Indications of benthic macroinvertebrate assemblage recovery following wastewater treatment upgrades in the central Grand River

by

Sean McLay

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Master of Science

in

Biology

Waterloo, Ontario, Canada, 2022

Author’s Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Treated effluent from wastewater treatment plants (WWTPs) can be a source of substantial nutrient and contaminant loading, altering water chemistry and ecological conditions in receiving waters. To meet new regulations and increasing demand, the Region of Waterloo invested in major upgrades to the Kitchener and Waterloo WWTPs, the two largest WWTPs in the Region, over the past decade. As a part of efforts to monitor the effectiveness of these investments, the Region of Waterloo (RoW) initiated a benthic macroinvertebrate (BMI) monitoring program sampling upstream and downstream of effluent outfalls. Using data provided by the RoW of that was collected by consultants in the fall every three years between 2009 to 2018, the response of the BMI assemblage to upgrades at the Waterloo and Kitchener WWTPs was assessed using multivariate analysis on the entire assemblage and univariate analysis on metrics such as family richness. Furthermore, BMI assemblage data was collected at five other WWTPs in the RoW (Preston, Hespeler, Wellesley, New Hamburg, and Ayr), allowing for their downstream effluent effects to be characterized. Additional sites were sampled separate from the RoW's sampling program as an independent characterization of BMI assemblages along the central Grand River. Spatial and temporal analyses of BMI suggested that the impacts of the Waterloo and Kitchener WWTPs are reducing over time, with upstream and downstream assemblages becoming more similar following upgrades, as evident in the decreasing of average Bray-Curtis dissimilarity in association with upgrades and increased similarity in metrics such as family richness between assemblages. These findings support that WWTP upgrades are succeeding in reducing the influence of effluent discharged into the receiving waters. With some upgrades to WWTPs occurring within a year prior to the last sampling year (2018), and delays in response to amelioration of stressors sometimes observed in BMI assemblages, recovery is likely still ongoing. At three of the five other WTTPs whose downstream impact was assessed (i.e., Hespeler, New Hamburg, and Ayr), there was dissimilarity between upstream and downstream BMI assemblages, although further study is required to confirm the stressors associated with these ecological responses, including natural variability. Continued monitoring using the existing Before-After/Control-Impact (BACI) design to isolate effluent effects from the interannual variation apparent in all analyses is therefore recommended to capture potential effects and/or further recovery.

Acknowledgements

A huge thank you to Dr. Mark Servos and Dr. Adam Yates for your mentorship, providing ample opportunity to learn and grow professionally. You were both so generous in your support and took every occasion to help me however you could, and for that I am very grateful. And thank you to my committee members Dr. Keegan Hicks and Dr. Paul Craig for your thoughtful guidance and expertise.

Thank you to Dominika Celmer-Repin and the Region of Waterloo for providing data and reports that were essential to my project, as well answering the many questions and requests I have had over the past two years.

Thank you to Dr. Edward Krynak for teaching me how to process my samples and providing insight to critical parts of my analyses. Thanks again to Dr. Keegan Hicks and Dr. Edward Krynak for allowing me to use the data that you both worked so hard to collect to support my thesis.

Thank you to Amory Conant, Erika Burton, Jaki Peters, Katelyn Gao, and especially Hadi Dhiyebi for helping with field sampling, toiling in the river while I played with the nets on shore. And thank you to Leslie Bragg for helping me with all the logistics of my project, be it shipping, ordering, etc., and always being available to answer questions. I also want to thank the Servos and StrEAMS labs for being a much needed social and collaborative outlet, online and in-person.

And thank you to my friends, family, and partner – I appreciate all the love and support I have received over the past two years and hope I make you proud.

Table of Contents

List of Figures

List of Tables

1.0 Introduction

In urbanized environments, effluent from wastewater treatment plants (WWTPs) is among the most ubiquitous stressors impacting stream health (Paul & Meyer, 2001). Effluent discharged from WWTPs can have pronounced effects on receiving waters by exerting strong influences on ecosystem processes such as nutrient cycling (Carey & Migliaccio, 2009), with effluent sometimes comprising a substantial portion of total nutrient loads in streams (Marti et al., 2004). Nutrient loading can influence the entire ecosystem, potentially increasing microbial activity as well as fungal and algal biomass (Gulis & Suberkropp, 2003), and reducing oxygen levels via increased organic matter decomposition (Carpenter et al., 1998; Poulton et al., 2015) and respiration (Venkiteswaran et al., 2015). With the consequences of effluent release well established, upgrading wastewater treatment plants to reduce nutrient and contaminant loading to the environment has become the focus of many remedial management plans.

Upgrades to WWTPs generally have the purpose of improving environmental quality by reducing nutrient loading (Gücker et al., 2006; Besley & Chessman, 2008), biochemical oxygen demand (Crawford et al., 1992), and organic micropollutants like biocides and pharmaceuticals (Ashauer, 2016). Infrastructure and process upgrades have been shown to successfully improve water quality, typically increasing dissolved oxygen (DO) and reducing ammonium concentrations (Crawford et al., 1992; Johnson et al., 2019). However, even tertiary treatment of effluent sometimes is insufficient to prevent eutrophication in some environments (Gücker et al., 2006). In addition, most treatment systems are not equipped to remove the diversity of chemicals that are now entering these systems, including pesticides, pharmaceuticals, and personal care products (Blair et al., 2013; Loos et al., 2013). With minimal funding for wastewater infrastructure (UN-Water, 2015) and new threats emerging continuously, it is paramount that the limited funding available is well allocated and that the effectiveness of upgrades in achieving intended environmental quality improvements in receiving waters are verified.

1.1 Assessing Wastewater Effluent Effects

One approach to investigate the success of wastewater treatment upgrades is using benthic macroinvertebrates (BMI). Benthic macroinvertebrates are aquatic, bottom-dwelling organisms that lack vertebrae, such as worms, molluscs, and insects. Some BMI are more tolerant to stream pollution, while others are more sensitive, with those that are more sensitive potentially being lost following exposure to environmental stressors (Hilsenhoff, 1977). Exposure to environmental stressors, such as wastewater effluent, can therefore lead to important changes in species assemblages in the downstream receiving environment (Mor et al., 2019). These changes in assemblage are important as BMI play a key role in many ecosystem processes, such as nutrient cycling and decomposition (Wallace & Webster, 1996). Benthic macroinvertebrates can be effective bioindicators of ecosystem health (Hilsenhoff, 1977) as they reflect local conditions over an extended period of time due to their relatively long life cycles and a sedentary nature. Given their ecological importance and inherent suitability as a bioindicator, they are often utilized to provide insight into the impacts of wastewater treatment upgrades on aquatic ecosystems (Ashauer, 2016; Arce et al., 2014; Crawford et al., 1992).

There have been a variety of observed impacts on BMI resulting from stressors contained in WWTP effluent released into rivers that can be used to garner insight on the effects of WWTP upgrades. Nutrient enrichment resulting from the release of effluent could result in the loss of sensitive species and thereby reduce taxa richness (Paul & Meyer, 2001). Changes in assemblage have been observed in BMI exposed to wastewater effluent, with decreases in the proportion of pollution-sensitive species (Grantham et al., 2012), and increases in tolerant species (Ortiz & Puig, 2007). Low DO concentrations as a consequence of eutrophication can be a stressor for some BMI (Sundermann et al., 2013), potentially causing changes in assemblage (Spänhoff et al., 2007). Oxygen-sensitive organisms, including particular mayfly species, can be reduced in proportion downstream of minimally treated wastewater outfalls (Suckling, 1982). Oxygen depletion can be further compounded by the release of ammonium and nitrite, which undergo nitrification and consume oxygen during transformation, resulting in reduced DO (Sánchez-Murillo et al., 2014). Additionally, BMI have traits that describe a taxon's morphology, physiology, life history, and dispersal ability, which dictate its ability to survive the conditions of a given habitat (Allan et al., 2021). As release of wastewater effluent has the ability to change habitat and water quality conditions, the presence and proportion of various BMI traits have been used to assess the effects of wastewater effluent exposure (Arce et al., 2014; Lecerf et al., 2006). One example of a trait that has been used to assess WWTP impacts is feeding type, with the relative abundance of different feeding types such as scrapers (mainly feed on periphyton) and collector-filterers (suspension feeders of mainly detritus; Allan et al., 2021). The proportion of these feeding types may change with exposure to effluent release as nutrient loading can affect the availability of food sources such as biofilm and fine particulate organic matter, potentially causing shifts in the BMI assemblage (Arce et al., 2014). As WWTP upgrades change the composition of effluent released, BMI assemblages very well may shift in response.

Benthic macroinvertebrates have been used to assess the effects of the upgrading or decommissioning of WWTPs in several past studies. For instance, Crawford et al. (1992) observed a shift in assemblage from mainly pollution-tolerant (e.g. oligochaetes) to more pollution-sensitive species (e.g. Hydropsychidae) following the implementation of advanced wastewater treatment at a WWTP in Indiana, United States. Likewise, after wastewater discharge cessation in France’s Vistre River, Arce et al. (2014) found a reduction in the proportion of species with polyvoltinism and an increase in species richness, likely due to release from effluent-induced stresses such as low DO and eutrophication. Additionally, Besley & Chessman (2008) found that BMI assemblages upstream and downstream of effluent release in the Blue Mountains, Australia became more similar following transferral of effluent treatment from five small WWTPs to a larger, more efficient WWTP.

Wastewater effluent is often not the only anthropogenic stressor present in aquatic systems. Inputs from agriculture can further increase nutrient loading via N and P, which combined with changes in hydrology, habitat conditions, and other contaminants (Walsh et al., 2005), have led to increasing proportions of BMI that are generalists, shorter lived, and more tolerant of stressors and their effects (Allan, 2004). If agricultural stressors are substantial enough, they may shift the BMI assemblage in a similar manner to municipal effluent release, making it difficult to discern whether effluent has an impact. With other stressors potentially confounding the effects of WWTP effluent, it may be difficult to detect a change related to treatment upgrades implemented. Therefore, to better isolate effluent effects, it is important to control for other stressors present in the system through robust study design.

