

2020

Recovery of Benthic Macroinvertebrate Communities after Restoration in Two Intensively Mined Watersheds

Brian R. Gordon
brianodrog@gmail.com

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>



Part of the [Terrestrial and Aquatic Ecology Commons](#)

Recommended Citation

Gordon, Brian R., "Recovery of Benthic Macroinvertebrate Communities after Restoration in Two Intensively Mined Watersheds" (2020). *Graduate Theses, Dissertations, and Problem Reports*. 7947.
<https://researchrepository.wvu.edu/etd/7947>

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

Recovery of Benthic Macroinvertebrate Communities after Restoration in Two Intensively Mined
Watersheds

Brian R. Gordon

A Thesis submitted to
The Davis College of Agriculture, Natural Resources and Design
at West Virginia University
in partial fulfillment of the requirements
for the degree of

Master of Science
in
Wildlife and Fisheries Resources

J. Todd Petty, Ph.D., Chair
Eric R. Merriam, Ph.D.
Nicolas P. Zégre, Ph.D.

Wildlife and Fisheries Resources Program
Division of Forestry and Natural Resources

Morgantown, WV

2020

Key words: acid mine drainage (AMD), benthic macroinvertebrates, lime dosing, biological communities,
field experiment, and watershed restoration.

Copyright 2020 Brian R. Gordon

Abstract

Recovery of Benthic Macroinvertebrate Communities after Restoration in Two Intensively Mined Watersheds

Brian R. Gordon

As the body of literature supporting watershed-scale restoration projects continues to grow, there is need for a greater understanding of the underlying mechanisms that drive it. With data spanning ten years capturing pre- and post-restoration benthic macroinvertebrate communities in two distinct watersheds of West Virginia, we seek to contribute to the developing pool of knowledge regarding these dynamics. Sites were laid out prior to remediation in such a manner as to capture the spectrum of water quality conditions present, these are divided into three treatment types: reference, treated, and abandoned mine drainage (AMD). Our objectives were to: 1) determine if previously documented trends of improvement persisted into 2018, 2) explore specific impacts on community structure, and 3) determine if distance to remediation impacted degree of recovery. We hypothesized that: 1) previously measured improvements would continue, 2) watershed-wide communities would exhibit less variation over time as they grew more similar overall to reference quality, and 3) that sites situated farther from active treatment locations would exhibit greater degrees of recovery. In order to provide meaningful and holistic interpretations of the community data collected, we analyzed how site distance from active treatment dosers impacts scores from indices of biotic integrity as well as Bray-Curtis distance coefficients on community similarity. We found that the previously documented recovery was still in evidence, however at a much lesser rate. Contrary to our other hypotheses, watershed-wide communities became generally more diverse over time; and distance to treatment yielded opposite trends in recovery between our two watersheds.

Acknowledgements

I would like to thank my committee, Dr. Todd Petty, Dr. Eric Merriam, and Dr. Nicolas Zégre for their commitment to bettering my understanding of what it means to be a scientist, and helping me grow into that role. Thank you Dr. Petty for granting me the opportunity to pursue knowledge as a member of your laboratory, and for your insight into research, study design, and the finer points of writing. I would be remiss not to give special thanks to Dr. Merriam for his constant encouragement and willingness to consult on all manner of problems, even outside of normal work hours. Gratitude is also due to Donna Hartman for constant support and encouragement. My fellow lab members: Rebecca Long, Kevin Eliason, and Josh Ankeny made work a joy and offered much needed kinship in a new home. Dr. Christopher Lituma deserves recognition for providing a framework for the graduate students of the Department of Forestry & Natural Resources to assemble, grow bonds, and make plans. Finally, I would not have experienced any of this if it were not for the mentorship of Dr. Mel Zimmerman, who first gave me a place in a laboratory, and showed me how fascinating stream ecology can be.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
Introduction	1
Objectives	5
Methods.....	5
Site Selection.....	5
Sampling Techniques	6
Analysis	7
Objective 1: Post-restoration recovery.....	7
Objective 2: Community composition	7
Objective 3: Recovery in relation to treatment distance	9
Results.....	10
Objective 1: Post-restoration recovery.....	10
ANOVA and Tukey’s post-hoc on GLIMPSS scores grouped by treatment.....	10
Objective 2: Community Composition.....	10
PERMANOVA and NMDS on Bray-Curtis dissimilarity matrices of community composition	10
Nested temperature of communities	11
Objective 3: Recovery in relation to treatment distance	12
Community composition and minimum distance to treatment PERMANOVA	12
Community integrity scores and minimum distance to treatment	12
Discussion.....	12
Objective 1: Post-restoration recovery.....	13
Objective 2: Community Composition.....	13
Objective 3: Recovery in relation to treatment distance	15
Implications for future research	16
Figures.....	18
Tables	28
Works Cited.....	30

List of Figures & Tables

Figure 1 Study watersheds.....	6
Figure 2 Box plot of treatment groups	18
Figure 3 NMDS of benthic macroinvertebrate communities.....	18
Figure 4 NMDS of treated communities.....	19
Figure 5 Individual year NMDS of Three Fork communities.....	19
Figure 6 Individual year NMDS of Abrams communities.....	20
Figure 7 Nested graph Three Fork Creek 2008	21
Figure 8 Nested graph Three Fork Creek 2017	22
Figure 9 Nested graph Three Fork Creek 2018	22
Figure 10 Nested graph Three Fork Creek for 2008-2018	23
Figure 11 Nested graph Abrams Creek 2008	24
Figure 12 Nested graph Abrams Creek 2013	24
Figure 13 Nested graph Abrams Creek 2017	25
Figure 14 Nested graph Abrams Creek 2018	25
Figure 15 Nested graph Abrams Creek for 2008-2018	26
Figure 16 GLIMPSS scores plotted against distance to treatment.....	27
Figure 17 Change in GLIMPSS scores plotted against distance to treatment.....	27
Table 1 Water quality summary.....	28-29

Introduction

Mankind's detrimental impacts on surface waters precede even the industrial revolution. Documented comments on severe pollution in the waters of the Eastern United States date at least as early as 1769 (Kauffman Jr., 2010). Tanneries, dyers, and fertilizer on agricultural fields all contributed to water which could not be safely drunk without treatment (Wijbenga and Hutzinger, 1984). With the advent of industry came disturbing new signs of pollution. The Cuyahoga River famously caught fire at least ten times (Stradling and Stradling, 2008) and it was not the only water to burn as a result of chemical pollution. The Chesapeake Bay and other coastal waters began to exhibit "dead zones" – anoxic areas devoid of organisms that need oxygen to survive – due to high nutrients in agricultural runoff (Diaz and Rosenberg, 2008). And metal contamination soon became known for its potential to decimate life in the waters it affects.

Abandoned Mine Drainage (AMD) is a source of metal contamination and subsequent pollution. To explain the impacts of AMD in brief: sulfides are exposed during the mining process, the sulfides form sulfuric acid after reacting with atmospheric air and water, the resulting acidic and oxidized conditions allow metals to leach into water, then as acidity drops these metals may precipitate into a toxic substrate coating the streambed (Hogsden and Harding, 2012). One study testing the comparative effects of aqueous vs precipitate AMD found that aqueous was more damaging, but it should be noted that both presentations are problematic (DeNicola and Stapleton, 2002). These mechanisms overwhelm the tolerance of many species when AMD affects a stream; and this negative effect propagates upwards through the trophic levels, even leaving the stream (Mulvihill et al., 2008).

To summarize: AMD has well documented potential to render water nearly abiotic, and emerging research indicates these effects are not confined to the boundary of the waters affected. While accurate figures on deep mines are elusive, there is an estimated 5,900 km² of surface mining in central Appalachia (Pericak et al., 2018), and approximately 9,600 km of streams impacted by AMD in

Pennsylvania and West Virginia (DeNicola and Stapleton, 2002). However AMD is not a uniquely Appalachian problem, occurring in New Zealand, South Africa, China, and Spain among other places. Watson (2017) calls mining a “crucial” regional issue, referencing another paper dissecting the impacts of mining on streams (Petty et al., 2010).

Considering this long history of poor stewardship over water resources, the 1948 passage of the Federal Pollution Control Act (FPCA) snaps into clear focus as a recent development. The FPCA was eventually updated into the modern day Clean Water Act (CWA), through a series of amendments taking place roughly between 1972 and 1983 which served to give the Act greater regulatory authority. Owing to this legislative framework requiring that waters have a minimal expectancy of “providing for the propagation of fish, shellfish, and wildlife as well as recreation in and on the water” (USEPA, 2017), individuals and corporations were given legislative and financial motivation to begin curtailing these centuries-old grievances. A study published in 2005 concluded that the United States spends over one billion dollars annually on river restoration (Bernhardt et al., 2005). As the study references a constant increase in this value, we can safely conclude that the current cost is even higher. With such a large national cost, one naturally questions: “is it worth it?”. The answer is at least hypothetically yes, one study suggests that nearly half a billion dollars in revenue are generated from nature-based recreation in the Apalachicola River region of Florida per year (Shrestha et al., 2007). Locally, West Virginia reported fishing-related license revenue in excess of 6.5 million for 2018 (WVDNR, 2019). Before coming to a full conclusion though, it is important to consider that the benefits of restoration are utterly contingent on its success. Defining and refining methods for measuring restoration success is a constant drive within the scientific community (DeNicola and Stapleton, 2016; Palmer et al., 2014; Walter et al., 2012; Watson et al., 2017).

In many cases restoration meets the goals it was intended to and the waters upon which it was implemented can again contribute to the region’s ecological and economic health, but this is not always

the case. Palmer et. al. (2014) denounce the popular movement of restoration targeting solely physical parameters, such as channel reconfiguration projects, warning that "...[this] method does not address chemical or biological processes." Buchanan et. al. (2012) highlight the necessity of post-restoration evaluations to further our understanding of how efforts at restoration can be best applied. Add to this call the sobering conclusion that time after restoration is not a reliable predictor of community improvement, and it becomes increasingly apparent that better restoration will stem from deeper understanding of the biological processes at play (Leps et al., 2016).

McClurg et. al. (2007) came to a similar conclusion as Leps whilst studying recovery of streams after remediation from acid-precipitation impacts. They posited that a possible cause for the failure of recovery over time was due to treatment isolation within an unhealthy network, and suggested that improved connectivity to healthier systems may bolster recruitment of pollution-intolerant taxa, a notion shared by Westveer et. al. (2018). Watershed-scale approaches to restoration are arguably still in nascent stages; however, there is a growing body of literature indicating the utility of this method (Bernhardt et al., 2005; McClurg et al., 2007; Watson et al., 2017). This new approach to broad-scale restoration is supported in large part by recent appreciation for the importance of metacommunity dynamics in recovery (Merovich et al., 2013; Merriam and Petty, 2016; Swan and Brown, 2017).

Two recent examples of research detailing recovery of streams after watershed-scale AMD remediation were conducted by West Virginia University. Watson (2017) studied water chemistry, habitat, benthic macroinvertebrates, and fishes before and after implementation of in-stream lime dosers aimed at restoring water chemistry to circumneutral conditions. Watson sampled the Abrams Creek sites three years after implementation and concluded that watershed scale recovery may take more time to become established and evident. Long (2019) continued Watson's work on Abrams Creek and expanded the research to capture restoration efforts in Three Fork Creek as well. Both authors included macroinvertebrate assemblages in their assessments of recovery, but both also noted that

further research was needed in order to explain trends within the recovering benthic communities.

Macroinvertebrates are an important component of most stream quality assessments. This widely accepted protocol makes practical sense: collecting kick samples is generally less labor intensive, less environmentally disruptive, and requires cheaper equipment than many other sampling methods (Barbour et al., 1999). Because they are a lower trophic level, macroinvertebrates can provide insight into fisheries and water quality that may not be currently evident in the fish population due to lag effects (Wallace and Webster, 1996). And transient episodes of toxicity can be evident in macroinvertebrates for longer than more mobile communities such as fish due to their inability to quickly disperse and return (Barbour et al., 1999). Specifically, AMD impacts result in a community shift towards more acid-tolerant genera, as well as changes in macroinvertebrate functional feeding group distributions (Gerhardt et al., 2004).

Both Watson and Long used a standard method of assessing macroinvertebrate samples called an Index of Biotic Integrity (IBI), which assigns value to different taxa and allows researchers to classify stream health based on the species it supports. However, IBIs may fail to capture finer-scale community shifts that do not reflect a change in pollution tolerance but may be ecologically significant nonetheless (Merovich and Petty, 2007). Conducting community similarity and nested temperature analyses should provide more holistic insight to the dynamics occurring in respect to – or else despite – restoration efforts. Nested temperature analysis will facilitate deeper comprehension of underlying watershed community dynamics through illustrating taxa by site type, oriented by decreasing richness.

