasing TFI scores), improvement (decreasing TFI scores) or are just natural variations, is important. To provide insight into changes not associated with natural variation precision estimates from repeated sites were evaluated. Precision measurements using coefficient of variation (CV) of the TFI score across replicate sites indicated that natural variation is low. The average CV across all DAGs was 4.3%, which translates to a TFI score ± 0.3 (Wertz 2019); well under recommended threshold ranges of 10-15% (Stribling et al. 2008). The highest CV (8.8%) was noted in the FS<150 DAG, this is likely caused by longitudinal shifts of assemblages that may be seasonally affected (i.e. cooler assemblages may retreat upstream as summer temperature increases). Herein, the TFI can be used as a tool to measure trends of improving/degrading conditions through time. As follow-up investigations are completed, TFI values greater than the average precision estimates per DAG (Table 1) are considered outside of the range of natural variation and are likely caused by changing conditions.

Table 1. Precision estimates using coefficient of variation (CV) and corresponding thermal fish index scores for repeated sites within each drainage area group (DAG).

DAG	CV %	TFI ±	*N*
LS<1000	4.0	0.3	16
FS<40	1.8	0.1	11
FS<150	8.8	0.7	59
FS<550	3.2	0.3	61
FS<6000	4.5	0.4	39
FS>6000	3.3	0.3	178

With appropriate implementation of this TFI-BASS method, DEP can use fish to: identify and list impaired waters, direct management strategies through source/cause evaluations, document a change in conditions through the implementation of cause and effect survey design, and measure incremental progress.

LITERATURE CITED

Baker, J.P., Van Sickle, J., Gagen, C.J., DeWalle, D.R., Sharpe, W.E., Carline, R.F., Baldigo, B.P., Murdoch, P.S., Bath, D.W., Krester, W.A. and Simonin, H.A., 1996. Episodic acidification of small streams in the northeastern United States: effects on fish populations. Ecological Applications, 6(2), pp.422-437.

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

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), pp.1251–1266.

DCNR, 2007. Digital data set of mapped karst features in south-central and southeastern Pennsylvania. Edition 1.1. http://www.pasda.psu.edu/

Detenbeck, N.E., DeVore, P.W., Niemi, G.J. and Lima, A., 1992. Recovery of temperate-stream fish communities from disturbance: a review of case studies and synthesis of theory. Environmental management, 16(1), p.33.

Dunham, J.B., Adams, S.B., Schroeter, R.E. and Novinger, D.C., 2002. Alien invasions in aquatic ecosystems: toward an understanding of brook trout invasions and potential impacts on inland cutthroat trout in western North America. Reviews in Fish Biology and Fisheries, 12(4), pp.373-391.

Fausch, K.D., Karr, J.R. and Yant, P.R., 1984. Regional application of an index of biotic integrity based on stream fish communities. Transactions of the American Fisheries Society, 113(1), pp.39-55.

Friday, M. and Shull, D.R. 2017. Acid precipitation source and cause determination method. Chapter 6, pages 6–10. In Shull, D. R., and M. Pulket. (editors). Assessment Methodology for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

Hughes, R.M., Howlin, S. and Kaufmann, P.R., 2004. A biointegrity index (IBI) for coldwater streams of western Oregon and Washington. *Transactions of the American Fisheries Society*, *133*(6), pp.1497-1515.

Johnson, D.W., Simonin, H.A., Colquhoun, J.R. and Flack, F.M., 1987. In situ toxicity tests of fishes in acid waters. Biogeochemistry, 3(1-3), pp.181-208.

Langdon, R.W., 2001. A preliminary index of biological integrity for fish assemblages of small coldwater streams in Vermont. *Northeastern Naturalist*, *8*(2), pp.219-233.

Lyons, J., Wang, L. and Simonson, T.D., 1996. Development and validation of an index of biotic integrity for coldwater streams in Wisconsin. *North American Journal of Fisheries Management*, *16*(2), pp.241-256.

Lyons, J., Piette, R.R. and Niermeyer, K.W., 2001. Development, validation, and application of a fish-based index of biotic integrity for Wisconsin's large warmwater rivers. *Transactions of the American Fisheries Society*, *130*(6), pp.1077-1094.

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

Poff, N.L., 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological society, 16(2), pp.391-409.

Schofield, C.L., 1976. Acid precipitation: effects on fish. Ambio, pp.228-230.

Shull, D.R. and Lookenbill, M.J., 2017. Assessing the expansion of wadeable benthic macroinvertebrate collection methods in large semiwadeable rivers. Freshwater Science, 36(3), pp.683-691.

Shull, D.R., and M. Pulket. 2018. Assessment Methodology for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

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

Troia, M.J., Denk, M.A. and Gido, K.B., 2015. Temperature-dependent performance as a driver of warm-water fish species replacement along the river continuum. Canadian journal of fisheries and aquatic sciences, 73(3), pp.394-405.

Troia, M.J. and McManamay, R.A., 2019. Biogeographic classification of streams using fish community–and trait–environment relationships. Diversity and Distributions.

USEPA. 1990. Biological Criteria: National Program Guidance For Surface Waters. EPA-440/5-90-004. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 2002. Consolidated assessment and listing methodology: Toward a compendium of best practices. U.S. Environmental Protection Agency, Office of Wetlands, Oceans, and Watersheds. Washington, DC.

Walters, G. 2017a. Source and Cause Determination Methods. Chapter 6, pages 1-5. In Shull, D.R., and M. Pulket. (editors). Assessment Methodology for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

Walters, G. 2017b. Physical Habitat Assessment Method. Chapter 4, pages 1-6. In Shull, D.R., and M. Pulket. (editors). Assessment Methodology for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

Walters, G., and M. Pulket 2017. Assessment Determination and Listing Methods. Chapter 5, pages 1-12. In Shull, D.R., and M. Pulket. (editors). Assessment Methodology for Streams and Rivers. Pennsylvania Department of Environmental Protection Harrisburg, Pennsylvania.

Wertz, T.A. 2017. Fish Data Collection Protocol. Chapter 3, pages 42-59. *In* Shull, D. R., and M. J. Lookenbill. Water quality monitoring protocols for streams and rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

Wertz, T.A. 2019. Technical Development of a Thermal Fish Index. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.