Incorporating a reference (upstream) site to compare to an impacted (downstream) site, and observing how this comparison changes over time, aids in detecting the response to changes in a stressor in a multi-stressor environment. In a Before-After/Control-Impact (BACI) design, samples are collected before and after the onset of an environmental stressor, both at control sites uninfluenced by the stressor of interest (e.g. municipal wastewater outfall) and impacted sites directly affected by it (Green, 1979). This allows for a comparison of conditions before and after, detecting differences attributable to changes in the stressor occurring between, and controlling for natural variation and other anthropogenic stressors. This study design isn't limited to detecting detriments resulting from stressors, but also benefits resulting from their reduction, as may be the case with wastewater treatment upgrades.

The BACI design has been employed in many studies investigating upgrades to wastewater treatment processes (Arce et al., 2014; Ashauer, 2016; Besley & Chessman, 2008; Crawford et al., 1992; Johnson et al., 2019). Some studies had multiple upstream sites (Johnson, 2019), or multiple wastewater treatment plants in the same reach (Arce et al., 2014; Johnson et al., 2019), however the upstream vs. downstream comparison was still used to tease out effluent-associated effects. One study included comparisons of effluent-impacted sites to least-impaired river reaches to determine the recovery seen in impacted habitats to similar habitats closer to reference/pristine conditions (Arce et al., 2014). However, the authors acknowledged that singular actions to improve local water quality will not return assemblages subject to multiple anthropogenic stressors to a state similar to unimpacted reaches. Therefore, a focus on comparisons between downstream, effluent-exposed BMI assemblages and upstream, non-exposed assemblages will still provide insight into the effectiveness of WWTP upgrades. Overall, these studies have provided evidence that the BACI design is well-established and can be effectively used to capture the effects of WWTP upgrades using BMI.

1.2 Wastewater Treatment in the Central Grand River

In the Region of Waterloo, Ontario, Canada, a 2007 study assessed what could be done to accommodate new wastewater treatment regulation requirements and a growing population (Region of Waterloo, 2007). Of primary concern were two major WWTPs, Kitchener and

Waterloo, located in the central portion of the Grand River (hereafter referred to as the "central Grand River") within the Region. Together, these two WWTPs discharged over 100,000 m^3/day of effluent into the Grand River in 2007 and have continued to into 2018 (Region of Waterloo, 2020). To meet regulations, increasing demand, and improve effluent quality, wastewater treatment upgrades have occurred over the past decade at the Kitchener and Waterloo WWTPs, costing hundreds of millions of dollars (Hicks, 2017).

Upgrades at both the Waterloo and Kitchener WWTPs were completed with the purpose of improving effluent quality, including the reduction of ammonia loading to receiving waters. The Kitchener WWTP began upgrades in 2012, increasing nitrification at plant 1 and providing full nitrification at plant 2 (**Table 1**), which were the major upgrades completed. These upgrades have had a marked effect on the water chemistry of effluent released from the Kitchener WWTP, with a large reduction in ammonia concentrations and a concomitant increase in nitrate following the implementation of full nitrification in January of 2013 (Hicks et al., 2017b; **Figure 1**). Another, more minor upgrade occurred in 2017, adding tertiary phosphorus filters to plant 2, reducing P loading (**Figure 2**). The Waterloo WWTP underwent similar upgrades (**Table 2**), starting at only partial nitrification (Nikel, 2020) with the aim of ultimately achieving full nitrification (Region of Waterloo, 2018). However, construction delays impeded treatment function, resulting in reduced nitrification from 2009 until 2015 when Return-Activated Sludge (RAS) re-aeration came online (Srikanthan, 2019). The effluent quality continued to improve with the implementations of the first aeration tanks in 2017 and the second aeration tank upgrade coming in March 2018, resulting in full, year-round nitrification (Srikanthan, 2019). Similar to the Kitchener WWTP, the water chemistry of effluent discharged from the Waterloo WWTP exhibited notable changes following wastewater treatment process upgrades. The ammonia concentration decreased and nitrate increased following the completion of the RAS re-aeration in 2015, continuing until full nitrification was reached in March of 2018 (Nikel, 2020; **Figure 3**). In parallel, Srikanthan (2019) showed that other key contaminants, such as estrogens, declined with the onset of nitrification at both Waterloo and Kitchener.

River water chemistry, such as ammonia and nitrate, was collected as a part of the Region of Waterloo's Surface Water Quality Monitoring Program (Region of Waterloo, 2018). The purpose of this program was to determine the impact of ten WWTPs and their upgrades in the

Region of Waterloo, including the Kitchener and Waterloo WWTPs. Not only was water chemistry collected as a part of this program, but routine BMI sampling occurred as well. Water chemistry data has been collected throughout the year since 2007, and BMI samples have been collected in the autumn every three years from 2009 to 2018, both by LGL Limited (except from 2007 to early 2009 where the Grand River Conservation Authority collected water chemistry). This BMI data spanning from before and after major upgrades provides an opportunity to assess the effectiveness of upgrades implemented to the Kitchener and Waterloo WWTP. In addition, the Surface Water Quality Monitoring Program collected BMI assemblages at other WWTPs in the Region of Waterloo as well. This includes the Preston, Hespeler, Wellesley, New Hamburg, and Ayr WWTPs, all of which have very similar data available as the Kitchener and Waterloo WWTPs. However, unlike with Kitchener and Waterloo, these other WWTPs have incurred little to no upgrades throughout the study period. This data thereby allows for the investigation of BMI assemblages downstream of each WWTP outfall to detect any impacts that may be occurring.

Table 1. Overview of upgrades implemented to the Kitchener wastewater treatment plant, including what was upgraded, the date the upgrade was completed, and its intended result on final effluent (Srikanthan, 2019; Nikel, 2020).

Upgrade	**Date Completed**	**Intended Result**
New tanks/aeration – Plant 2 – Pass 1	August 2012	Increased nitrification
New aeration – Plant 1	October 2012	Increased nitrification
New tanks/aeration – Plant 2 – Passes 2 and 3	January 2013	Full nitrification at plant 2
New Diffuser System	November 2016	Better dispersed effluent across the channel
Tertiary filters commissioned (Plant 1 not being filtered)	August 2017	Decreased phosphorus load
Optimized coagulant dose for TP control	October 2017	Decreased phosphorus load
Plant 3 online	2020	
Plant 4 online	2020	Provide full nitrification and tertiary phosphorus treatment plant-wide
Plant 1 offline	2020	

Table 2. Overview of upgrades implemented to the Waterloo wastewater treatment plant, including what was upgraded, the date the upgrade was completed, and its intended result on final effluent (Srikanthan, 2019; Nikel, 2020).

Upgrade	Date Completed	Intended Result
RAS re-aeration online*	2015	Increased nitrification
Aeration tank 1 upgrades	March 2017	Increased nitrification
Aeration tank 2 online	March 2018	Full nitrification

* Interim dewatering occurred from 2009-2014 due to construction delays, resulting in reduced nitrification until RAS re-aeration came online

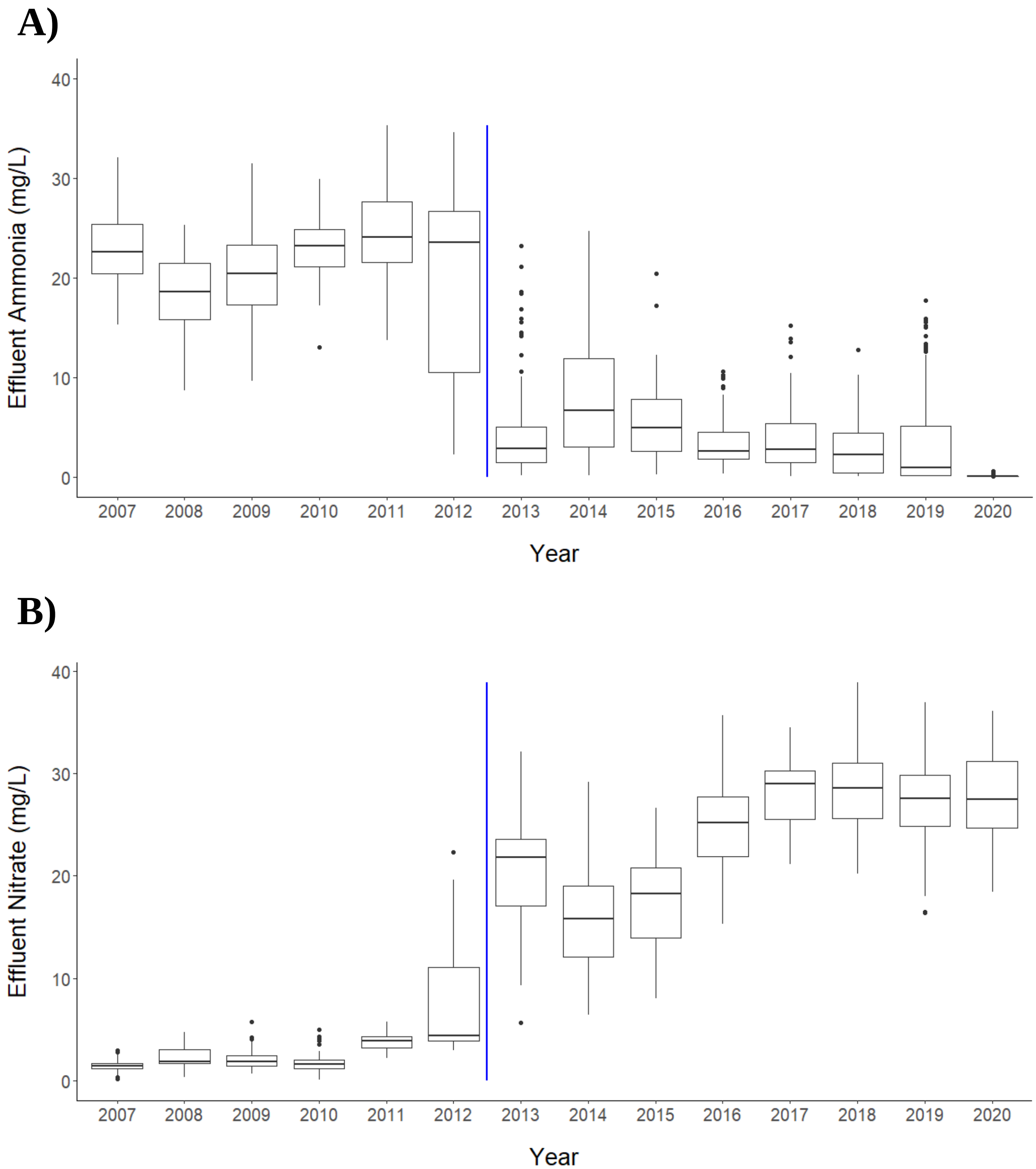

Figure 1. Boxplots showing range of effluent ammonia (A) and nitrate (B) concentrations measured approximately weekly at the Kitchener wastewater treatment plant (Dominika Celmer-Repin, Region of Waterloo, personal communication). The line within boxplots represents the median. The blue line indicates the implementation of the major nitrification upgrades.