While some research has shown distance to treatment to be a crucial factor in determining a site's probability of successful recovery (Keener and Sharpe, 2005; McClurg et al., 2007), this subject appears to be underexplored in the literature surrounding restoration science. In continuation of the research started by Watson and upheld by Long, this project will contribute an additional year of benthic macroinvertebrate data to both watersheds, conduct more detailed community analysis, and will test

the hypothesis that distance to treatment may explain the discrepancies in macroinvertebrate recovery previously observed.

Objectives

The objective of this research is to contribute towards a greater understanding of constraints on benthic macroinvertebrate recovery in restored AMD watersheds by building upon the previous studies of Long (2019) and Watson (2017).

My specific objectives are to:

1. Test whether macroinvertebrate communities are continuing to improve 7-8 years post restoration, as defined by biotic indices used by the state of WV
2. Characterize the effect of restoration on genus-level community composition and nestedness.
3. Examine the effect of distance to treatment on macroinvertebrate community recovery.

Methods

Site Selection

Two distinct watersheds were studied. The Abram Creek (AC) watershed is within the North Branch of the Potomac River Basin, and Three Fork Creek's (TFC) watershed is located in the Tygart Valley River Basin. Both watersheds are dominated by forest with less than 20% agricultural land use. Both historically suffered extensive abandoned mine drainage (AMD) impacts, and have been receiving active treatment via alkaline dosers implemented in 2010 (AC) and 2011 (TFC) by the West Virginia Department of Environmental Protection's Office of Abandoned Mine Land and Reclamation.

Study sites within each watershed were strategically placed in consideration of active treatment sites and confluences of tributaries (see Fig.1). The selection process was designed to capture the spectrum of water qualities within the watershed. Sites were broken into three distinct treatment types: treated (downstream of active treatment), AMD (no upstream treatment), and reference

(naturally circumneutral/unimpacted sites). Three Fork Creek has 19 sites established (5 reference, 4 AMD, 10 treated) and Abram Creek has 17 sites (4 reference, 1 AMD, 12 treated).

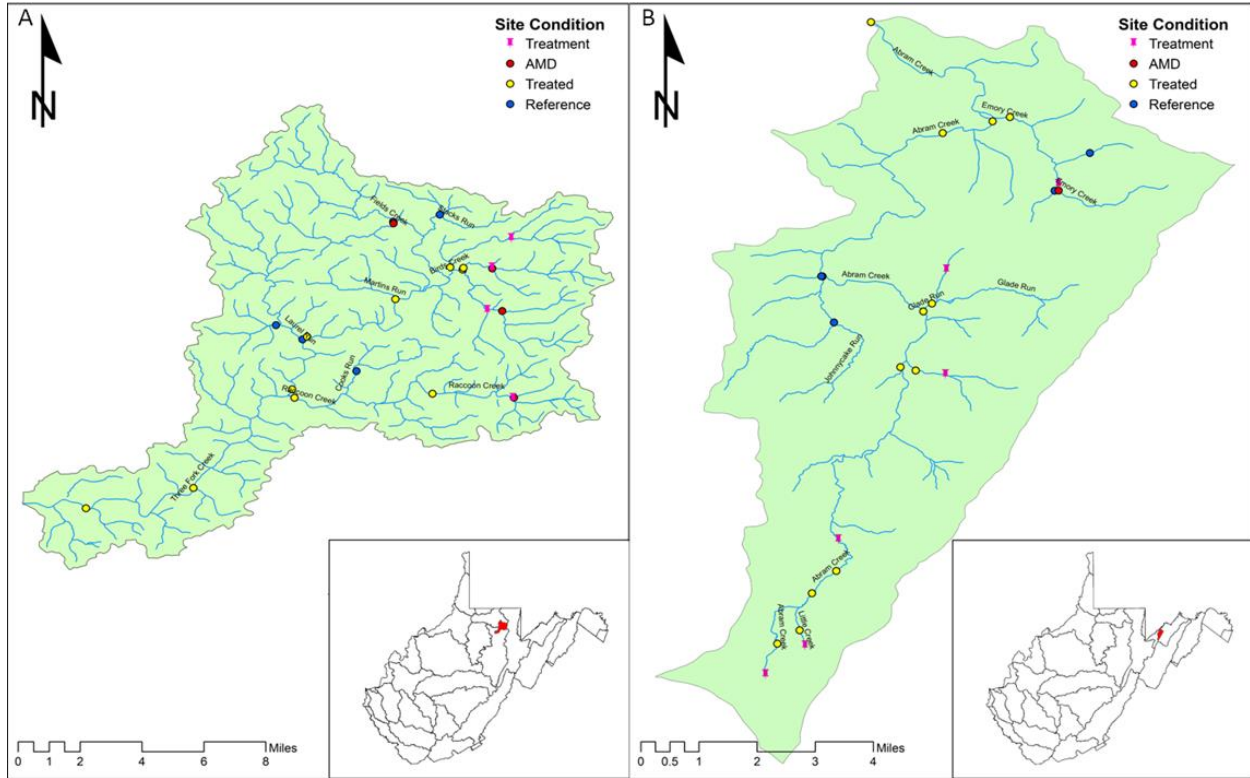


Figure 1. Study watersheds Three Fork Creek (A) and Abrams Creek (B) with sites labeled by treatment type

Sampling Techniques

Water chemistry samples were collected to inform the context of collected macroinvertebrates. Samples were collected during baseflow conditions between May and June, then stored at 4°C until West Virginia University's National Research Center for Coal and Energy analyzed them for alkalinity/acidity, sulfate, total dissolved aluminum, barium copper, chloride ion, cobalt, chromium, cadmium, calcium, sodium, nickel, selenium, zinc, iron, magnesium, and manganese content.

On the same days that water quality was sampled, benthic macroinvertebrate samples were collected via kick sampling following standard protocol as set forth by the West Virginia Department of

Environmental Protection (WVDEP, 2009). Samples were collected from four distinct riffles to yield one square meter of sampled substrate for each respective site. The contents of each kick were combined into a single container for each site per year and stored with 95% ethanol. Invertebrates were sorted in the lab via 200 count subsampling (WVDEP, 2009) before identifying to genus via dichotomous keys (Merritt et al. 2008).

Analysis

Objective 1: Post-restoration recovery

We summarized (mean and standard deviations) each water chemistry attribute by year and site type to characterize the effect of restoration on water chemistry within each watershed (Table 1). Subsampled macroinvertebrate collections were scored using the Genus Level Index of Most Probable Stream Status (GLIMPSS), an index of biotic integrity (IBI) developed and employed by West Virginia's Department of Environmental Protection (Pond, G.J. et al. 2013). We ran a linear mixed effects model to test for efficacy of treatment over time with site as a random effect to account for repeated measures. The model showed differences between pre- and post-restoration treated GLIMPSS scores so we then ran independent sample T-tests to compare mean GLIMPSS scores between treated and reference sites within years. We hypothesized that a significant difference would exist in mean treated GLIMPSS scores from pre- to post-treatment in both watersheds, and that 2017 and 2018 mean treated scores may not be significantly different from mean reference scores from the same years in one or both watersheds.

Objective 2: Community composition

Trends in a community can often be lost in the course of its wholesale reduction to an indexed score via IBI. The Bray-Curtis distance coefficient was used to provide a more holistic analysis by converting community abundance data into dissimilarity measures between sites by years. Essentially

each permutational pairing of sites by year was assigned a value based on the similarity of the composition of their macroinvertebrate communities. Taxa occurring in five percent or fewer of total samples per watershed were deemed “rare” and removed from the dataset prior to creating the matrix due to their potential to disproportionately drive dissimilarity (McCune et al. 2002). The matrices were then run through the function ADONIS2 - an analysis of variance through distance matrices - found under the *vegan* package in R (Oksanen et al., n.d.). This function incorporates permutational multivariate analyses of variance (PERMANOVA), and is based on the work detailed in Anderson’s methods paper (2001). Values were tested with 999 permutations to assess the significance of year and treatment type on community composition similarity (constrained to treated sites across all years), grouping permutations by site name to account for resampling. Unless otherwise noted, $\alpha=.05$ for this and all subsequent tests. We expected to find that year was a significant predictor of community dissimilarity in both watersheds. We then conducted the same test analyzing the difference in community dissimilarities based on treatment type between reference and treated sites within individual year subsets of the same dataset of rare-removed taxa for each watershed. For this follow-up testing we expected to see treatment type being a strong predictor of community dissimilarities before treatment with its strength as a predictor diminishing over time after treatment implementation in both watersheds. Due to the conceptually abstract nature of interpreting PERMANOVA results on Bray-Curtis dissimilarity, we elected to pair a graphical representation of the Bray-Curtis ordination of the community compositions using the above described matrices with non-metric multidimensional scaling (NMDS). Each community composition PERMANOVA has an NMDS that was constructed with an identical matrix. We felt this would facilitate comprehension and interpretation of the somewhat obscure results by allowing the audience to visualize the dimensionality of the site communities and personally observe treatment type clustering and variations over the course of the study.

Communities were also run through nested temperature analysis (after rare taxa were removed,

as described above) in *vegan* in R to facilitate comprehension of taxa persistence and trends in treatment over time. This analysis packs sites and taxa to optimize a curve of expected occurrence and then assigns a T statistic of 0-100 to the dataset according to how nested the communities are. In a dataset with a T of zero the site with the greatest richness would represent every taxa found in the data, and every subsequent site would be a shrinking subset of the richer sites. A T of 100 indicates that the sample is not nested at all. A T statistic alone cannot be compared between matrices of different sizes and fills but comparison is valid when considering the p values. Each analysis of nestedness was run with 99 simulations. (Rodríguez-Gironés and Santamaría, 2006). We hypothesized that nestedness would increase (meaning the T statistic would decrease) from 2008 to 2013 in AC and then decrease over time – reflecting the increased richness from a new “disturbance”, and then later a patch-sorting dynamic in regards to the remaining AMD sites. The sorting dynamic would present as a checkerboard pattern on the nestedness graphs as it reflected different sites’ water quality dictating the communities which could thrive there. To elaborate on this concept: degraded water quality in an AMD site would preclude colonization of a pollution intolerant community, and in circumneutral/restored streams pollution intolerant taxa would be better adapted and should outcompete pollution tolerant taxa as long as regional recruitment was available and there was not an overwhelming mass effect of pollution tolerant taxa infiltration. A similar result could be expected in TFC, but as we do not have data from 2013 we instead hypothesize that nestedness would increase over time in our dataset due to the addition of more varied water quality in the system.

Objective 3: Recovery in relation to treatment distance

PERMANOVA run with 999 permutations was used to test the strength of minimum distance to lime dosers as a predictor of average community composition within each individual year. Additionally, GLIMPSS scores were plotted versus site distance from lime dosers to visualize the trend. A second-order polynomial trendline was fitted to these points, as it had the highest R^2 of the tested lines. Given

the context of the study in relation to restoration we also elected to plot change over time in GLIMPSS (final year/site – first year/site) versus site distance from dosers to concisely illustrate IBI improvement as it pertains to treatment proximity. We hypothesized that minimum distance to treatment would be a significant predictor of community composition in both watersheds after treatment implementation, we expected this to manifest on the graphs as a trending increase in GLIMPSS and change-in-GLIMPSS as distance to treatment increased.

Results

Objective 1: Post-restoration recovery

ANOVA and Tukey's post-hoc on GLIMPSS scores grouped by treatment

In TFC the mean GLIMPSS score of treated sites was significantly different between pre- and post-remediation treatment years (2008-2017 $t=4.90$, $p<.001$; 2008-2018 $t=7.09$, $p<.001$). Mean score of treated sites was significantly different from reference in both 2017 ($t=-3.96$, $p=.002$) and 2018 ($t=-.245$, $p=.02$). In AC mean GLIMPSS score of treated sites was significantly different between pre- and post-remediation treatment (2008-2013 $t=6.50$, $p<.001$; 2008-2017 $t=6.54$, $p<.001$; 2008-2018 $t=8.45$, $p<.001$). Mean score of treated sites was significantly different from reference in 2013 ($t=-3.82$, $p=.004$), 2017 ($t=-3.59$, $p=.008$), and 2018 ($t=-2.82$, $p=.040$). See figure 3 for a boxplot of treated GLIMPSS scores with average reference and AMD scores.

Objective 2: Community Composition

PERMANOVA and NMDS on Bray-Curtis dissimilarity matrices of community composition

Figures 4-7 display graphic representations in NMDS of the Bray-Curtis dissimilarity of site communities used in this calculation, the following results are most easily understood when considered in conjunction with the figure. In TFC statistically significant differences existed between the average dissimilarity measures of treated communities by year ($F=6.3$, $p<.001$) and the stress on the NMDS of

treated sites across all years was 9.81%. By comparison, the stress on the NMDS for treated and reference communities was 12.25%. Dissimilarity could not be tested between treatments for 2008 in either watershed as only one reference was present in each at the time (NMDS stress=3.84%), in 2017 a significant distinction existed ($F=3.0$, $p=.021$, NMDS stress=10.26%), and in 2018 there was once more no significant difference in the average community dissimilarity measures between reference and treated sites ($F=1.2$, $p=.260$, NMDS stress=9.27%).