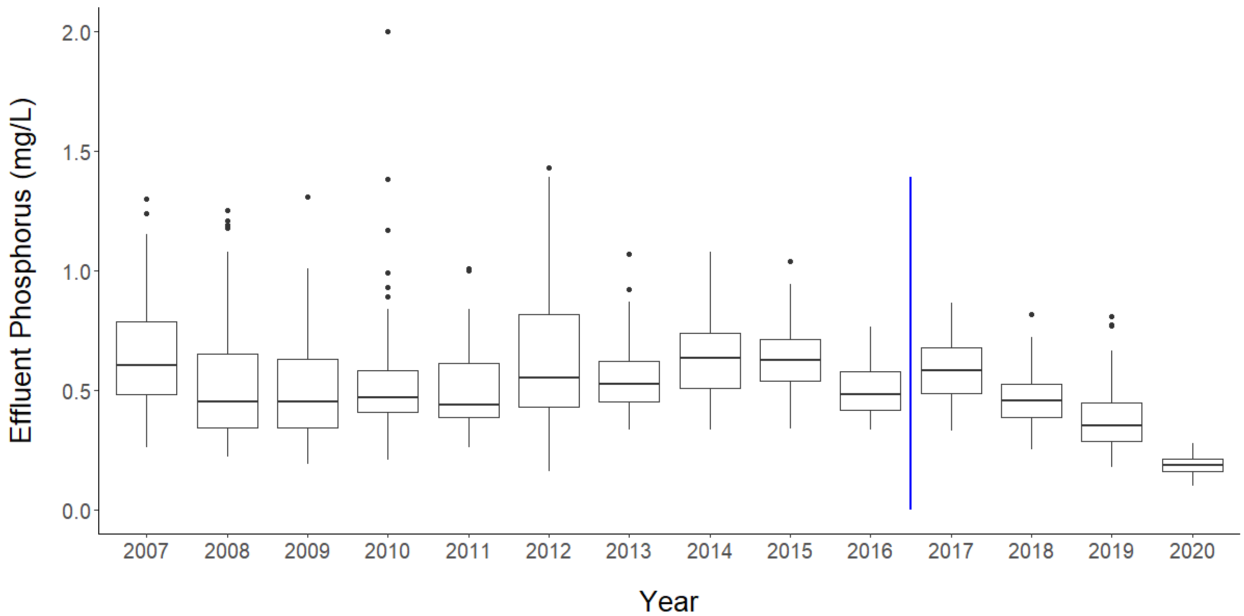

Figure 2. Boxplots of effluent phosphorus concentrations measured approximately weekly at the Kitchener wastewater treatment plant (Dominika Celmer-Repin, Region of Waterloo, personal communication). The line within boxplots represents the median. The blue line indicates the implementation of tertiary phosphorus filters and optimization of the coagulation dose for total phosphorus control.

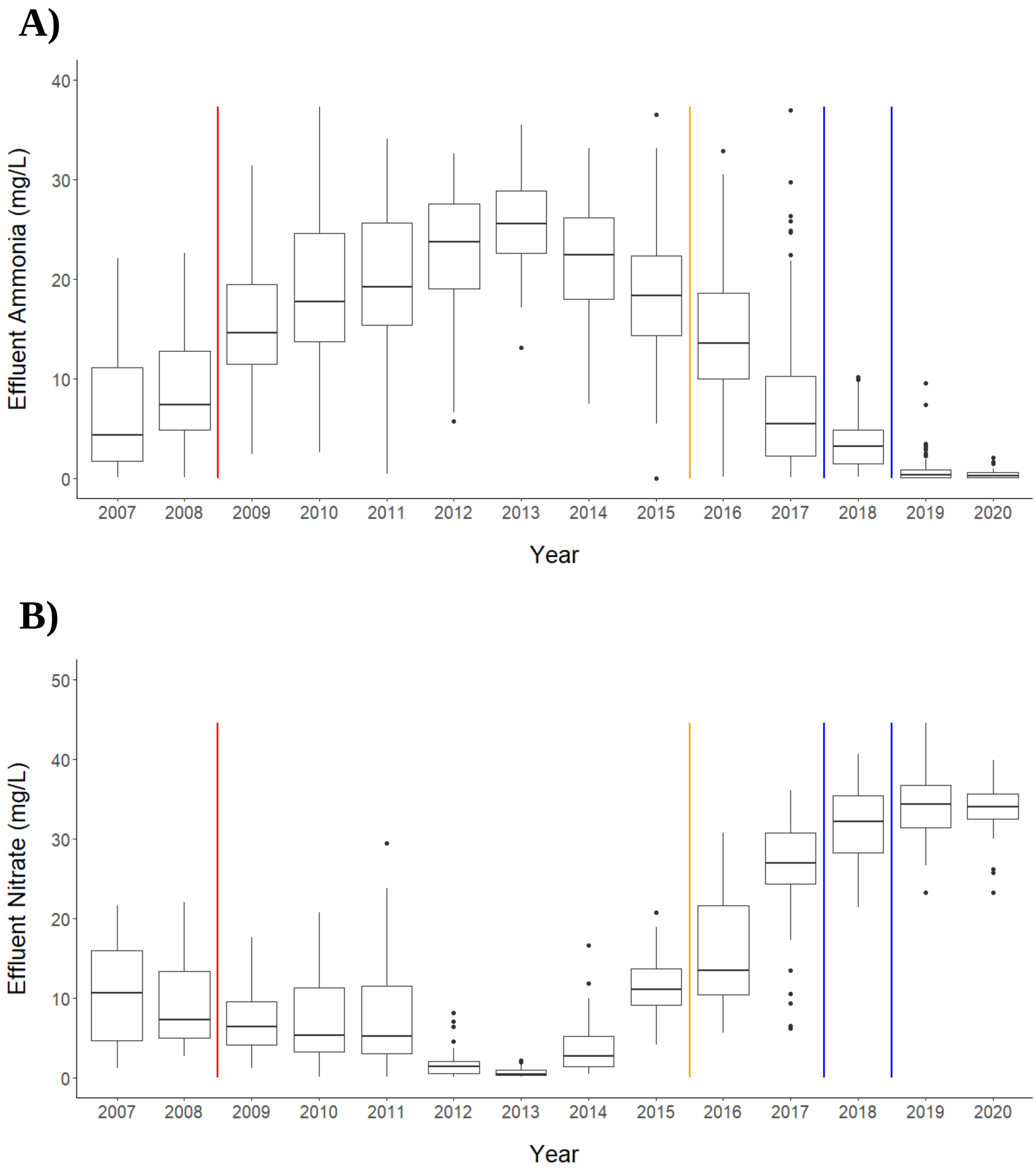

Figure 3. Boxplots of effluent ammonia (A) and nitrate (B) concentrations measured approximately weekly at the Waterloo wastewater treatment plant (Dominika Celmer-Repin, Region of Waterloo, personal communication). The line within boxplots represents the median. The red line indicates the beginning of construction, orange line indicates the completion of construction, and dark blue lines indicate the implementation of upgrades.

1.3 Previous Wastewater Studies in the Grand River

Much of the research regarding wastewater effects and treatment upgrades in the Grand River regarding biota revolve around fish, especially the rainbow darter (*Etheostoma caeruleum*). Previous investigations have observed changes in stable isotope ratios due to nutrient cycling (Hicks et al., 2017b) and increased incidence of intersex (presence of testes and ovaries) in male rainbow darters caused by endocrine-disrupting compounds sourced from effluent (Tetreault et al., 2011, Bahamonde et al., 2014; Fuzzen et al., 2016; Hicks et al., 2017a). Recovery of these endpoints were used to assess changes resulting from wastewater treatment upgrades. After major upgrades at the Kitchener WWTP came online in 2012/2013, there was a major reduction in the incidence of intersex, showing a notable change associated with treatment alterations (Hicks et al., 2017a). Upgrades at the Kitchener WWTP were also associated with the $\delta^{15}N$ isotopic signature of primary consumers and rainbow darters (Hicks et al., 2017b) as well as *in vitro* 11-ketotestosterone production (Marjan et al., 2018) more closely resembling those observed upstream. Most recently, similar, but more muted responses were seen in these same endpoints for rainbow darters (isotopic signature, intersex, steroid production) associated with the Waterloo WWTP upgrades, with downstream sites more closely resembling reference sites (Nikel, 2020). While relating upgrades to changes in rainbow darters downstream of the Waterloo WWTP proved difficult due to variability in endpoints in both impacted and reference sites, there have nevertheless been shifts in biota following upgrades.

While most research surrounding WWTPs in the Grand River watershed have focused on the rainbow darter, there has also been some research regarding other organisms. One study looked into wastewater effects on the isotopic signatures of macrophytes downstream of both the Waterloo and Kitchener WWTPs (Hood et al., 2014). Macrophytes sampled in 2007, 2008, and 2009 downstream of the WWTPs and within the effluent plume incorporated primarily effluent-derived NH_4^+ which increased with distance from the outfall, exhibiting the effects the WWTPs exerted on some primary producers. Moreover, a study investigated stable isotopes in aquatic invertebrates and fish sampled downstream of the Waterloo, Kitchener, and Guelph WWTPs in May, July, and September of 2007 (Loomer et al., 2015). Nitrogen pollution sourced from the WWTPs appeared to make $\delta^{15}N$ values more variable both spatially and temporally downstream of effluent release. Additionally, there has been research on ammonia-oxidizing bacteria downstream of the Waterloo WWTP, with ammonia-oxidizing bacteria more abundant within the effluent plume compared to ammonia-oxidizing archaea being more abundant outside of the plume in 2010 (Sonthiphand et al., 2013). Furthermore, a study on caged freshwater mussels downstream of the Kitchener WWTP showed clear effluent effects, with signs of oxidative stress, possible estrogenicity in males, and stimulated immune response (Gillis et al., 2014). Population effects have also been studied, with a 98% decrease in abundance of freshwater mussels downstream of the Hespeler WWTP outfall compared to upstream (Gillis et al., 2017a). Most pertinently, a study on freshwater mussels investigated the effects of wastewater treatment upgrades at the Kitchener WWTP, sampling Unionid mussels in 2012 before major upgrades were implemented, and after in 2014 (Gillis et al., 2017b). Mussel surveys in 2012 demonstrated the impact of effluent release, with mussels found at a relatively high frequency within 2 km upstream of the Kitchener WWTP, but none found for 7 km downstream of the outfall. In the 2014 survey, while live mussels were found about 1 km downstream of the outfall, these were between 4-8 years old, indicating they were either also present in 2012 or were transported downstream during a high flow event. Although there was no evidence that upgrades modified the effect effluent had on freshwater mussels downstream of the outfall, the authors posited that recovery may take decades as a result of their sedentary nature and extended life cycle reliant on host fish, necessitating long-term monitoring.

Although not extensively studied, there has been some investigation of fish and BMI assemblages in the Grand River in relation to effluent effects. The fish assemblage upstream and downstream of both the Kitchener and Waterloo WWTPs were sampled in September of 2007 and 2008, as well as November of 2008 to assess their health status and characterize WWTP impacts (Tetreault et al., 2013). This investigation into fish assemblages found that assemblages downstream of WWTP outfalls differed in metrics such as abundance, diversity, and family/species richness relative to upstream assemblages and those further downstream. Additionally, to assess if the effects seen in rainbow darters following upgrades to the Kitchener and Waterloo WWTPs were indicative of the entire fish assemblage, collections of both fish and, for comparative purposes, BMI, were completed (Hicks, 2017). The fish assemblage followed a spatial gradient along the central Grand River, whereas patterns in BMI assemblages were more associated with environmental variables such as water quality, which is affected by effluent release. This association was apparent as BMI assemblages downstream of the WWTP shifted in relation to the poorer water quality relative to upstream. While some fish, such as rainbow darters, can act as a sentinel species, the BMI assemblage appeared to be both a more sensitive indicator of wastewater effects and reflective of local water quality conditions in the central Grand River (Hicks, 2017). This study has shown that BMI respond to the effects of wastewater treatment effluent in the central Grand River and provides insight into the need for further investigating the effectiveness of upgrades (Hicks, 2017).

1.4 Research Objectives and Hypotheses

The impact of wastewater on ecosystem health is a major concern in the Region of Waterloo. Two of the largest WWTPs in the watershed are located in the central Grand River and have recently undergone major upgrades. Although upgrades to the Kitchener and Waterloo WWTPs have altered the impact of effluent on their receiving waters, additional confirmations must be made to verify whether these changes reduce impacts in the biota therein. The Region of Waterloo has provided an opportunity to investigate these impacts having maintained water quality and BMI collections at each of the WWTPs in the Region since 2009, prior to the upgrades. Although the data has been collected and reported to the Region of Waterloo, it has not been explored fully. In addition, a longitudinal survey, using kick nets with additional sites, was completed to put the variability into context and independently determine patterns in 2020. By looking at the effects of WWTP upgrades on BMI assemblages over space and time using a BACI approach, the responses can be better elucidated in the context of interannual variability. This data presents a unique opportunity to determine if macroinvertebrate assemblages have changed over time across the Region of Waterloo in association with WWTP upgrades and/or other environmental changes. This thesis addresses three key questions**:**

1. *Do benthic macroinvertebrate assemblages change longitudinally along the central Grand River in 2020; how does this compare to previous similar collections?*
2. *Is there a difference between benthic macroinvertebrate assemblages before and after upgrades were completed at the Waterloo and Kitchener wastewater treatment plants based on the Region of Waterloo collections?*

3. *Are there differences in benthic macroinvertebrate assemblages upstream and downstream of other WWTPs in the Region of Waterloo?*

It is expected that BMI assemblages in the central Grand River will follow a longitudinal pattern, with sites adjacent in the network being most similar to each other, but influenced by point source inputs such as WWTPs. Moreover, it is expected that downstream BMI assemblages and metrics describing them for both the Waterloo and Kitchener WWTPs will have become more similar to upstream BMI following the implementation of upgrades. Finally, it is expected that the BMI assemblage response associated with the outfalls of other WWTPs in the Region of Waterloo will be site specific.

2.0 Methods

2.1 Study Area

All sample collections were made in the Region of Waterloo (**Figure 4A**), which is located in southwestern Ontario, Canada, within the Grand River watershed of the Great Lakes basin. The Region includes the cities of Waterloo, Kitchener, and Cambridge, and the townships of Woolwich, Wellesley, Wilmot, and North Dumfries. Within the Region of Waterloo are three major subwatersheds, including the central Grand, Conestogo, and Nith Rivers.

The central Grand River has a stream order of 5 before meeting the heavily agriculturally-influenced Conestogo River, and an order of 6 afterwards (Hicks, 2017; **Figure 4C)**. The watershed is influenced by intensive agriculture as well as urbanization (Waterloo/Kitchener). Flow in the central Grand River is controlled by two major upstream dams (Shand Dam on the Grand River and Conestogo Dam on the Conestogo River). The subwatershed is characterized by diamicton tills, moraines, and gravel and sand spillways (Lake Erie Source Protection Region Technical Team, 2008). Although spikes in flow can occur (e.g., precipitation and snow melt), upstream dams moderate the flow for flood protection as well as maintenance of summer flow.

The geology of the Nith River subwatershed changes from north to south, with predominately diamicton around Wellesley and New Hamburg, and sand, gravel and diamicton around Ayr (GRCA, 2018; **Figure 4C**). This basin is also subject to urban and intense agricultural influences, with the headwaters having poor water quality (Loomer & Cooke, 2011).

The Speed River subwatershed is characterized mainly by sand and diamicton throughout its course (GRCA, 2018; **Figure 4C)**. While the Speed River subwatershed is characterized by agricultural land use, it is less intensive and has relatively high water quality compared to the Nith and Grand Rivers (Loomer & Cooke, 2011). The city of Guelph and several dams (e.g. the Guelph Dam) exert a strong urban influence on this relatively small river, especially during low flows in the summer (GRCA, 2022; Loomer & Cooke, 2011).

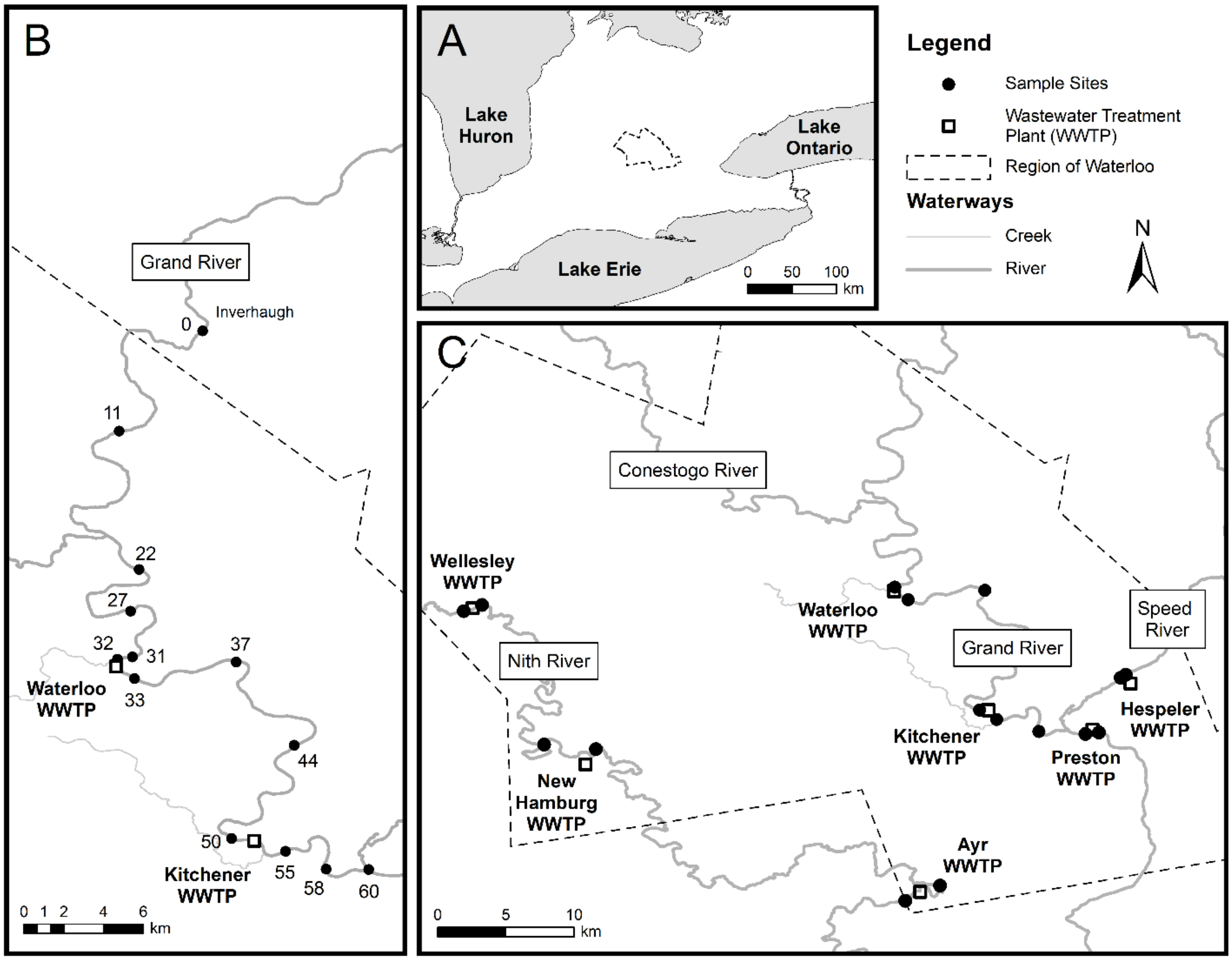

Figure 4. Map showing the study area (Region of Waterloo) in a Great Lakes context (A) and the locations of sampling sites to investigate longitudinal patterns in the central Grand River (B). Additionally, it includes the locations of sampling sites to determine if benthic macroinvertebrate assemblages upstream and downstream of the Kitchener and Waterloo WWTPs are more similar following upgrades, and if the other WWTPs in the Region of Waterloo are impacting their downstream assemblages (C). Contains information licensed under the Open Government License – Canada.