When considering the total dataset, AC had results similar to TFC with statistically significant differences existed between the average dissimilarity measures of treated communities by year ($F=8.1$, $p<.001$) and the stress on the NMDS of treated sites across all years was 17.30%. By comparison, the stress on the NMDS for treated and reference communities was 19.41%. When looking at individual years, AC had a much more linear trend over time, with no significant differences in average community dissimilarity by type shortly after treatment (2013 $F=1.7$, $p=.080$, NMDS stress=9.15%), but establishing and maintaining a significant difference in average community dissimilarity between the two treatment types as time after restoration increased (2017 $F=2.8$, $p=.022$ NMDS stress=12.76%; 2018 $F=2.6$, $p=.014$, NMDS stress=12.11%).

Nested temperature of communities

The communities within TFC were most nested before treatment (2008 $T= 12.69$, $p=.01$), and became moderately less nested as time progressed (2017 $T= 20.98$, $p=.01$; 2018 $T= 22.43$, $p=.01$). Communities exhibited nestedness comparable to that of post-treatment when observed across all years sampled (08-18 $T= 19.92$, $p=.01$). See figures 8-11 for displays of TFC community composition by year with a curve for expected taxa occurrences.

Degree of nestedness in AC was slightly more variable than TFC. Nestedness in AC decreased from 2008 to 2013 (2008 $T= 15.08$, $p=.01$; 2013 $T= 32.50$, $p=.01$), the communities then exhibited a continued increase in nestedness from 2013 through to 2018 (2013 $T=32.50$, $p=.01$; 2017 $T=30.84$,

$p=.01$; 2018 $T = 20.54$, $p=.01$). The degree of nestedness within AC communities was similar to the highest measured individual year value when testing across all years (2008-2018 $T= 24.18$, $p=.01$) but this effect could be attributed to increased matrix size and the nature of the calculation. See figures 12-16 for displays of AC community composition by year with a curve for expected taxa occurrences.

Objective 3: Recovery in relation to treatment distance

Community composition and minimum distance to treatment PERMANOVA

Minimum distance to treatment was not a statistically significant predictor of community composition in sites downstream of lime dosers in TFC before treatment (2008: $F=1.59$, $R^2=.15$, $p=.17$); however it was a significant predictor in both years sampled post-treatment (2017: $F=4.90$, $R^2=.35$, $p=.001$; and 2018: $F=2.81$, $R^2=.24$, $p=.034$).

Minimum distance to treatment was a statistically significant predictor of community composition in sites downstream of lime dosers for all years sampled in AC, even before treatment was implemented (2008: $F=3.55$, $R^2=.31$, $p=.026$; 2013: $F=3.82$, $R^2=.32$, $p=.003$; 2017: $F=2.45$, $R^2=0.23$, $p=.03$; and 2018: $F=2.88$, $R^2=.26$, $p=.016$).

Community integrity scores and minimum distance to treatment

In both TFC and AC GLIMPSS scores improved as treatment distance increased throughout the years sampled (see Fig. 17). When the change over time in site scores was plotted and fit to a trendline, TFC sites had a rising growth in GLIMPSS scores associated with increasing distance ($R^2=0.660$), while AC had a decreasing growth in GLIMPSS scores over time as distance increased ($R^2=0.581$) (see Fig.18).

Discussion

Our results support previous conclusions that lime dosing of streams has a measurable positive impact on water quality (Long, 2019; Watson et al., 2017). Though they demonstrate that these positive

effects have not persisted to the same degree into 2018. Though the body of research concerning watershed-scale restoration is increasing (Palmer et al., 2014), it will be some time before the finer mechanisms occurring within that broader context are well understood (Brown et al., 2011; Idígoras Chaumel et al., 2019; Stoll et al., 2016). Later we will discuss some of the areas we believe are most in need of further investigation.

Objective 1: Post-restoration recovery

While we observed a definite trend of treated sites' IBI scores rising to approach those of reference, we also observe that the average score of 2018 references in both watersheds were higher than previous references. It is possible that this increase in reference scores is evidence of watershed-scale inflation of community condition as a result of a healthier metacommunity pool. However, studies have indicated that hydrology can be a significant driver of benthic macroinvertebrate community health and composition (Sarremejane et al., 2018). To safely conclude these results should be attributed to metacommunity dynamics, one would first have to account for hydrological influence. Unfortunately numerous and inconsistent freezing and low water events rendered the applicable USGS gaging station data unreliable during the study period. As a result, though it seems likely these watershed-wide IBI increases are due to net metacommunity improvement it would be irresponsible to conclusively state this is so. While treatment implementation was a significant predictor of IBI scores in treated sites for both watersheds, our study period did not capture a year in which average treated site scores were not statistically different from average reference site scores.

Objective 2: Community Composition

Taking a step back from the simplified scores of IBI's and looking at the communities as dynamic entities through Bray-Curtis dissimilarity measures provides a different story. It is with these tests that obvious differences begin to arise between TFC and AC. Results indicate that sample year was a

significant predictor of average community dissimilarity when comparing treated sites across all years in the study for each respective watershed. However testing between treatment types within individual years showed quite the opposite effect from what we had expected. If the Bray-Curtis dissimilarity measure is to be believed as a direct measure of community composition then the PERMANOVA results would indicate that both watersheds had reference communities thriving in pre-treatment treated sites (or else vice versa); however this is likely due to having too few pre-treatment reference sites to detect a significant difference. Other results provide further insight, and by consulting plots for GLIMPSS metrics it is clear that the treated communities were not exhibiting healthy reference scores/taxa. Continuing our exploration of these unexpected results, we look to the NMDS plots, and observe what appears to be visual grouping of treatment types with a persistently larger spread representing the treated communities; it is likely that this larger degree of variation in treated sites is responsible for their continued failure to obtain reference quality. While some treated sites have responded well, others have maintained low quality, as evidenced by the few that remain persistently low on nestedness graphs – failing to add taxa even as quickly as some AMD sites despite an empirically improving watershed. The word “similarity” in this context may be unnecessarily confusing, if we think of the PERMANOVA measurement on Bray-Curtis dissimilarities instead as representing the range of possible site-to-site deviation within each treatment type’s total community, then it becomes more meaningful. Possible causes of the varied results between watersheds are myriad; but three likely solutions present themselves: there was a difference in regional recruitment which allowed for TFC to recolonize acid-sensitive taxa more rapidly (Merovich et al., 2013; Merriam and Petty, 2016), the more dendritic network of TFC allowed for a broader range of community compositions resulting in more detectable differences (Vannote et al., 1980), or the higher degree of initial impairment in TFC captures a broader spectrum of recovery. The differences noted could be due to one of these factors, an interplay between the three, or something else entirely.

The nestedness results of the two watersheds were also quite opposite of expectations. The decreasing degree of nestedness observed in TFC communities appears to be a result of rapid recruitment to treated sites, which supports the above suggestion that there may be a difference in regional recruitment between the two watersheds (even further support exists in the unexpected presence of a 2018 AMD site near the top of the TFC all-year nestedness graph). The initial decrease of nestedness in AC after restoration was likely due to the beginning of taxa sorting by patch dynamics, wherein the variance of habitat suitability could begin to affect the taxa recruited, as opposed to the original regime of isolated references within an impaired regional pool of taxa. It follows logic that while the sorting regime persisted, nestedness increased in AC as more taxa were recruited overall. Of particular interest are the treated sites from 2017 and 2018 that occupy the bottom of the graph representing all years for AC. It was our initial hope that incorporating the distance to treatment testing would explain these variations.

Objective 3: Recovery in relation to treatment distance

Distance to treatment has been repeatedly called a critical aspect of AMD restoration studies (Long, 2019; McClurg et al., 2007; Watson et al., 2017). Anomalous results in AC, indicating that distance to treatment was a significant predictor of community composition even before treatment implementation, were concerning. A glance at the layout of our study sites quickly demonstrates issues with this approach. It is well known that stream order dictates community composition (Vannote et al., 1980). AMD impacts in our watersheds were mostly situated in headwaters. Therefore, there is an inherent correlation between distance to treatment and stream order. Add to this the dendritic nature of the studied watersheds, and we gain the second confounding factor: as distance from treatment increases, so too does number of upstream impacts/treatments.

Due to these setbacks PERMANOVA results probably do not bear a direct interpretation on the influence of distance to treatment on community composition, but the suite of results compiled here do demonstrate treatment's impact on TFC, which appear to have allowed for dynamics of the river continuum concept to begin functioning after treatment. Due to these confounding factors, we elected to use the coarser evaluation of IBI's as they are designed to operate uniformly across stream orders. Plotting GLIMPSS scores of treated against distance to nearest treatment returned a concise narrative: as distance from treatment increases, site scores increase, though this effect appears to diminish after an initial jump in quality. Perhaps the most perplexing result of this study lies in the graphs of delta GLIMPSS (final-initial) versus minimum distance to treatment. In TFC sites tended to improve more as distance from treatment increased, however AC sites tended to show less improvement as distance from treatment increased. Possible explanations for these variances are similar to those posed earlier, it could be due to the initial degree to which each system's mainstem sites were impacted (as those are most distant from treatments), or it could reflect differences in regional recruitment pools as opposed to efficacy of restoration.

Implications for future research

It appears safe to conclude that the trends of recovery documented by Long and Watson are largely concluded, though watershed-scale restoration has demonstrated some unexpected positive results (such as surprising richness of AMD sites in TFC). Some of our difficulties could have been avoided with a few study design alterations had we known in 2008 that the research would take this direction. To address the issue of treatment distance being a proxy for stream order, one could sample at regular and identical intervals down the length of several treated systems and compare these communities to those sampled in the same manner in untreated systems. This would allow the researcher to conclude at what distance from treatment the system returns to the dynamics of the river

continuum concept, as well as document if and how this distance changes over time after implementation. Another possibility would be situating a “bad” reference watershed in an area with obvious recruitment issues, such as the heart of the southern coal fields, in order to experimentally isolate the impacts of regional recruitment. A problem with this sampling scheme is the issue of locating and gaining access to so many qualified watersheds. At the least, our research has provided valuable insight into community trends to inform the context of previous studies on the efficacy of watershed-scale restoration efforts. We hope that our efforts to illuminate the finer processes involved in systemic restoration may inspire another to do the same.

Figures

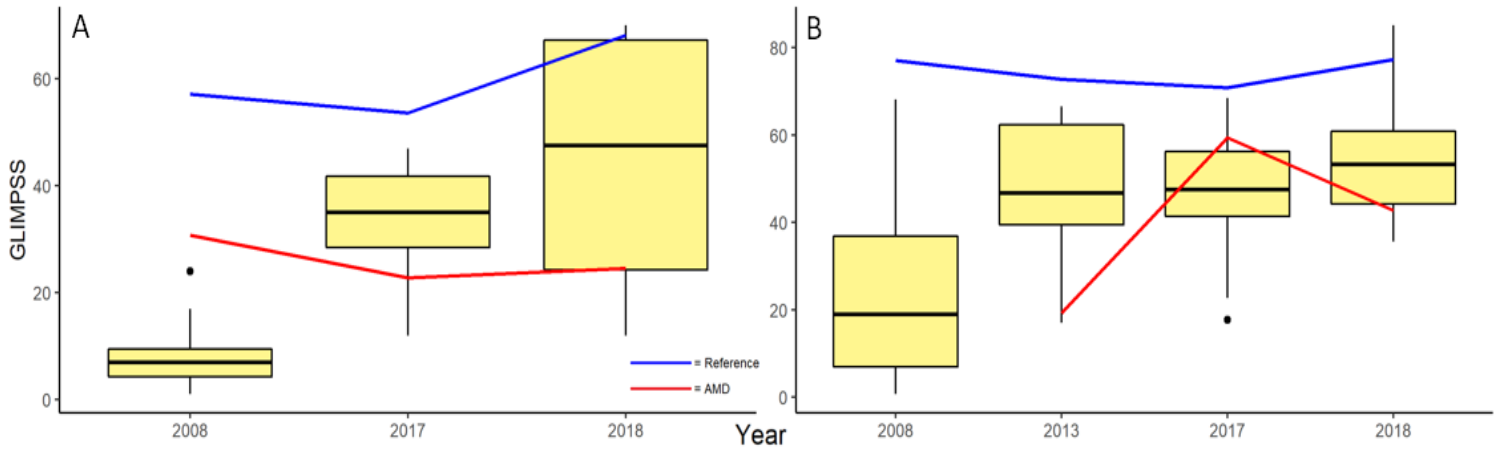


Figure 2 Boxplot of treated site GLIMPSS scores with mean reference and AMD scores plotted as lines by year for Three Fork Creek (A) and Abrams Creek (B)