2.2 Experimental Design

This study has three components. The first is a longitudinal survey in the central Grand River to investigate BMI assemblages along the Grand River and how this pattern compares to previous years. The second component focusses on BMI assemblages before and after upgrades at the Kitchener and Waterloo WWTPs, investigating how they changed in association with upgrades. The third is looking into potential impacts of five other WWTPs in the Region of Waterloo, investigating if downstream assemblages differ from upstream as a result of effluent exposure.

2.2.1 Longitudinal survey study in the central Grand River

BMI samples were collected in autumn of 2020 at thirteen sites to assess the longitudinal pattern in benthic macroinvertebrate assemblages in the central Grand River (**Figure 4B**; **Table 3**). Sites were selected to span the central Grand River and capture major changes in river conditions, such as tributaries or WWTP outfalls, with samples collected from Inverhaugh and as far downstream as the confluence of the Grand and Speed Rivers. A subset of six of these sites were previously sampled in 2012 and 2014 by the research groups of Adam Yates (including E. Krynak), Western University, and Mark Servos (including K. Hicks), University of Waterloo, using the same methods as collections in 2020 (Servos et al., 2015), allowing for the assessment of how the longitudinal pattern in BMI assemblages has changed over time, with one sample collected per site per year. This subset consisted of site “11”, which is located upstream of the Grand River’s confluence with the Conestogo River and all WWTPs in this study. Additionally, the subset includes site “27”, located upstream of the Waterloo WWTP, and site “33”, located downstream. And lastly, site “50”, located upstream of the Kitchener WWTP, and sites “55” and “58”, located downstream.

Table 3. Sites used to investigate the longitudinal pattern of benthic macroinvertebrate assemblages. Site names are based on the longitudinal distance from the most upstream site, “0”, in kilometers. For example, site “44” is approximately 44 km downstream of site “0”.

Site Name	Latitude	Longitude	Sampling Years
“0”	43.63110	-80.44290	2020
“11”	43.58576	-80.48071	2012, 2014, 2020
“22”	43.52325	-80.47218	2020
“27”	43.50448	-80.47535	2012, 2014, 2020
“31”	43.48358	-80.47483	2020
“32”	43.48225	-80.48188	2020
“33”	43.47376	-80.47369	2012, 2014, 2020
“37”	43.48131	-80.42796	2020
“44” *	43.44428	-80.40095	2020
“50”	43.40978	-80.43173	2012, 2014, 2020
“55”	43.39576	-80.40575	2012, 2014, 2020
“58”	43.38810	-80.38736	2012, 2014, 2020
“60”	43.38753	-80.36768	2020

* Sampled a little over a month after all other sites (early December vs. late October).

2.2.2 BMI assemblages before and after upgrades at the Kitchener and Waterloo WWTPs

To assess if BMI assemblages exposed to Waterloo and Kitchener WWTP effluent differ before and after upgrade implementation, samples collected as a part of the Region of Waterloo Surface Water Quality Monitoring Program surrounding the WWTP outfalls were analyzed (**Table 4**). Site locations were chosen based on a Control-Impact design, sampling an upstream control site (approximately 500 m upstream of outfalls), a downstream-near site (approximately 1 km downstream of outfalls) and a downstream-far site (approximately 5 km downstream of outfalls). All sites were located in riffles of similar habitat conditions (e.g. substrate, depth, and flow rate). Sites were chosen based on suitability to the Surface Water Quality Monitoring Program, selecting the closest riffles to effluent release (except for the downstream-far site), while ensuring that these sites are consistently accessible even with variable flow (LGL Limited, 2010). Each site was sampled in the autumns of 2009, 2012, 2015, and 2018. There were plans to complete another sample collection in 2021, however this was delayed because of the COVID-19 pandemic.

2.2.3 Impacts in BMI assemblages downstream of other WWTPs in the Region of Waterloo

To investigate if the other five WWTPs in the Region of Waterloo appear to impact their downstream BMI assemblages, sample collections from the Region's Surface Water Quality Monitoring Program were analyzed (**Table 4**). Samples were collected in riffle habitats upstream and downstream of three WWTPs discharging to the Nith River west of the central Grand (Wellesley, New Hamburg, and Ayr), one WWTP discharging to the Grand River (Preston), and one discharging to the Speed River east of the central Grand (Hespeler). Sites were chosen based on suitability to the Surface Water Quality Monitoring Program, selecting the closest riffles to effluent release while ensuring that these sites are consistently accessible even with variable flow (LGL Limited, 2010). Five samples were collected at each site per year in 2009, 2012, 2015, and 2018. Data was not available for the Ayr and Wellesley WWTPs in 2009. There were plans to complete another sample collection in 2021, however this was delayed because of the COVID-19 pandemic.

Table 4. Sites included in assessing the effects of upgrades and if wastewater treatment plants in the Region of Waterloo impact downstream benthic macroinvertebrate assemblages. Samples were collected in 2009, 2012, 2015, and 2018 unless otherwise noted. US = Upstream, DSN = Downstream-near, DSF = Downstream-far, DS = Downstream.

River	Wastewater Treatment Plant	Site Code	Latitude	Longitude
Grand	Waterloo	US	43.48189	-80.4817
		DSN	43.47344	-80.47283
		DSF	43.48064	-80.42379
	Kitchener	US	43.40079	-80.42604
		DSN	43.39466	-80.4147
		DSF	43.38664	-80.38681
	Preston	US	43.38469	-80.35598
		DS	43.38622	-80.34704
Speed	Hespeler	US	43.42422	-80.3294
		DS	43.42217	-80.3327
Nith	Wellesley*	US	43.46571	-80.76611
		DS	43.47031	-80.75399
	New Hamburg	US	43.37802	-80.71297
		DS	43.37547	-80.67884
	Ayr*	US	43.28503	-80.45195
		DS	43.27486	-80.47475

* Data from 2009 is not available.

2.3 Water Chemistry

Instream water chemistry data was sourced from the Region of Waterloo (contracted to LGL Limited) as a part of the Surface Water Quality Monitoring Program (LGL Limited, 2019). When possible, three samples were collected at each bank and in the middle of the channel. Samples were later composited based on sample location and analyzed for nitrate (detection limit of 0.1 mg/L), ammonia (detection limit of 0.01 mg/L), and total phosphorous (detection limit of 0.002 mg/L) by Maxxam Analytical Laboratories. Samples were collected approximately twice a month from late July to mid November, with only some years having samples collected during winter months. Only data collected from July to November was assessed in this study due in part to the lack of winter data for some years, and also because BMI sample collections occurred in the Fall, meaning that the water chemistry during summer and early autumn is most relevant to assemblage composition. For details on the analytical methods regarding water chemistry, please refer to LGL Limited, 2010.

2.4 Longitudinal Survey Sample Collections

A 3-minute kick net sample was collected by McLay et al. at each riffle site on either October 29 or 30, 2020 using the Canadian Biomonitoring Network (CABIN) field protocol (ECCC, 2012a). The only exception is for site “44”, which was collected on December 3rd, 2020. The sampler travelled along a zigzag path using an A-frame net with a 500 µm mesh in a single riffle habitat, disturbing the substrate using kicks to suspend material and biota in the water column, allowing them to flow downstream into the net. Samples were preserved using 100% ethanol to account for dilution from river water. Leeches and crayfish were excluded from collections. Sampling was completed in late October as BMI are generally in greater abundance and more mature during this period in temperate streams (Hynes, 1970), making identification easier.

In 2012 and 2014, kick net sample collections were completed by the research groups of Adam Yates (including E. Krynak), Western University, and Mark Servos (including K. Hicks), University of Waterloo. The samples in 2012 and 2014 were also collected in a 3-minute kick using the CABIN protocol. However, the mesh size was 400 μm. Kick net sampling in 2020 included the same sampling sites as in 2012 and 2014 to ensure that BMI samples were directly comparable.

To prepare BMI samples for subsampling, all large materials such as rocks and leaves were removed to allow for effective homogenization. Once removals were complete, the sample was transferred to a 27 cm x 27 cm x 15 cm Marchant box to evenly disperse the sample among its 100 individual cells and allow for random sampling. Cells were then randomly selected and all the material within was then placed into a glass dish and observed using a boom microscope, repeatedly scanning from left to right and top to bottom until a scan of the entire sample was complete and no BMI were found. Cells were selected and all BMI within were picked until at least 300 organisms or 5% of the sample was processed, whichever came last. Any BMI found were then stored in 75% ethanol for preservation.

After subsampling was complete for a given sample, picked debris and material were collected and stored in bottles and QA/QC'd by Dr. Edward Krynak. During the QA/QC, Dr. Krynak looked through the debris (e.g. rocks, twigs) after a sample was picked to ensure all BMI were found. If any were found, they were counted and returned into the same vial as the BMI initially picked during subsampling. The number of BMI found during the QA/QC were divided by the total number of BMI picked, calculating a % efficiency to measure the effectiveness of subsampling. This process was completed on sites "11", "22", "27", "58", and "60", all of which had > 95% efficiency, which meets the ≥ 95% guideline provided by the Canadian Biomonitoring Network (ECCC, 2020). Given the consistent meeting of the 95% standard for the samples previously listed, QA/QC was not completed for the remaining samples.