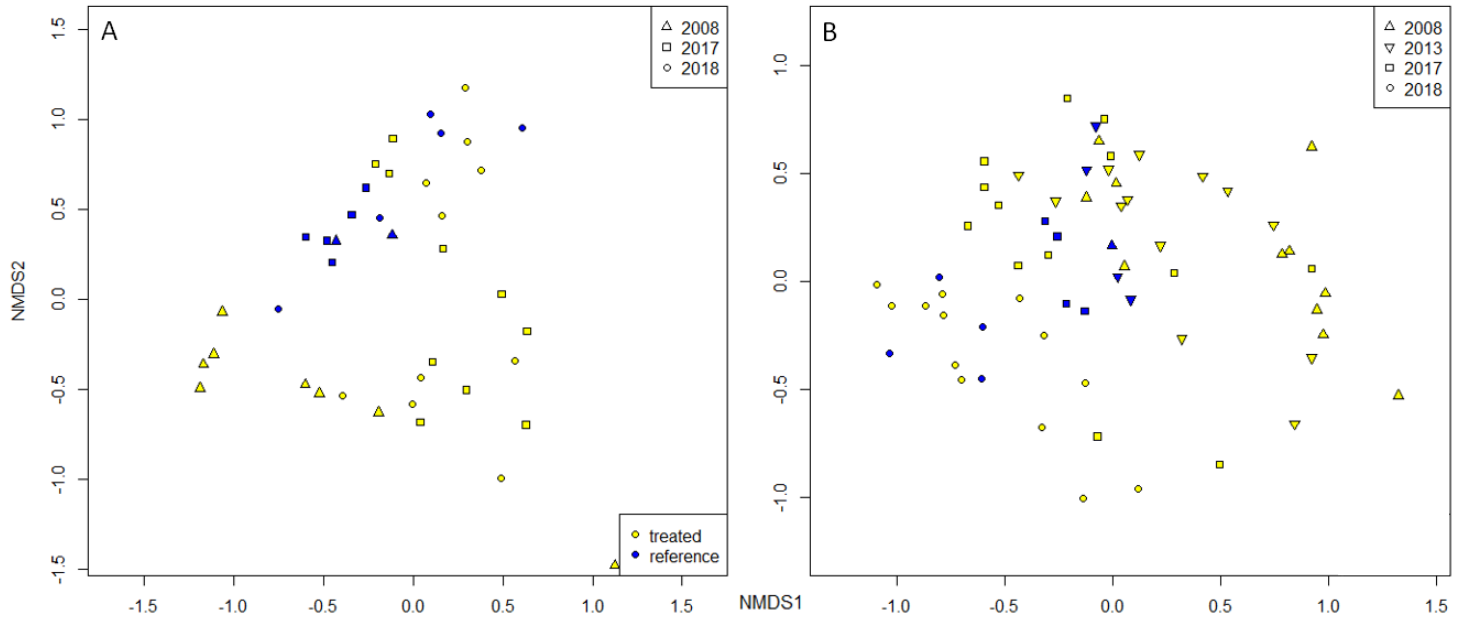


Figure 3 NMDS of Bray-Curtis dissimilarity of benthic macroinvertebrate communities by year and treatment for Three Fork Creek (A) and Abrams Creek (B)

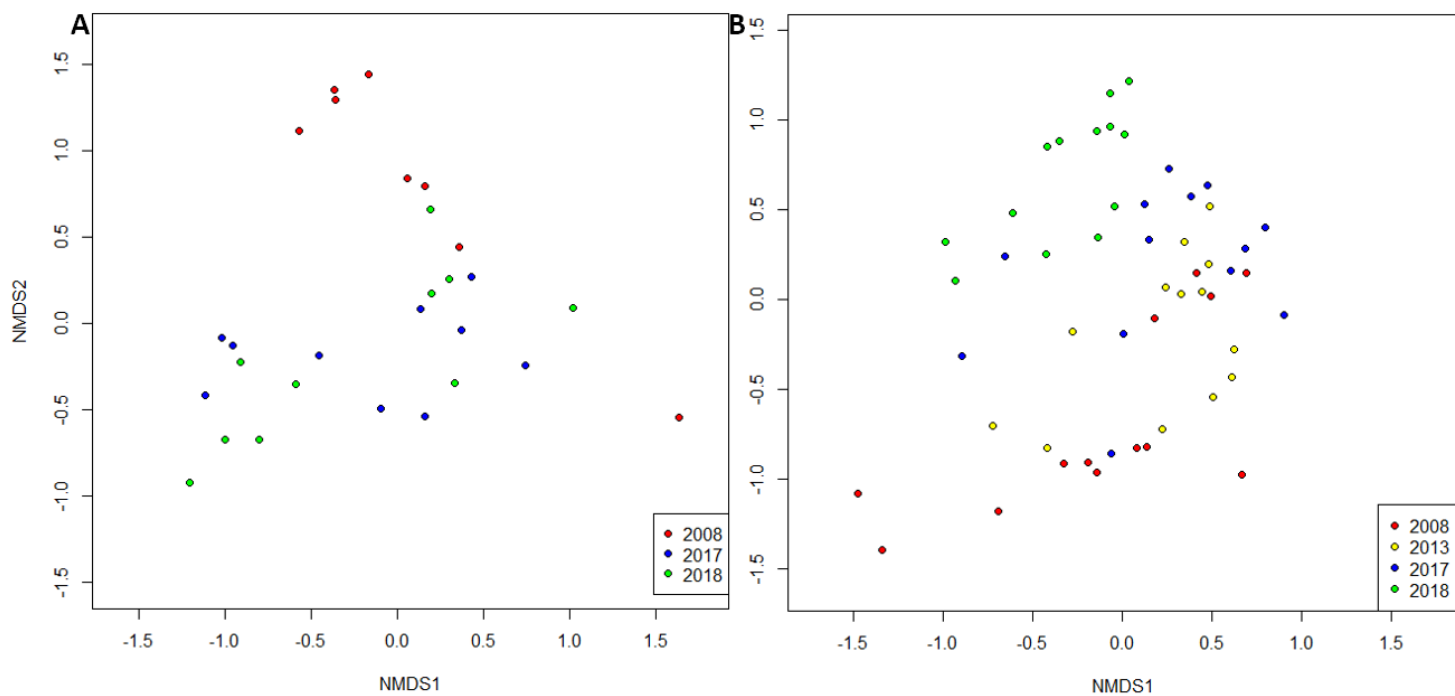


Figure 4 NMDS of Bray-Curtis dissimilarity of benthic macroinvertebrate communities by year and treatment for Three Fork Creek (A) and Abrams Creek (B)

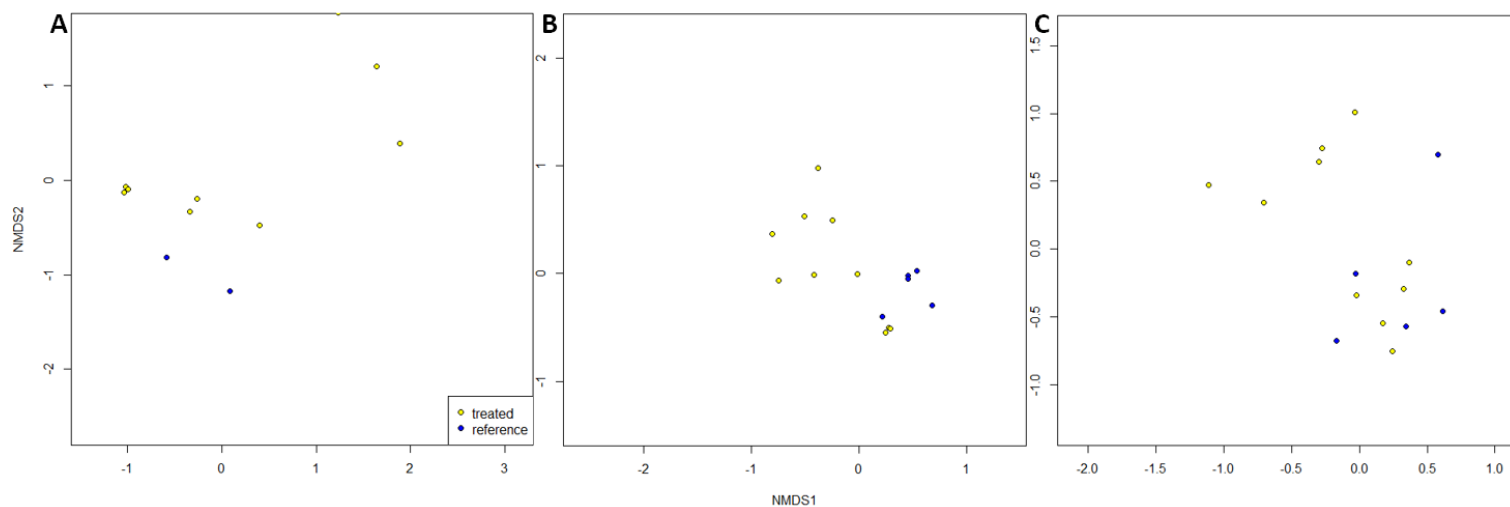


Figure 5 NMDS of Bray-Curtis dissimilarity of benthic macroinvertebrate communities by treatment for Three Fork Creek in 2008 (A), 2017 (B), and 2018 (C)

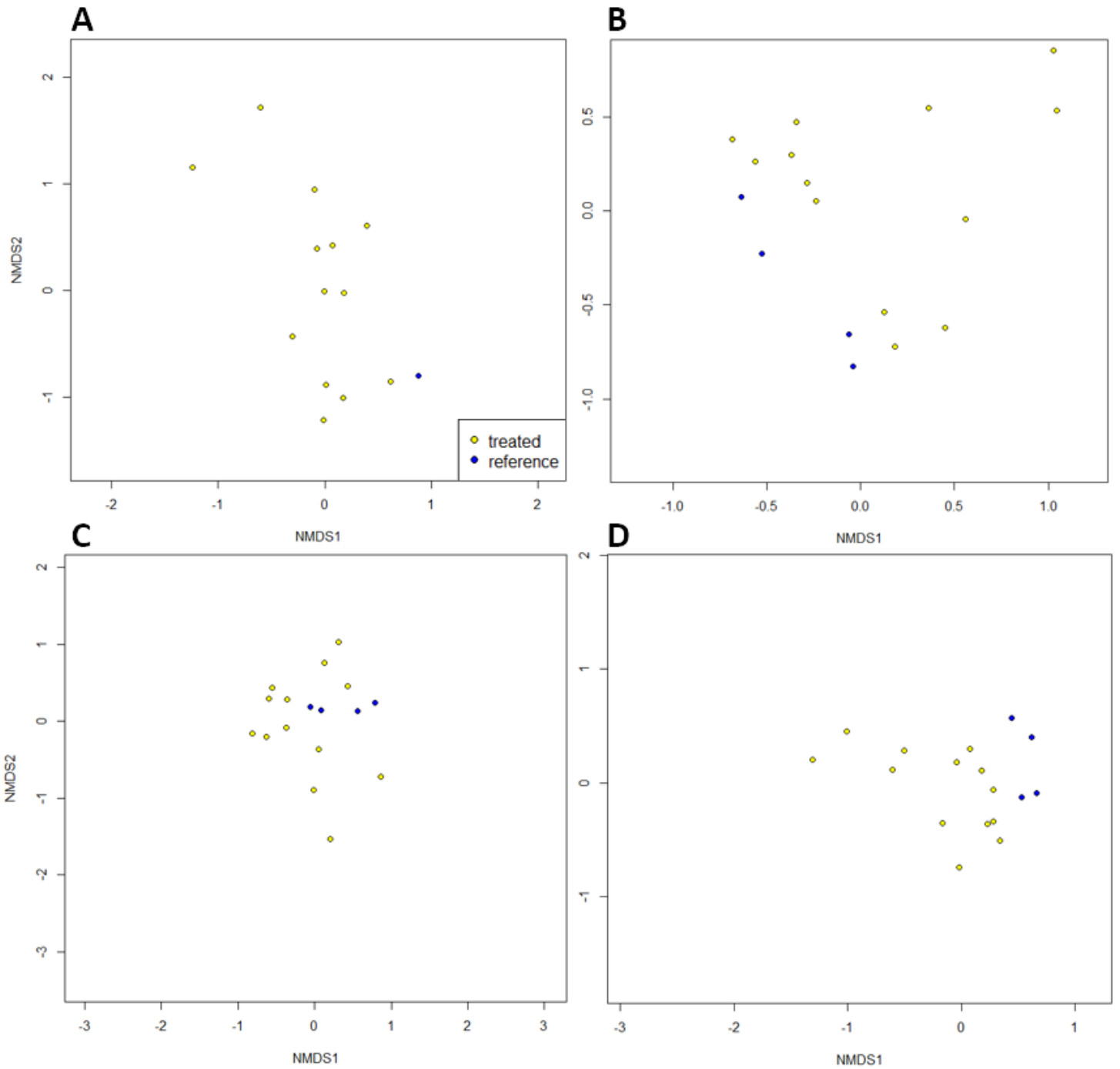


Figure 6 NMDS of Bray-Curtis dissimilarity of benthic macroinvertebrate communities by treatment for Abrams Creek in 2008 (A), 2013 (B), 2017 (C), and 2018 (D)