The 13 enumerated subsamples were shipped to Cordillera Consulting in Summerland, British Columbia, Canada for taxonomic identification. Identification was completed to the genus/species level where possible. Taxa were excluded following guidelines found in the CABIN Lab Manual (ECCC, 2012a). Subsampling was already completed for the 2012 and 2014 kick net samples and had been previously identified by Cordillera Consulting to the species/genus level where possible.

2.5 Region of Waterloo Sample Collections

The following methods were employed to collect samples pertaining to the investigation of changes in BMI assemblages before and after upgrades at the Kitchener and Waterloo WWTPs, as well as detection of impacts downstream of the other five WWTPs in the Region of Waterloo.

Full details of the collection methods are presented in reports provided by LGL Limited (LGL Limited, 2010, 2013, 2016, 2019) under contract to the Region of Waterloo and summarized here.

Surber samplers were used to sample BMI, as collected by LGL Limited, the data from which was provided by the Region of Waterloo for use in this study. Collections followed the Environmental Effects Monitoring (EEM) program protocol (ECCC, 2012b) using a 30.5 x 32.0 cm Surber in 2009 and a 30.0 x 30.0 cm Surber in 2012, 2015, and 2018, all with a 500 µm mesh.

To collect a BMI sample, the Surber frame was pressed firmly into the substrate. All rocks and surface materials inside the frame were scrubbed clean and the top 2 cm of sediment disturbed so that the sample would flow downstream into the net. For all years, five composite samples consisting of three samples each were collected at each site (i.e. riffle) in and outside of the effluent plume.

All subsampling and processing of BMI samples was conducted by LGL Limited. In 2009, samples were put into a gridded tray with four sections for subsampling, after which each quarter was placed into a separate tray (LGL Limited, 2010). All BMI were systematically picked from the first quarter to collect a minimum of 250 individuals, with the entire sample being picked if 250 was not reached. In 2012, a known volume of the sample was extracted and moved to a tray to be picked and sorted of all BMI to collect at least 300 organisms, where the entire sample was picked if this threshold was not met (LGL Limited, 2013). In both 2009 and 2012, identification was completed by Dr. Richard Bland (Bland Richard Associates) to species-level where possible. In 2015 and 2018, samples were picked in the same manner as 2012 with the exception of samples being fixed in formalin and dyed with a protein dye to improve discernibility (LGL Limited, 2016, 2019). Additionally, for 2015 and 2018, taxonomic identification was completed to the lowest practical taxonomic level by ZEAS Incorporated.

2.6 Data Analysis

All data analyses were completed using R, version 3.6.3 (Holding the Windsock), using packages as outlined hereafter, or SPSS (Version 28.0.1). All data visualization for the results section, including ordinations and bar plots, were created using the R package ‘ggplot2’ (Wickham, 2016). All analyses were run with an $\alpha = 0.1$ to reduce the probability of Type II errors as the data utilized is inherently variable and the sample sizes are small.

Four metrics, taxa richness, Modified Family Biotic Index (MFBI) (Barbour et al., 1999; Bode et al., 1991; Hauer & Lamberti, 1996; Hilsenhoff, 1988), % scrapers, and % filterers, were used to characterize BMI assemblage response to the effects of WWTP effluent exposure (**Table 5**). These metrics were selected based on their potential to change following implementation of WWTP upgrades.

Table 5. Metrics used to characterize BMI assemblage response to WWTP effluent exposure. Included are details of what each metric means and how they're calculated, as well as the predicted change in that BMI assemblage metric downstream of the Kitchener and Waterloo WWTPs following upgrades.

Metric	**Description**	**Expected Change Downstream After Upgrades**
Family Richness	The # of unique families present in a given sample.	Increase in family richness
Modified Family Biotic Index (MFBI)	An index estimating water quality using the count of each taxon present and that taxon's tolerance to organic pollution.	Decrease in MFBI value
% Scrapers	A scraper is a feeding type that some BMI possess where individuals feed on microflora, microfauna, and algae attached to substrates. % scrapers is the relative abundance of this feeding type in a given sample.	Increase in % scrapers
% Filterers	A collector-filterer is a feeding type that some BMI possess where individuals feed on small particles/detritus suspended in the water column. % filterers is the relative abundance of this feeding type in a given sample.	Decrease in % filterers

2.6.1 Instream water chemistry (Kitchener and Waterloo WWTPs)

Annual instream ammonia collected at the US (upstream) and DSN (downstream-near) sites for the Kitchener and Waterloo WWTPs were analyzed using General Linear Models (GLMs, $\alpha = 0.1$). The median ammonia concentration was calculated for each year (Jun-Nov) for a total of 13 concentrations. Annual median concentrations were split into two time periods based on the timing of major change(s) in wastewater treatment processes. For Kitchener, the first time period (pre-upgrade) ranged from 2007-2012, while the other time period (post-upgrade) ranged from 2013-2019. For Waterloo, the first time period (pre-upgrade) ranged from 2007-2016 and the second time period (post-upgrade) ranged from 2017-2019. It should be noted that while the Waterloo WWTP underwent a period of construction increasing ammonia concentrations (2009-2015), this division of time periods should still indicate whether concentrations have improved given that construction was finished by 2016, and all upgrades were either implemented or soon-to-be implemented after 2017. The GLM was performed with effects of Site (upstream and downstream) and Time Period (pre-upgrade and post-upgrade), as well as their interaction (Site x Time Period) using SPSS (Version 28.0.1). When the GLM yielded a significant interaction term, a post-hoc test was completed using estimated marginal means comparing upstream and downstream sites within each time-period.

2.6.2 Taxonomic resolution harmonization

To ensure consistency in taxonomic level and that no taxonomic group is represented at different levels, all BMI assemblage data had their taxonomic resolution shifted based on the proportion of organisms identified to family versus genus. An insufficient number of organisms were identified to species to remain at the species-level, so the lowest level of taxonomy used was genus. To keep as much of the data to the genus level as possible without removing too many individuals who were only identified to family, certain guidelines were followed to perform taxonomic harmonization following similar methods to Krynak & Yates (2018) and Verdonschot (2006). The rule was such that if greater than 20% of individuals were identified only to family level (versus genus), all individuals' taxonomic level was adjusted to the family level.

2.6.3 Longitudinal survey study in the central Grand River

To assess the longitudinal pattern of BMI assemblages at the 13 sites in the central Grand River sampled in 2020 by kick net, a non-metric multidimensional scaling (NMDS) ordination was calculated using Bray-Curtis dissimilarity on Site x Taxa data via the R function *vegdist* in the package 'Vegan' (Oksanen et al., 2020). As a part of this function, a Hellinger transformation was applied to the Site x Taxa count data to prevent much of the similarity in the resultant ordinations being due to zero counts at both sites (Legendre & Gallagher, 2001). Rare taxa (i.e., any taxon appearing in < 5% of the samples) were excluded to reduce the number of zero values for taxa following Krynak & Yates (2018).

A separate NMDS ordination was calculated using the subset of six sites sampled by kick net in 2012, 2014, and 2020 in the same manner outlined previously.

2.6.4 BMI assemblages before and after upgrades at Kitchener and Waterloo WWTPs

To assess differences between BMI assemblages upstream and downstream of the Kitchener and Waterloo WWTPs, the five composite Surber samples collected at each of the US (upstream), DSN (downstream-near), and DSF (downstream-far) sites in 2009, 2012, 2015, and 2018 were used to calculate two NMDS ordinations using Bray-Curtis dissimilarity on Hellinger-transformed Site x Taxa count data via the R function *vegdist* in the package 'Vegan' (Oksanen et al., 2020). Taxa appearing in less than 5% of the samples were considered rare and were excluded from analyses (Krynak & Yates, 2018).

Assessment of each WWTP was done using separate two-way PERMANOVAs ($\alpha = 0.1$) on the Hellinger-transformed Site x Taxa data with the R studio function *adonis* in the package 'vegan' (Oksanen et al., 2020). Factors included were Site (Three levels: upstream, downstream near-field, and downstream far-field) and Year (Four levels: 2009, 2012, 2015, 2018) and their interaction (Site x Year). If a significant interaction term was yielded, a post-hoc pairwise PERMANOVA was conducted between upstream and downstream sites within each year using the R function *adonis* in the package 'vegan' (Oksanen et al., 2020). Average dissimilarity was calculated between US and DSN, and US and DSF sites within each year using the same dissimilarity matrix created for the two-way PERMANOVA, taking the mean dissimilarity of each pairwise comparison for each replicate upstream and downstream (5 replicates for each site, for a total of 25 comparisons).

A SIMPER analysis was completed on assemblages upstream and downstream (US vs DSN and US vs DSF) of the Kitchener and Waterloo WWTPs to see which taxa were most responsible for dissimilarity within each sampling year and how this changed over time. The SIMPERs were performed using the function *simper* in the package 'vegan' (Oksanen et al., 2020) on the Hellinger transformed Site x Taxa count data.

Separate GLMs ($\alpha = 0.1$) were run on the following univariate metrics: modified family biotic index (MFBI), % scrapers, % filterers, and family richness for each WWTP using SPSS (Version 28.0.1). To calculate % scrapers and % filterers, primary feeding type was assigned to taxa using the U.S. Freshwater Traits Database (EPA, 2012) using methods outlined in Krynak & Yates (2018). To best capture traits associated with the BMI of Southern Ontario, any information contained in the traits database sourced outside of Ecological Region 8.0 (EPA, 1997) and south of Tennessee and North Carolina, USA, was excluded. For taxa with multiple entries of varying modality for primary feeding type, the most abundant modality was applied. Effects for the GLM included Site and Year, as well as their interaction. When a significant interaction term was yielded from the GLM, a post-hoc test was completed using estimated marginal means comparing up and downstream sites among years.