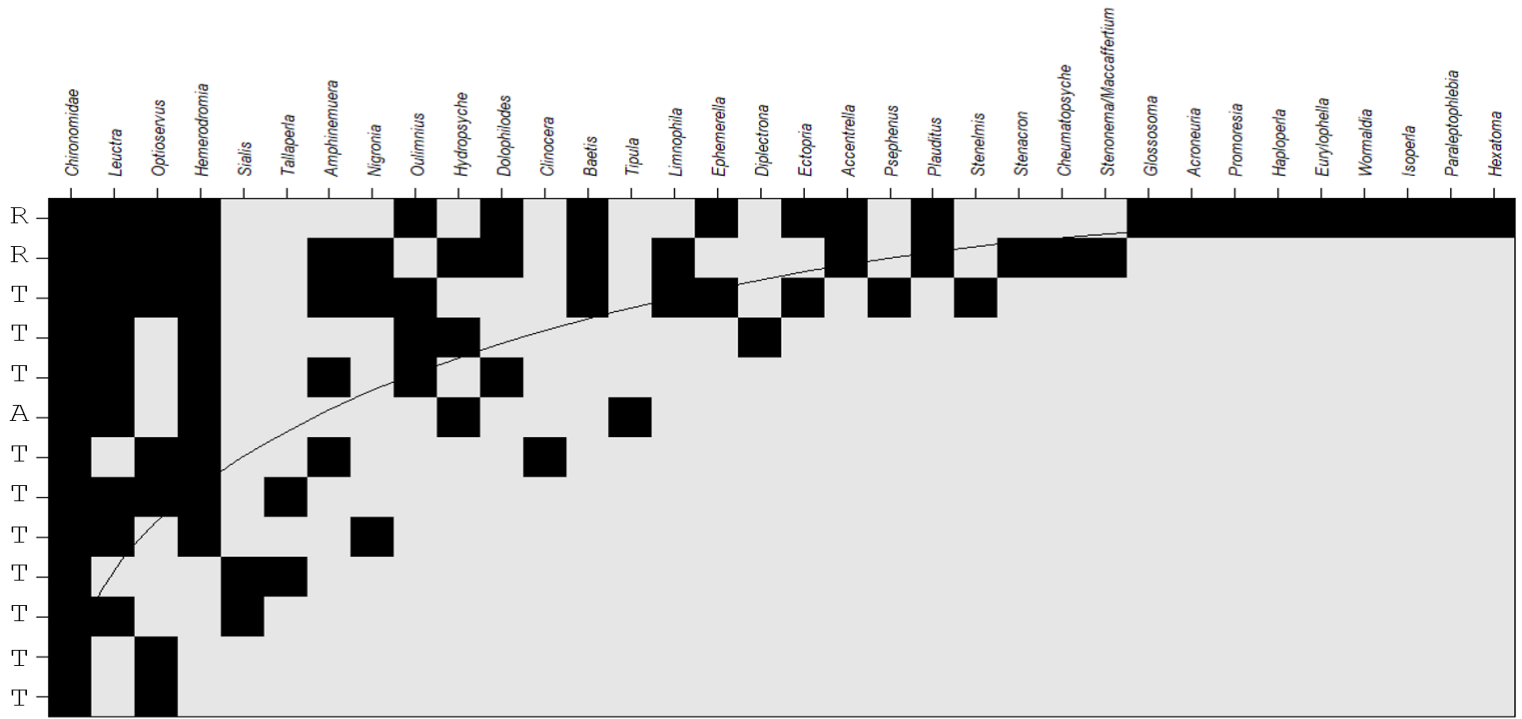


Figure 7 Nestod graph with occurrence curve of taxa, and treatment types (R=reference, T=Treated, A=AMD) in Three Fork Creek 2008

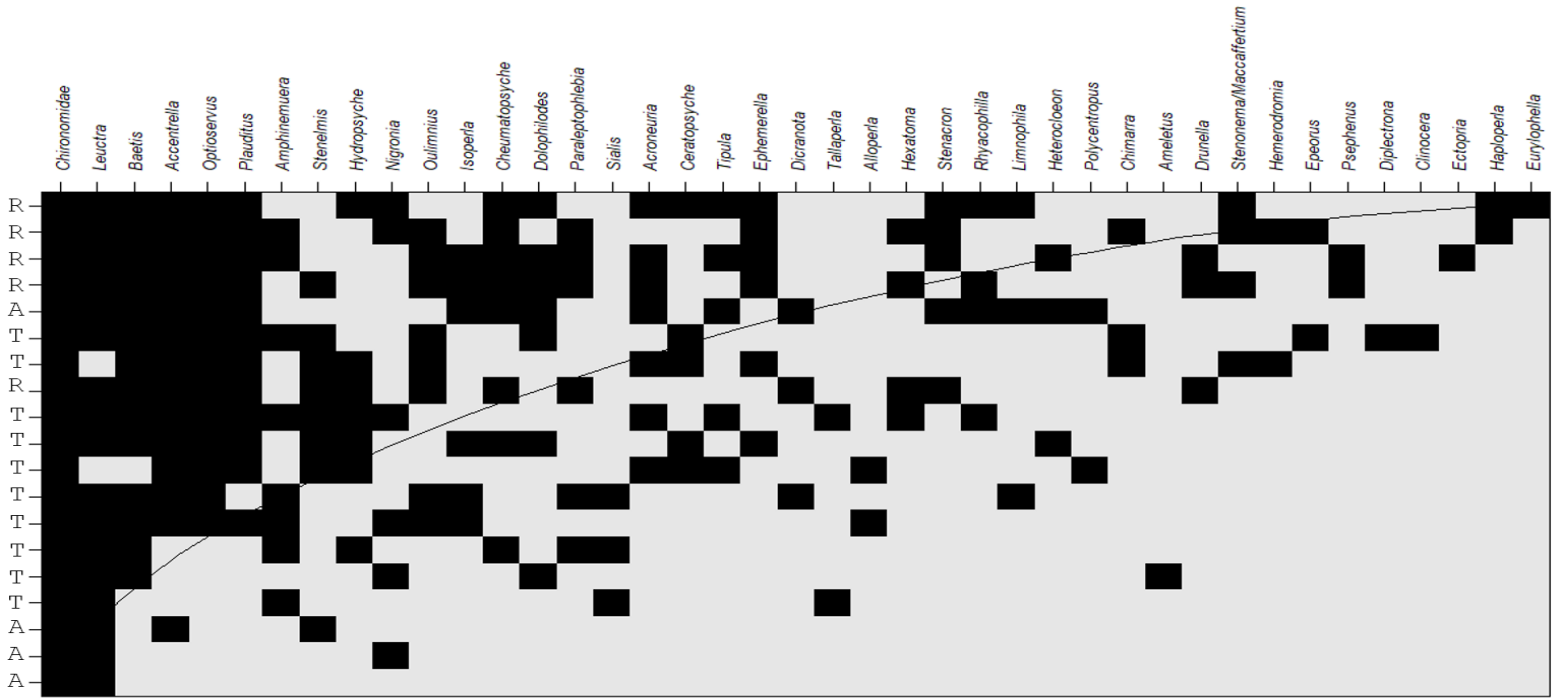


Figure 8 Nestled graph with occurrence curve of taxa, and treatment types (R=reference, T=Treated, A=AMD) in Three Fork Creek 2017

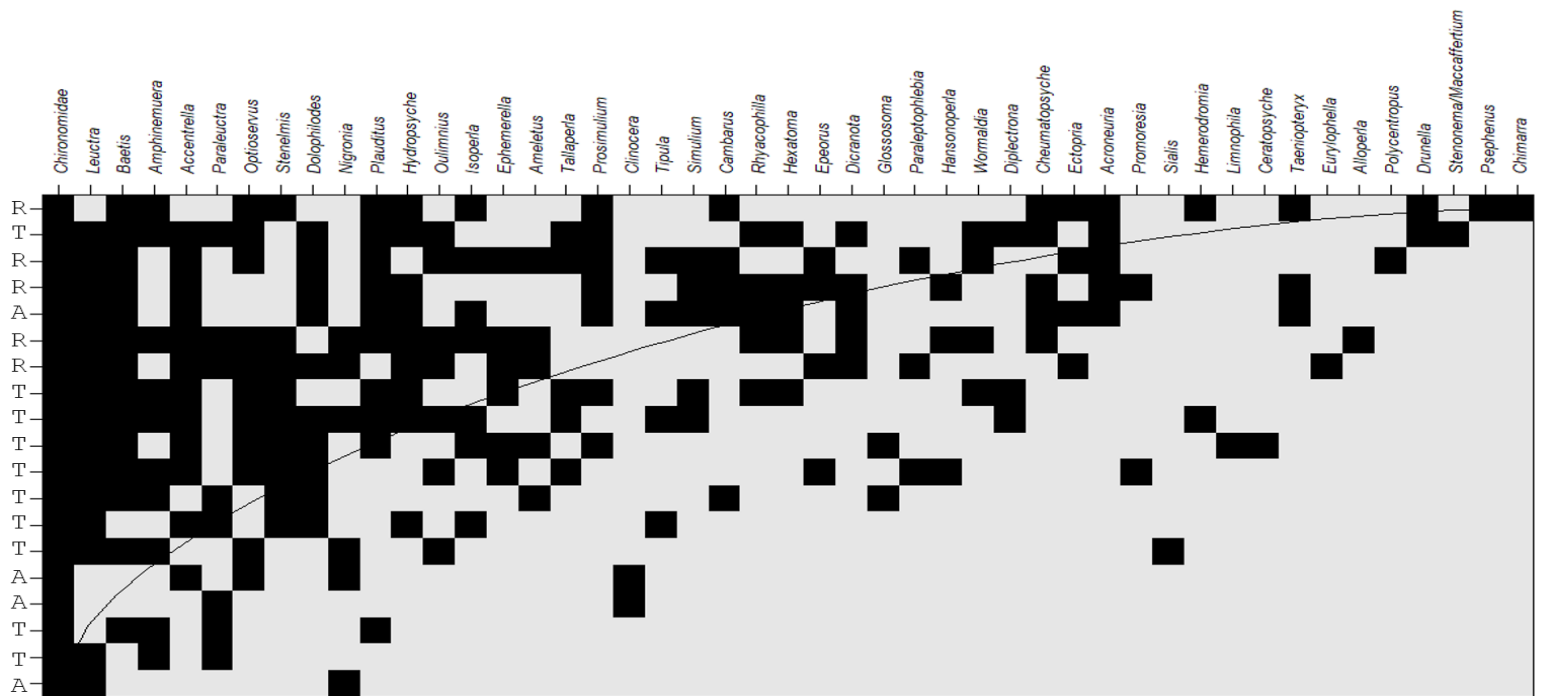


Figure 9 Nestled graph with occurrence curve of taxa, and treatment types (R=reference, T=Treated, A=AMD) in Three Fork Creek 2018