2.6.5 Impacts in BMI assemblages downstream of other WWTPs in the Region of Waterloo

To assess patterns between BMI assemblages upstream and downstream of the Wellesley, New Hamburg, Ayr, Hespeler, and Preston WWTPs, separate NMDS ordinations were calculated using the five composite Surber samples collected each year at the US and DS sites in 2009, 2012, 2015, and 2018 using the R function *vegdist* in the package 'Vegan' (Oksanen et al., 2020). However, Ayr and Wellesley did not have sufficient data in 2009 and as such do not have this year in their ordinations. As a part of the calculation of these ordinations, the Site x Taxa count data was Hellinger transformed to reduce the degree of similarity due to absence (Legendre & Gallagher, 2001). Rare taxa were excluded from the analyses.

Assessment of the Wellesley, New Hamburg, Ayr, Hespeler, and Preston WWTPs was done using separate two-way PERMANOVAs ($\alpha = 0.1$) on the Hellinger-transformed Site x Taxa count data with the R studio function *adonis* in the package ‘vegan’ (Oksanen et al., 2020). Factors included were Site (Two levels: upstream, downstream), Year (Four levels: 2009, 2012, 2015, 2018) and their interaction (Site x Year). However, for the Ayr and Wellesley WWTPs, there were only three levels for year (2012, 2015, 2018). If a significant interaction term was yielded, a post-hoc pairwise PERMANOVA was conducted between the US and DS sites within each year using the R function *adonis* in the package ‘vegan’ (Oksanen et al., 2020). Average dissimilarity was calculated between US and DSN sites within each year using the same dissimilarity matrix created for the two-way PERMANOVA, taking the mean dissimilarity of each pairwise comparison for each replicate upstream and downstream (5 replicates for each site, for a total of 25 comparisons).

GLMs ($\alpha = 0.1$) were run for each of the univariate metrics, including modified family biotic index (MFBI), % scrapers, % filterers, and family richness for all five other WWTPs using SPSS (Version 28.0.1). To assign feeding type to each taxon and thereby calculate % filterers and % scrapers, methods outlined in Krynak & Yates, 2018 were followed utilizing the U.S. Freshwater Traits Database (EPA, 2012). To ensure that the data used pertained to BMI local to Southern Ontario, any trait data collected outside of Ecological Region 8.0 (EPA, 1997) and south of Tennessee and North Carolina was excluded. Some taxa had multiple sources with differing assignments of traits, in which case the most abundant modality was used. Effects for the GLM were site and year, as well as their interaction. In the case that the GLM produced a significant interaction term, post-hoc analysis using estimated marginal means was executed to compare upstream and downstream sites among years.

3.0 Results

3.1 Longitudinal Pattern in BMI Assemblages

NMDS ordination of BMI assemblages collected at 13 sites along the central Grand River in 2020 showed a lack of a longitudinal pattern, with only some assemblages grouped in association with the longitudinal sequence (**Figure 5A**). Assemblages at sites in the central reaches of the Grand River, such as “27”, “32”, and “33”, are grouped in the middle of the ordination, along with “37”, “44”, and “50” on the left. However, there are many assemblages distant from adjacent ones in the longitudinal sequence, such as those from the most upstream sites, “0” and “11”, with downstream assemblages “58” and “50” more similar to “11” than “0” is. Additionally, two of the more downstream assemblages, “60” and “55”, are placed far out of sequence, with “55” placed among more central sites (“27”, “31”, and “33”), and “60” isolated in the bottom-right corner of the ordination.

The ordination of six BMI assemblages collected along the central Grand River in 2012, 2014, and 2020 showed that BMI assemblages were more similar among sites within a year than within a site between years (**Figure 5B**). For example, in 2012, sites were predominately located in the left and middle regions of the ordination, whereas in 2014 sites shifted to the middle and right regions. Moreover, assemblages at the two upstream most sites (“11” and “27”) were more similar to one another in 2012 than assemblages at “11” were in 2012 and 2014. Finally, in 2020, all sites moved to the lower-right region of the ordination, further demonstrating the interannual variation occurring throughout the study period.

Additionally, the ordination of six BMI assemblages collected along the central Grand River in 2012, 2014, and 2020 suggested that the organization of sites within years varied over time (**Figure 5B**). Indeed, in 2012, sites generally followed a longitudinal gradient, except for “33”, which was grouped out of longitudinal sequence. In 2014, these groups shifted, with the most upstream sites (“11” and “27”) now separate from the three most downstream sites (“50”, “55”, and “58”), with a central site (“33”) far removed from all others. In 2020, these groupings largely dissolved, with all sites near each other and the three upstream sites (“11”, “27”, and

"33") in close proximity despite being the sites featuring the most dissimilarity among each other the previous sampling year.

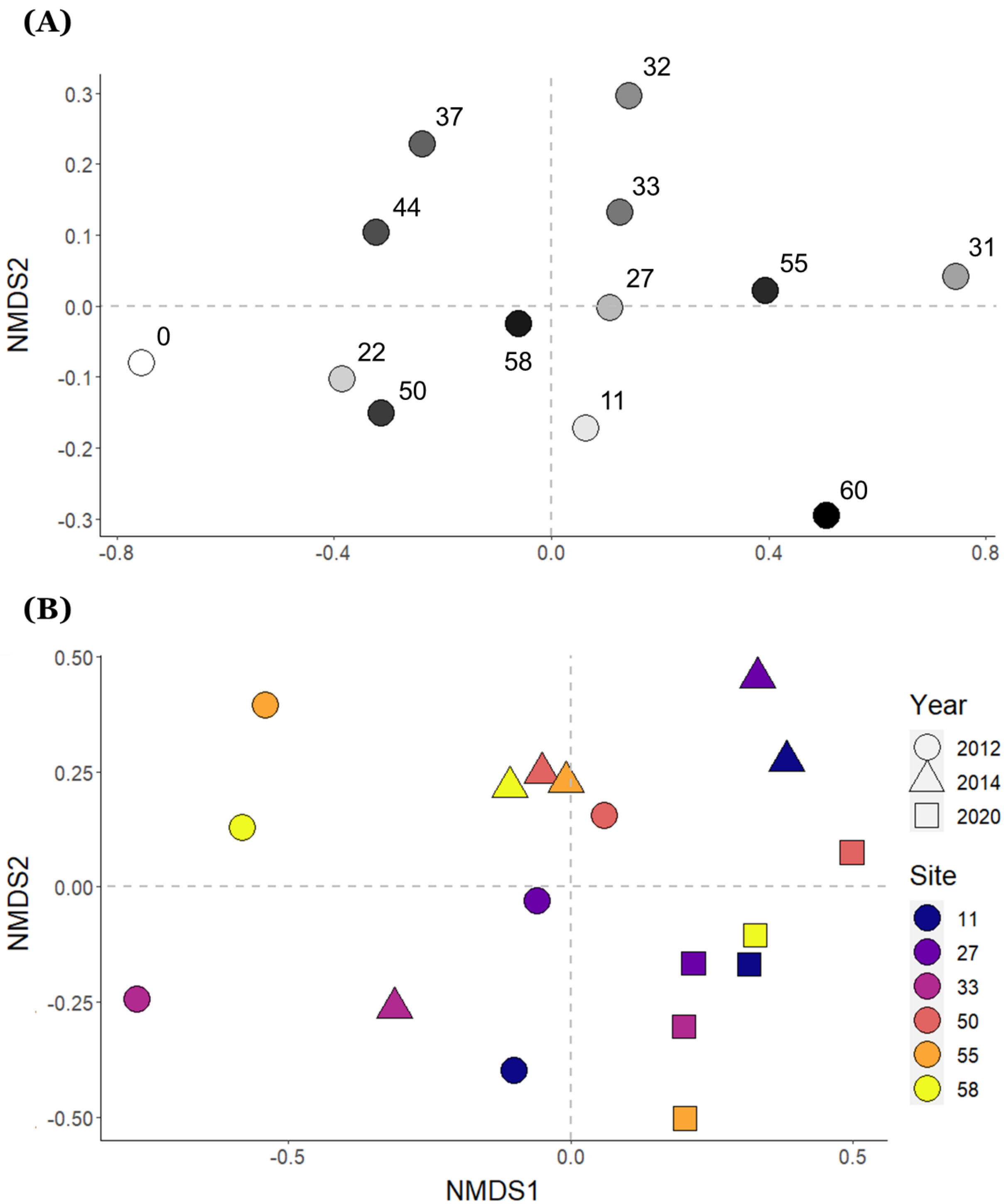

Figure 5. Ordination (NMDS) of benthic macroinvertebrate assemblages collected in the central Grand River in 2020 (A) and a subset of those sites in 2012, 2014, and 2020 (B). Site names denote longitudinal distance from the most upstream site of sampling (site "0").

3.2 Instream Water Chemistry

General Linear Models (GLMs) assessing instream ammonia and nitrate concentrations downstream of the Waterloo WWTP yielded significant interaction effects ($p = 0.003$ and 0.005, respectively; **Figure 6A, B**). Post-hoc, pairwise comparisons based on estimated marginal means indicated that ammonia concentrations were significantly greater at the downstream site during the pre-upgrade period ($p < 0.001$), but were not significantly different in the post-upgrade period ($p = 0.814$). For nitrate, post-hoc comparisons revealed that concentrations between up and downstream sites were significantly different during both the pre- and post-upgrade periods ($p = 0.029$ and $p < 0.001$, respectively).

The GLMs assessing instream ammonia and nitrate concentrations upstream and downstream of the Kitchener WWTP outfall revealed significant interaction terms ($p < 0.001$ and $p = 0.008$, respectively; **Figure 6C, D)**. Post-hoc, pairwise comparisons based on estimated marginal means indicated that ammonia concentrations were significantly greater downstream than upstream in the pre-upgrade period ($p < 0.001$), but not significantly different post-upgrade ($p = 0.477$). The opposite was the case for nitrate, with concentrations not significantly different between upstream and downstream sites pre-upgrade ($p = 0.690$), but significantly greater downstream during the post-upgrade period ($p < 0.001$).