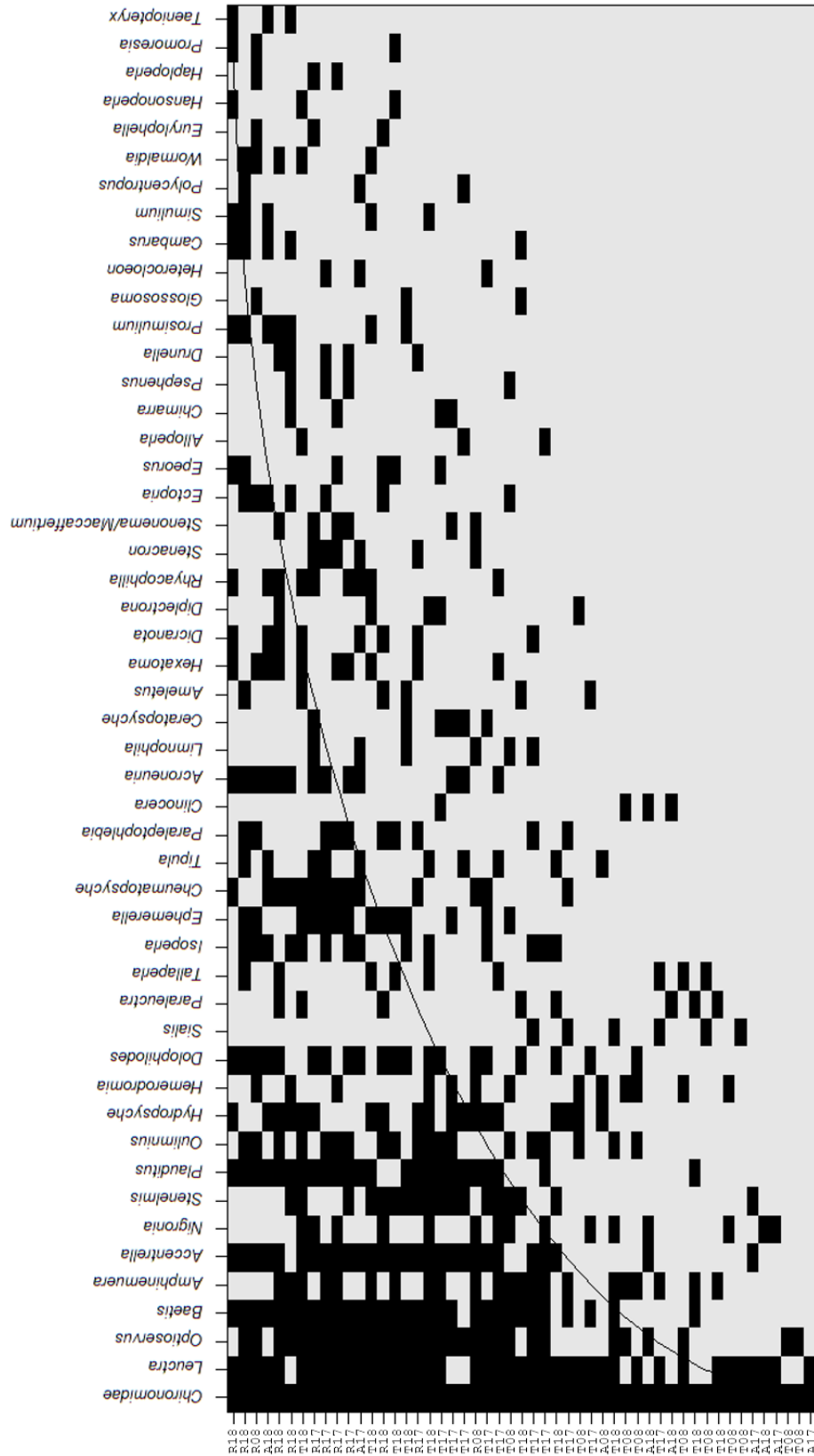


Figure 10 Nested graph with occurrence curve of taxa, and treatment types (R=reference, T=Treated, A=AMD) in Three Fork Creek for 2008-2018

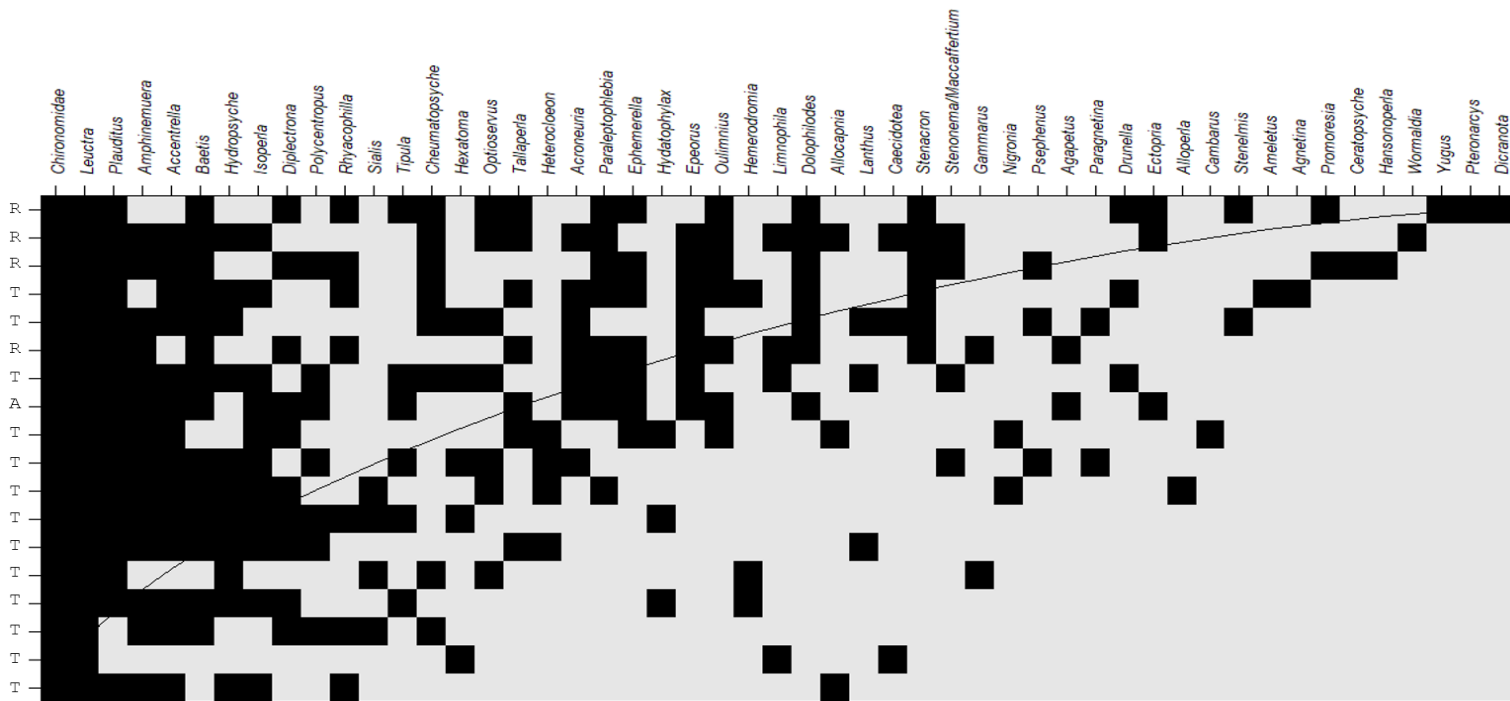


Figure 23 Nested graph with occurrence curve of taxa, and treatment types (R=reference, T=Treated, A=AMD) in Abrams Creek 2017

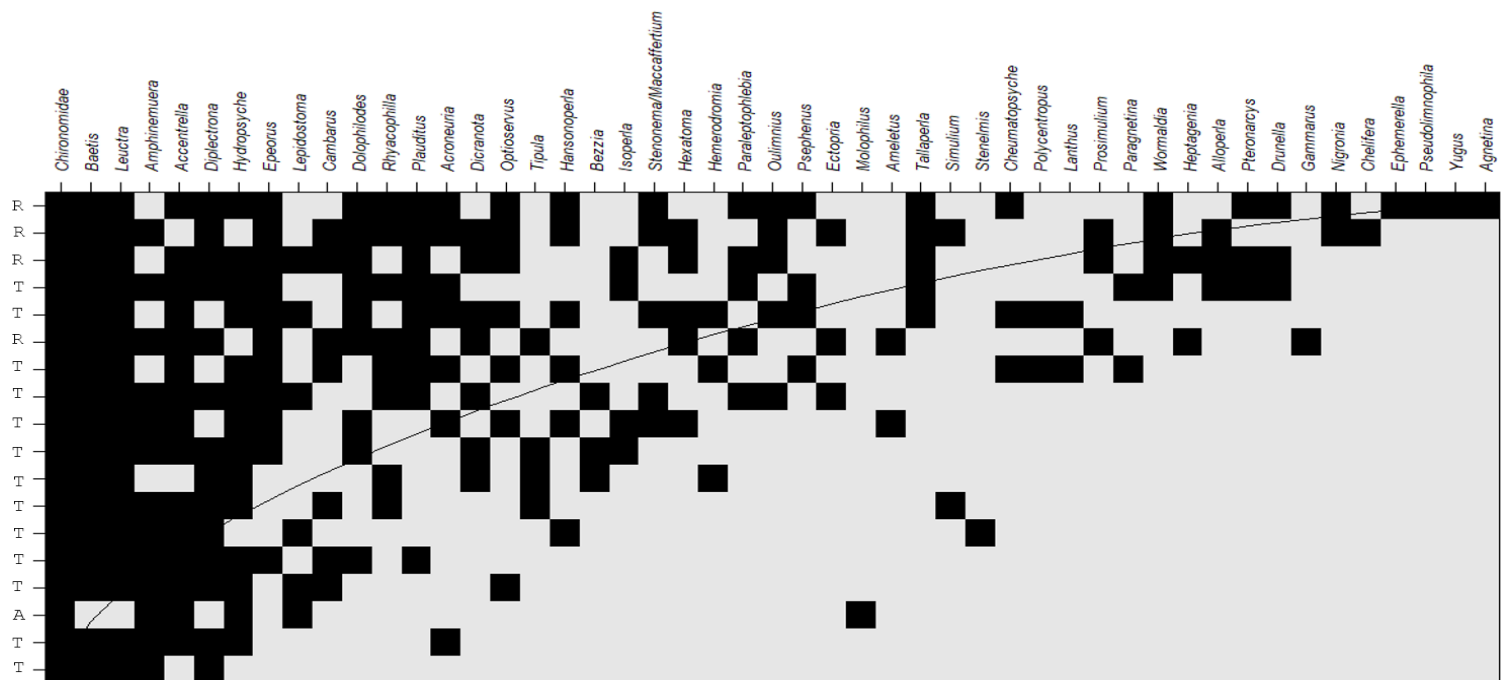


Figure 14 Nested graph with occurrence curve of taxa, and treatment types (R=reference, T=Treated, A=AMD) in Abrams Creek 2018

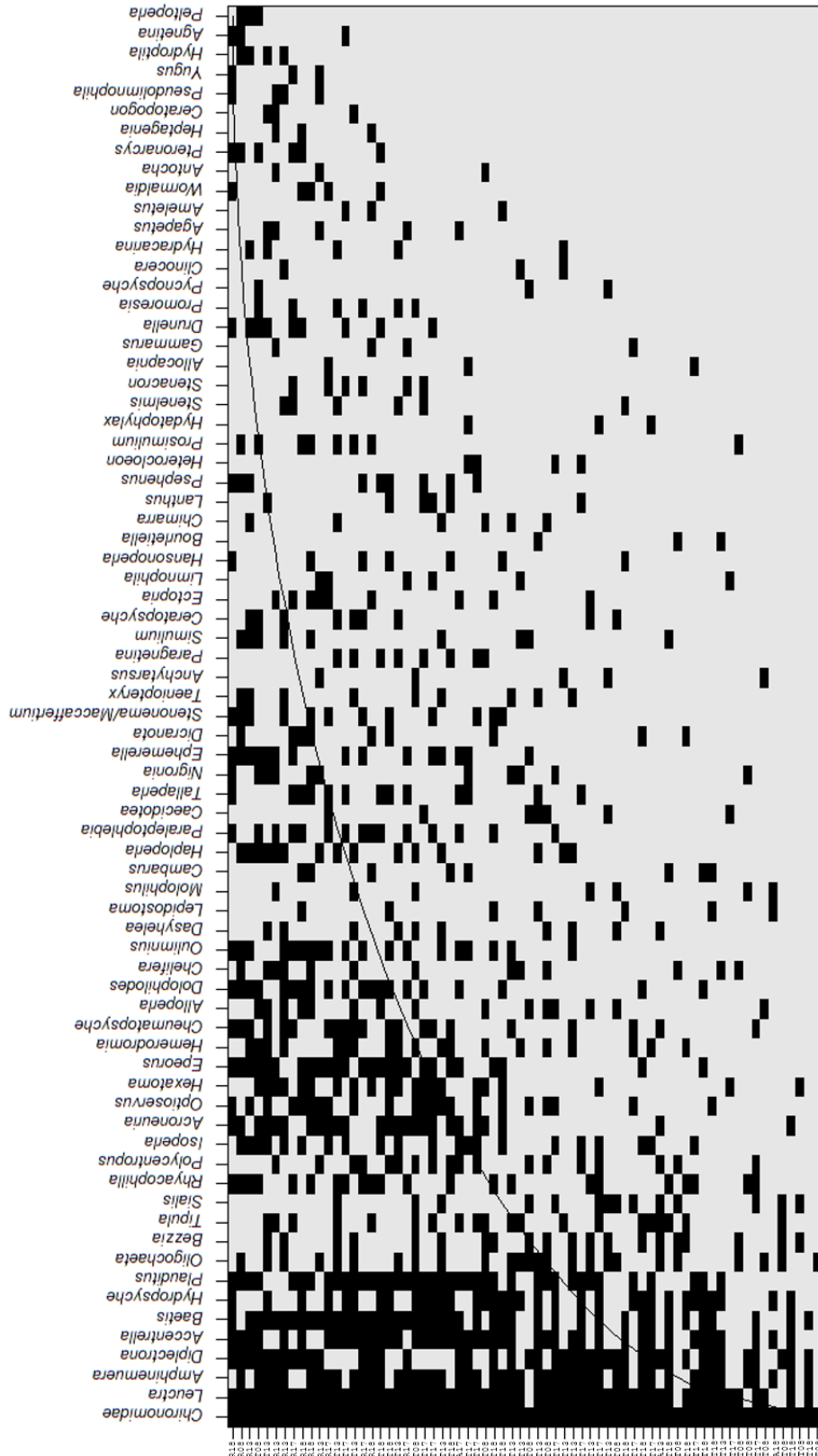


Figure 35 Nested graph with occurrence curve of taxa, and treatment types (R=reference, T=Treated, A=AMD) in Abrams Creek for 2008-2018

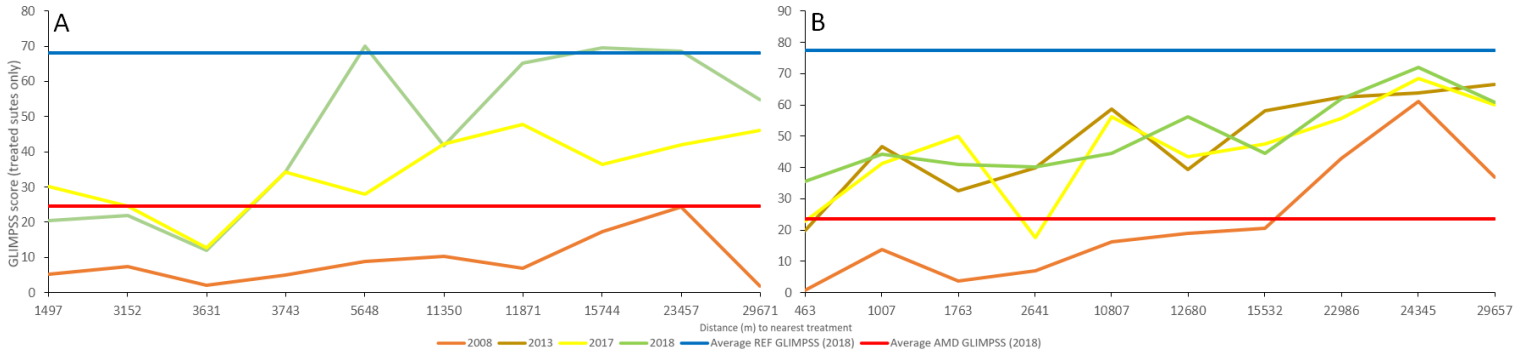


Figure 16 GLIMPSS scores plotted against distance to nearest lime doser for Three Fork Creek (A) and Abrams Creek (B) with a band for average reference score and average AMD score (respective to each watershed) in 2018

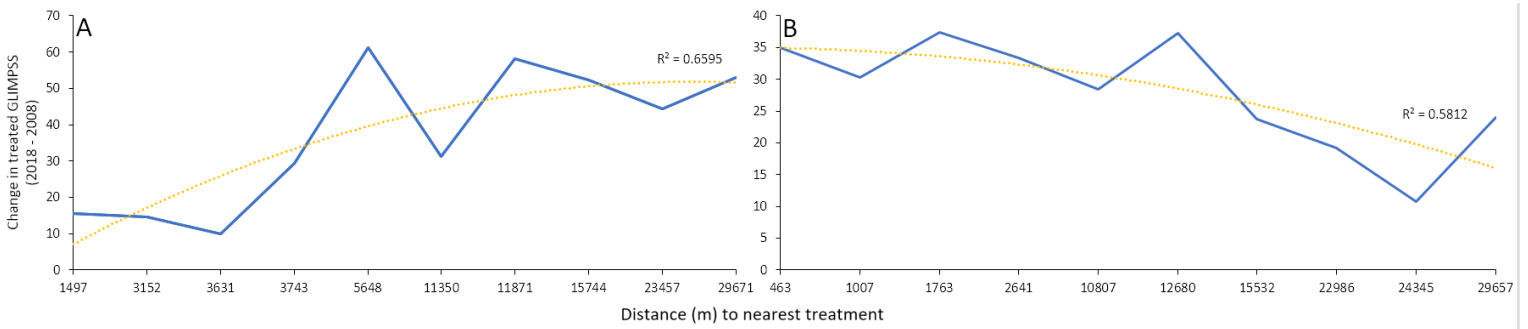


Figure 17 Change over time of GLIMPSS scores per site in Three Fork Creek (A) and Abrams Creek (B) plotted against minimum distance to treatment and fitted with a trendline

Tables

Table 1 Water quality summarized by treatment type for Three Fork Creek (A) and Abrams Creek (B). Averages are listed by site type with standard deviation provided in parentheses. All units are mg/L unless otherwise noted.

	2008			2017			2018		
	AMD (n=3)	Reference (n=6)	Treated (n=11)	AMD (n=3)	Reference (n=6)	Treated (n=11)	AMD (n=3)	Reference (n=6)	Treated (n=11)
Alk	0.68 (0)	49.79 (41.36)	1.43 (0.86)	0.68 (0)	21.14 (12.46)	12.86 (7.45)	0.68 (0)	24.53 (6.91)	19.85 (10.61)
Acid	140.18 (59.34)	8.54 (16.49)	88.34 (75.91)	70.86 (10.61)	1.81 (0)	8.77 (11.26)	70.48 (13.78)	1.81 (0)	5.01 (7.14)
SO4	444.67 (279.06)	206.37 (223.11)	509.45 (130.91)	129.67 (6.03)	25 (13.73)	137.63 (49.72)	163.71 (94.29)	107.32 (59.11)	100 (76.44)
Cl	3.69 (1.43)	5.97 (4.36)	4.44 (1.7)	2.54 (2.03)	3.82 (3.29)	3.04 (0.93)	2 (1.46)	3.73 (0.57)	3.07 (1.34)
Al	14.52 (6.28)	0.12 (0.19)	9.01 (6.94)	6.49 (0.81)	0.02 (0)	0.12 (0.34)	7.31 (1.77)	0.07 (0.05)	0.12 (0.18)
Ba	0.02 (0.01)	0.07 (0.02)	0.04 (0.02)	0.04 (0)	0.03 (0)	0.03 (0)	0.03 (0)	0.03 (0.01)	0.03 (0.01)
Ca	80.39 (54.68)	38.97 (24.27)	89.47 (20.56)	18.6 (5.42)	11.98 (4.7)	40 (13.97)	32.87 (24.82)	38.64 (16.43)	35.03 (19.36)
Cd	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Co	0.1 (0.07)	0 (0)	0.09 (0.05)	0.04 (0.01)	0 (0)	0.01 (0.01)	0.05 (0.02)	0.01 (0.01)	0.01 (0.01)
Cr	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Cu	0.01 (0.01)	0 (0)	0 (0.01)	0 (0)	0 (0)	0 (0)	0.01 (0.01)	0 (0)	0 (0)
Fe	0.69 (0.66)	0.32 (0.47)	1.88 (4.34)	1.01 (0.91)	0.04 (0.03)	0.22 (0.57)	0.91 (0.35)	0.02 (0.01)	0.45 (1.3)
Mg	23.18 (13.96)	11.08 (11.56)	26.43 (7.34)	6.72 (0.22)	2.62 (1.07)	8.42 (2.71)	9.22 (3.01)	7.74 (3.44)	7.02 (4.06)
Mn	3.28 (2.42)	0.17 (0.25)	3.06 (1.35)	0.79 (0.15)	0.06 (0.05)	0.51 (0.28)	1.03 (0.53)	0.28 (0.25)	0.29 (0.3)
Na	5.24 (3.9)	4.92 (3.73)	19.73 (10.24)	1.81 (1.52)	2.81 (1.38)	7.94 (4.86)	7.07 (10.64)	6.38 (3.57)	5.53 (3.97)
Ni	0.15 (0.07)	0.01 (0)	0.16 (0.1)	0.04 (0.03)	0.01 (0)	0.03 (0.02)	0.08 (0.02)	0.02 (0.01)	0.03 (0.02)
Se	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.01 (0.01)	0 (0)	0 (0.01)
Zn	0.22 (0.12)	0 (0)	0.23 (0.23)	0.17 (0.04)	0.01 (0.01)	0.04 (0.05)	0.21 (0.06)	0.02 (0.01)	0.04 (0.06)
TSS	4.67 (4.62)	12 (11.31)	11.45 (15)	2.37 (0)	5.79 (3.04)	14.79 (10.73)	2.37 (0)	6.25 (6.14)	10.75 (12.47)
pH	3.47 (0.23)	7.04 (0.6)	5.78 (0.59)	3.48 (0.09)	7.58 (0.45)	6.99 (0.68)	3.46 (0.38)	7.16 (0.43)	7.34 (0.67)
Cond (mS/cm)	384.95 (179.78)	102.92 (22.76)	304.73 (144.75)	389.67 (35.23)	106.83 (41.06)	309.36 (63.46)	466.33 (195.66)	112.17 (12.29)	337.64 (90.37)
DO	11.23 (0.37)	8.76 (1.89)	15.64 (22.22)	11.25 (0.79)	8.81 (4.25)	10.87 (0.84)	11.21 (0.36)	10.18 (0.4)	10.64 (0.41)
TDS(g/L)	0.28 (0.07)	0.07 (0.02)	0.2 (0.09)	0.25 (0.02)	0.07 (0.03)	0.2 (0.04)	0.3 (0.13)	0.07 (0.01)	0.22 (0.06)

B	2008				2013				2017				2018			
	AMD (n=1)	Treated (n=13)	Reference (n=4)		AMD (n=1)	Treated (n=13)	Reference (n=4)		AMD (n=1)	Treated (n=13)	Reference (n=4)		AMD (n=1)	Treated (n=13)	Reference (n=4)	
Alk	1.75 (n/a)	10.86 (15.52)	60.26 (43.65)		1.24 (n/a)	15.31 (7.84)	28.26 (12.71)		1.24 (n/a)	12.49 (4.11)	28.46 (22.83)		2.7 (n/a)	16.38 (4.94)	27.09 (19.09)	
Acid	36.14 (n/a)	10.37 (8.99)	1.58 (0.47)		27.64 (n/a)	3.72 (3.34)	1.81 (0)		27.04 (n/a)	4.4 (5.11)	1.81 (0)		11.92 (n/a)	2.1 (1.01)	1.81 (0)	
S04	324 (n/a)	310.15 (103.08)	21.53 (6.3)		89.7 (n/a)	136.61 (61.13)	11.71 (1.62)		124 (n/a)	129.52 (47.56)	12.16 (4.42)		102.28 (n/a)	127.93 (46.29)	11.14 (3.1)	
Cl	9.7 (n/a)	4.22 (1.68)	8.05 (5.07)		9.8 (n/a)	5.05 (2.82)	7.28 (5.23)		9.03 (n/a)	7.45 (5.02)	6.01 (4.24)		8.08 (n/a)	7.45 (5.19)	5.97 (5.19)	
Al	4.61 (n/a)	0.69 (1.19)	0.03 (0.01)		1.2 (n/a)	0.04 (0.03)	0.02 (0)		2.07 (n/a)	0.02 (0.01)	0.02 (0)		0.28 (n/a)	0.12 (0.08)	0.05 (0.01)	
Ba	0.02 (n/a)	0.03 (0.01)	0.06 (0.01)		0.02 (n/a)	0.03 (0)	0.04 (0.01)		0.03 (n/a)	0.03 (0)	0.04 (0.01)		0.02 (n/a)	0.03 (0)	0.03 (0.01)	
Ca	61.91 (n/a)	84.2 (18.65)	20.27 (14.39)		15.68 (n/a)	39.41 (16.26)	12.6 (4.84)		27.15 (n/a)	36.42 (11.08)	11.71 (8.56)		25.26 (n/a)	39.27 (12.09)	11.78 (6.97)	
Cd	0 (n/a)	0 (0)	0 (0)		0 (n/a)	0 (0)	0 (0)		0 (n/a)	0 (0)	0 (0)		0 (n/a)	0 (0)	0 (0)	
Co	0.11 (n/a)	0.04 (0.03)	0 (0)		0.03 (n/a)	0.01 (0.01)	0 (0)		0.05 (n/a)	0.01 (0.02)	0 (0)		0.03 (n/a)	0.01 (0.01)	0 (0)	
Cr	0 (n/a)	0 (0.01)	0 (0)		0 (n/a)	0 (0)	0 (0)		0 (n/a)	0 (0)	0 (0)		0 (n/a)	0 (0)	0 (0)	
Cu	0 (n/a)	0 (0.01)	0 (0)		0 (n/a)	0 (0)	0 (0)		0 (n/a)	0 (0)	0 (0)		0 (n/a)	0.01 (0.01)	0 (0)	
Fe	0.31 (n/a)	0.46 (0.68)	0.07 (0.07)		0.07 (n/a)	0.09 (0.15)	0.03 (0.01)		0.14 (n/a)	0.09 (0.19)	0.03 (0.01)		0.05 (n/a)	0.08 (0.13)	0.08 (0.02)	
Mg	24.14 (n/a)	25.35 (6.38)	2.97 (1.4)		7.19 (n/a)	12.29 (5.39)	2.51 (0.52)		12.14 (n/a)	12.53 (4.3)	2.03 (0.79)		10.18 (n/a)	12.28 (4.42)	2.01 (0.56)	
Mn	4.74 (n/a)	2.25 (1.82)	0.07 (0.08)		1.38 (n/a)	0.73 (0.48)	0.02 (0)		2.42 (n/a)	0.81 (0.73)	0.01 (0)		1.83 (n/a)	0.89 (0.64)	0.01 (0)	
Na	3.77 (n/a)	5.13 (2.27)	2.96 (2.96)		4.77 (n/a)	5.72 (1.63)	4.32 (2.96)		6.17 (n/a)	6.27 (2.28)	3.56 (2.77)		4.86 (n/a)	6.05 (2.76)	3.68 (3.45)	
Ni	0.14 (n/a)	0.07 (0.04)	0.01 (0)		0.05 (n/a)	0.03 (0.01)	0.01 (0)		0.07 (n/a)	0.05 (0.09)	0.01 (0)		0.06 (n/a)	0.03 (0.02)	0.02 (0.02)	
Se	0 (n/a)	0.01 (0.02)	0 (0)		0 (n/a)	0 (0)	0 (0)		0 (n/a)	0 (0)	0 (0)		0.03 (n/a)	0.02 (0.01)	0 (0)	
Zn	0.2 (n/a)	0.06 (0.07)	0.01 (0.01)		0.1 (n/a)	0.03 (0.03)	0.01 (0)		0.16 (n/a)	0.12 (0.31)	0.01 (0)		0.11 (n/a)	0.03 (0.02)	0.01 (0)	
TSS	6 (n/a)	20.77 (42.14)	24.5 (37)		4 (n/a)	9.9 (6.48)	15.59 (16.74)		14 (n/a)	6.91 (3.95)	11.09 (8.28)		22 (n/a)	13.72 (6.04)	8.19 (6.71)	
pH	5.15 (n/a)	6.41 (0.48)	7.02 (0.44)		4.22 (n/a)	7.02 (0.45)	7.37 (0.43)		6.17 (n/a)	7.66 (0.34)	7.77 (0.1)		5.98 (n/a)	7.19 (0.11)	6.82 (0.49)	
Cond (mS/cm)	439 (n/a)	298.31 (39.31)	112.5 (71.08)		517 (n/a)	386.85 (90.4)	127 (76.43)		514 (n/a)	531.82 (144.71)	143.75 (93.76)		240 (n/a)	313.08 (92.12)	88.25 (53.95)	
DO	4 (n/a)	8.53 (0.45)	15.21 (7.43)		8.24 (n/a)	9.23 (0.85)	9.48 (1.48)		9.71 (n/a)	9.6 (1.98)	12.74 (3.96)		9 (n/a)	9.44 (1.96)	11.36 (2.45)	
TDS(g/L)	0.43 (n/a)	0.19 (0.03)	0.1 (0.07)		0.34 (n/a)	0.28 (0.07)	0.09 (0.04)		0.33 (n/a)	0.35 (0.09)	0.09 (0.06)		0.16 (n/a)	0.2 (0.06)	0.06 (0.04)	

Works Cited

- Anderson, M., 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* 26, 32–46.
- Barbour, M.T., Gerritsen, J., Snyder, B.D., Stribling, J.B., 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish, Second Edition. US Environ. Prot. Agency Off. Water Washingt. DC 339. [https://doi.org/EPA 841-B-99-002](https://doi.org/EPA%20841-B-99-002)
- Bernhardt, E.S., Palmer, M.A., Allan, J.D., Alexander, G., Barnas, K., Brooks, S., Carr, J., Clayton, S., Dahm, C., Follstad-Shah, J., Galat, D., Gloss, S., Goodwin, P., Hart, D., Hassett, B., Jenkinson, R., Katz, S., Kondolf, G.M., Lake, P.S., Lave, R., Meyer, J.L., O'Donnell, T.K., Pagano, L., Powell, B., Sudduth, O., 2005. Synthesizing U.S. river restoration efforts. *Science* (80-.). <https://doi.org/10.1126/science.1109769>
- Brown, B.L., Swan, C.M., Auerbach, D.A., Campbell Grant, E.H., Hitt, N.P., Maloney, K.O., Patrick, C., 2011. Metacommunity theory as a multispecies, multiscale framework for studying the influence of river network structure on riverine communities and ecosystems. *J. North Am. Benthol. Soc.* 30, 310–327. <https://doi.org/10.1899/10-129.1>
- DeNicola, D.M., Stapleton, M.G., 2016. Using macroinvertebrates to assess ecological integrity of streams remediated for acid mine drainage. *Restor. Ecol.* 24, 656–667. <https://doi.org/10.1111/rec.12366>
- DeNicola, D.M., Stapleton, M.G., 2002. Impact of acid mine drainage on benthic communities in streams: The relative roles of substratum vs. aqueous effects. *Environ. Pollut.* 119, 303–315. [https://doi.org/10.1016/S0269-7491\(02\)00106-9](https://doi.org/10.1016/S0269-7491(02)00106-9)
- Diaz, R.J., Rosenberg, R., 2008. Spreading dead zones and consequences for marine ecosystems. *Science* (80-.). 321, 926–929. <https://doi.org/10.1126/science.1156401>
- Gerhardt, A., Janssens De Bisthoven, L., Soares, A.M.V.M., 2004. Macroinvertebrate response to acid mine drainage: Community metrics and on-line behavioural toxicity bioassay. *Environ. Pollut.* 130, 263–274. <https://doi.org/10.1016/j.envpol.2003.11.016>
- Hogsden, K.L., Harding, J.S., 2012. Consequences of acid mine drainage for the structure and function of benthic stream communities: a review. *Freshw. Sci.* 31, 108–120. <https://doi.org/10.1899/11-091.1>
- Idígoras Chaumel, A.L., Armanini, D.G., Schwindt, J.A., Yates, A.G., 2019. Interannual Variation of Benthic Macroinvertebrate Communities at Long-Term Monitoring Sites Impacted by Human Activities: Implications for Bioassessment. *Diversity* 11, 167. <https://doi.org/10.3390/d11090167>
- Kauffman Jr., G.J., 2010. The Delaware River revival: Four centuries of historic water quality change from Henry Hudson to Benjamin Franklin to JFK. *Pa. Hist.* 77.
- Keener, A.L., Sharpe, W.E., 2005. The effects of doubling limestone sand applications in two acidic southwestern Pennsylvania streams. *Restor. Ecol.* 13, 108–119. <https://doi.org/10.1111/j.1526-100X.2005.00013.x>
- Leps, M., Sundermann, A., Tonkin, J.D., Lorenz, A.W., Haase, P., 2016. Time is no healer: Increasing restoration age does not lead to improved benthic invertebrate communities in restored river reaches. *Sci. Total Environ.* 557–558, 722–732. <https://doi.org/10.1016/j.scitotenv.2016.03.120>

- Long, R.A., 2019. Ecological Benefits of Watershed-scale Restoration in Two Intensively Mined Cold- and Warm-water Ecosystems. West Virginia University.
- McClurg, S.E., Petty, J.T., Mazik, P., Clayton, J.L., 2007. Stream Ecosystem Response to Limestone Treatment in Acid Impacted Watersheds of the Allegheny Plateau. *Ecol. Appl.* 17, 1087–1104.
- McCune, B., Urban, D., Grace, J., 2002. Data Transformations, in: *Analysis of Ecological Communities*. pp. 67–79.
- Merovich, G.T., Petty, J.T., 2007. Interactive effects of multiple stressors and restoration priorities in a mined Appalachian watershed. *Hydrobiologia* 575, 13–31. <https://doi.org/10.1007/s10750-006-0277-y>
- Merovich, G.T., Petty, J.T., Strager, M.P., Fulton, J.B., 2013. Hierarchical classification of stream condition: a house–neighborhood framework for establishing conservation priorities in complex riverscapes. *Freshw. Sci.* 32, 874–891. <https://doi.org/10.1899/12-082.1>
- Merriam, E.R., Petty, J.T., 2016. Under siege: Isolated tributaries are threatened by regionally impaired metacommunities. *Sci. Total Environ.* 560–561, 170–178. <https://doi.org/10.1016/j.scitotenv.2016.04.053>
- Merritt, R., Cummins, K.W., Berg, M., 2008. *An Introduction to the Aquatic Insects of North America*, Fourth. ed. Kendall Hunt Publishing, Dubuque, Iowa.
- Mulvihill, R.S., Newell, F.L., Latta, S.C., 2008. Effects of acidification on the breeding ecology of a stream-dependent songbird, the Louisiana waterthrush (*Seiurus motacilla*). *Freshw. Biol.* 53, 2158–2169. <https://doi.org/10.1111/j.1365-2427.2008.02032.x>
- Oksanen, J., Guillaume Blanchet, K., Legendre, P., Minchin, P., O’Hara, R., Simpson, G., Solymos, P., Henry, M., Stevens, H., Wagner, H., n.d. *vegan: Community Ecology Package*.
- Palmer, M.A., Hondula, K.L., Koch, B.J., 2014. Ecological Restoration of Streams and Rivers: Shifting Strategies and Shifting Goals. *Annu. Rev. Ecol. Evol. Syst.* 45, 247–269. <https://doi.org/10.1146/annurev-ecolsys-120213-091935>
- Pond, G.J., Bailey, J.E., Lowman, B.M., Whitman, M.J., 2013. Calibration and validation of a regionally and seasonally stratified macroinvertebrate index for West Virginia wadeable streams. *Environ. Monit. Assess.* 185, 1515–1540.
- Rodríguez-Gironés, M.A., Santamaría, L., 2006. A new algorithm to calculate the nestedness temperature of presence-absence matrices. *J. Biogeogr.* 33, 924–935. <https://doi.org/10.1111/j.1365-2699.2006.01444.x>
- Sarremejane, R., Mykrä, H., Huttunen, K.L., Mustonen, K.-R., Marttila, H., Paavola, R., Sippel, K., Veijalainen, N., Muotka, T., 2018. Climate-driven hydrological variability determines inter-annual changes in stream invertebrate community assembly. *Oikos* 127, 1586–1595. <https://doi.org/10.1111/oik.05329>
- Shrestha, R.K., Stein, T. V., Clark, J., 2007. Valuing nature-based recreation in public natural areas of the Apalachicola River region, Florida. *J. Environ. Manage.* 85, 977–985. <https://doi.org/10.1016/j.jenvman.2006.11.014>
- Stoll, S., Breyer, P., Tonkin, J.D., Früh, D., Haase, P., 2016. Scale-dependent effects of river habitat quality

- on benthic invertebrate communities - Implications for stream restoration practice. *Sci. Total Environ.* 553, 495–503. <https://doi.org/10.1016/j.scitotenv.2016.02.126>
- Stradling, D., Stradling, R., 2008. Perceptions of the burning river: Deindustrialization and Cleveland's Cuyahoga River. *Environ. Hist. Durh. N. C.* 13, 515–535. <https://doi.org/10.1093/envhis/13.3.515>
- USEPA, 2017. United States - Environmental Protection Agency. Water quality standards handbook: Chapter 3: Water quality criteria. EPA-823-B-17-001.
- Vannote, R.L., Minshall, G.W., Cummins, K.W., Seddell, J.R., Cushing, C.E., 1980. The River Continuum Concept. *Can. J. Fish. Aquat. Sci.* 37, 130–137. <https://doi.org/10.1139/f80-017>
- Wallace, J.B., Webster, J.R., 1996. The Role of Macroinvertebrates in Stream Ecosystem Function. *Annu. Rev. Entomol.* 41, 115–139. <https://doi.org/10.1146/annurev.en.41.010196.000555>
- Walter, C.A., Nelson, D., Earle, J.I., 2012. Assessment of stream restoration: Sources of variation in macroinvertebrate recovery throughout an 11-year study of coal mine drainage treatment. *Restor. Ecol.* 20, 431–440. <https://doi.org/10.1111/j.1526-100X.2011.00845.x>
- Watson, A.S., Merovich, G.T., Petty, J.T., Gutta, J.B., 2017. Evaluating expected outcomes of acid remediation in an intensively mined Appalachian watershed. *Environ. Monit. Assess.* 189. <https://doi.org/10.1007/s10661-017-6036-x>
- Westveer, J.J., Van Der Geest, H.G., Emiel van Loon, E., Verdonschot, P.F.M., 2018. Connectivity and seasonality cause rapid taxonomic and functional trait succession within an invertebrate community after stream restoration. *PLoS One* 13, 1–17. <https://doi.org/10.1371/journal.pone.0197182>
- Wijbenga, A., Hutzinger, O., 1984. Chemicals, man and the environment. *Naturwissenschaften* 71, 239–246. <https://doi.org/10.1007/bf00441332>
- WVDEP, 2009. Standard Operating Procedures. West Virginia Dep. Environ. Prot. Charleston, West Virginia.
- WVDNR, 2019. West Virginia Division of Natural Resources 2018-2019 Annual Report.