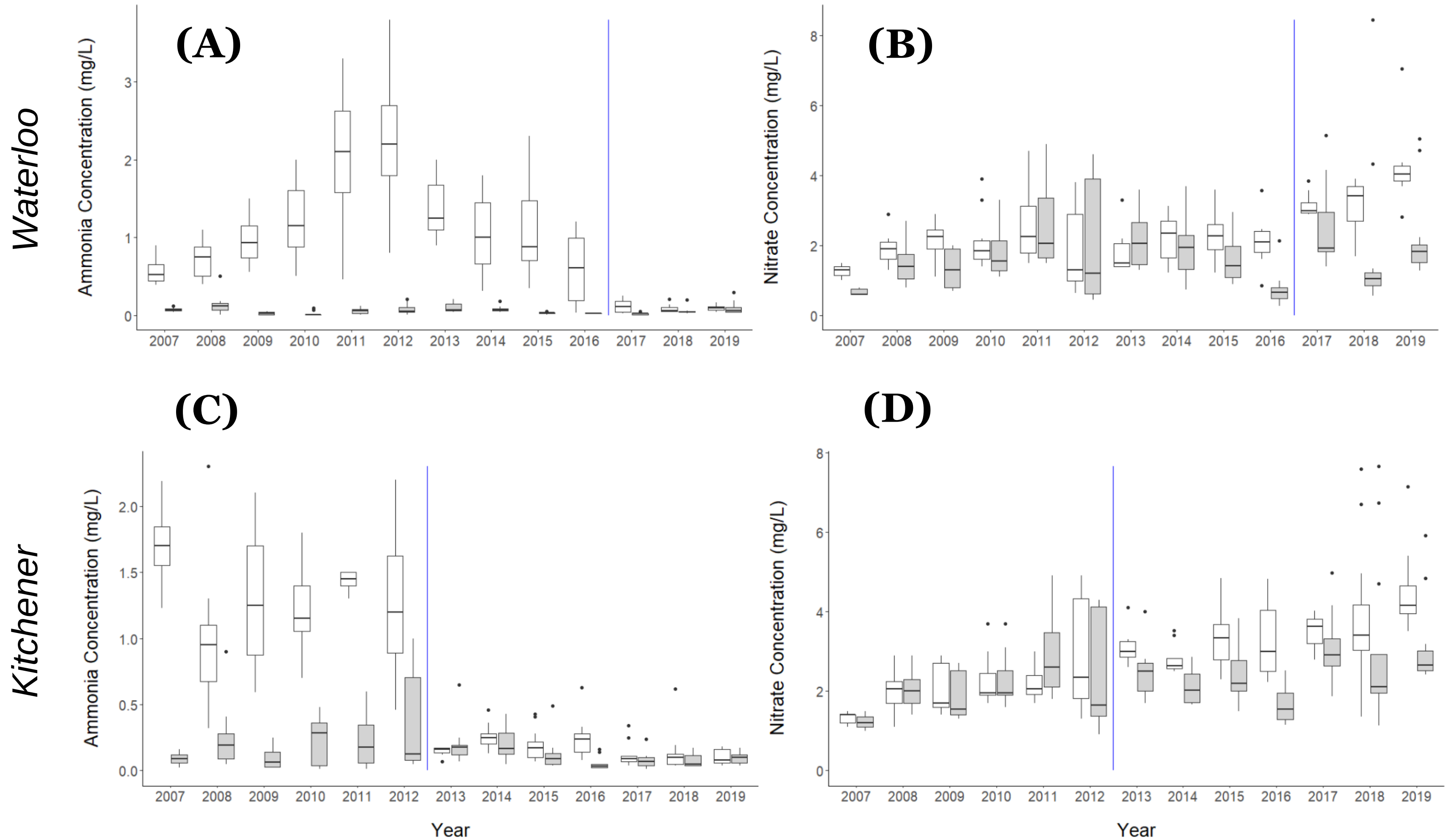

Figure 6. Boxplots showing the range of values of instream ammonia (A) and nitrate (B) concentrations at the Waterloo WWTP, as well as ammonia (C) and nitrate (D) concentrations at the Kitchener WWTP from Jun-Nov at US (grey) and DSN (white) sites collected as a part of the Region of Waterloo Surface Water Quality Monitoring Program. The years preceding the blue line (2007-2012 for Kitchener; 2007-2016 for Waterloo) compose the “pre-upgrade” time-period before major upgrades had been completed, and the years thereafter (2013-2019 for Kitchener; 2017-2019 for Waterloo) compose the “post-upgrade” time-period following the implementation of major upgrades. The black lines in boxplots denote the median concentration.

3.3 Assessment of WWTP Upgrade Effects

Waterloo WWTP

The NMDS ordination depicting BMI samples collected upstream and downstream of effluent release at the Waterloo WWTP showed increased dissimilarity between US & DSN and US & DSF sites from 2009 and 2012 in association with construction and decreased in dissimilarity following upgrades (**Figure 7A**). Additionally, assemblages collected within a given year were separate from assemblages collected in other years. The two-way PERMANOVA assessing site and year effects of effluent release on assemblage composition revealed a significant interaction between site and year ($p < 0.001$). Post-hoc, pairwise PERMANOVAs between US & DSN, and US & DSF sites were significant within all years (all p-values ≤ 0.032). Average dissimilarity between US and DSN sites trended in association with upgrades, increasing from 0.46 (± 0.11) in 2009 to 0.60 (± 0.14) in 2012, coincident with reduced treatment during construction, and then decreasing following the upgrades in 2015 to 0.52 (± 0.12) and 2018 to 0.35 (± 0.08). Average dissimilarity between US and DSF followed a similar trend, with an increase from 0.44 (± 0.06) to 0.48 (± 0.11) between 2009 and 2012 and then decreasing to 0.42 (± 0.08) in 2015 and 0.33 (± 0.04) by 2018.

SIMPER analysis revealed that the amount of dissimilarity between the Waterloo WWTP US and DSN sites associated with taxa that contributed less than 5% to dissimilarity decreased from 79% to 56% between 2009 and 2012, then increased to 74% between 2012 and 2018. Additionally, SIMPER analysis between US and DSF revealed that the dissimilarity explained by taxa that contributed less than 5% to this dissimilarity changed over time, decreasing from 94% to 61% between 2009 and 2012, then increasing to 95% between 2012 and 2018. There were also patterns in the taxa that contributed more than 5% of dissimilarity between sites. Similar patterns were observed for the isopod *Caecidotea spp.* at DSN and DSF, which had initially contributed more than 5% to between site dissimilarity only at DSN in 2009 (9%). Contributions exceeded 5% for both DSN and DSF in 2012 (15% and 5%, respectively) and 2015 (8 and 9%, respectively) due to high abundance at both sites. Thereafter, relative abundances of *Caecidotea spp.* at both DSN and DSF reduced to values similar to US in 2018, resulting in a dissimilarity contribution below 5% at both sites.

GLMs assessing % scrapers and family richness at the Waterloo WWTP indicated significant interaction effects ($p = 0.044$ and $p = 0.013$, respectively; **Figure 7**). Pairwise comparisons based on estimated marginal means revealed that % scrapers at US was significantly greater than at DSN in 2009, 2012 and 2015 ($p \leq 0.014$), but not in 2018 ($p = 0.188$). Likewise, % scrapers were significantly greater at US than at DSF for 2009 ($p = 0.020$) and 2012 ($p = 0.002$), but not for 2015 ($p = 0.321$) or 2018 ($p = 0.449$). For % scrapers between both US & DSN and US & DSF, pairwise differences in average values generally increased between 2009 and 2012, then decreased thereafter in association with upgrades. Pairwise comparisons based on estimated marginal means comparing family richness between US and DSN showed that richness was significantly greater US for all years ($p \leq 0.025$), whereas the pairwise comparisons for US and DSF sites indicated that values were significantly greater US for 2009 ($p < 0.001$) and 2012 ($p = 0.008$), but not for 2015 ($p = 0.446$) or 2018 ($p = 0.328$). For family richness between both US & DSN and US & DSF, differences in average values generally increased between 2009 and 2012, then decreased thereafter in association with upgrades. GLMs for the MFBI and % filterers did not yield a significant site by year interaction for the Waterloo WWTP ($p = 0.211$ and $p = 0.151$, respectively).

Kitchener WWTP

The NMDS ordination depicting BMI samples collected at the Kitchener WWTP showed BMI assemblages of the US, DSN, and DSF sites within years grouping progressively tighter over time, with a high degree of similarity among sites within years, but substantial separation between years regardless of site (**Figure 8A**). The two-way PERMANOVA assessing site and year effects of effluent release on assemblage composition indicated a significant interaction term ($p < 0.001$). Post-hoc, pairwise PERMANOVAs between US & DSN and US & DSF sites were all significant (all p-values ≤ 0.025). Average dissimilarity between US and DSN decreased from 0.49 (± 0.06) to 0.40 (± 0.06) between 2009 and 2012, then increased slightly to 0.41 (± 0.10) between 2012 and 2015 during which major upgrades occurred in 2012 and 2013, finally decreasing to 0.34 (± 0.06) in 2018. A similar pattern was observed for average dissimilarity between US and DSF, decreasing from 0.70 (± 0.08) to 0.33 (± 0.05) between 2009 and 2018, although there was a decrease between 2012 and 2015 of 0.50 (± 0.11) to 0.38 (± 0.05) coincident with upgrades.

Based on SIMPER analysis, the percent of overall dissimilarity between US and DSN assemblages associated with taxa that individually contributed less than 5% to dissimilarity decreased from 62% to 56% between 2009 and 2012, but increased thereafter in 2015 (59%) and 2018 (75%). A separate SIMPER analysis showed that the percent of total dissimilarity between US and DSF associated with taxa individually contributing less than 5% decreased from 65% to 47% between 2009 and 2012, then increased by 2015 (67%) before slightly decreasing by 2018 (61%). Between US and DSN, *Cheumatopsyche spp.* were a top contributor to dissimilarity in 2009 (12%) and 2012 (10%), featuring greater relative abundance at DSN compared to US, before lowering in contribution in 2015 (5%) and finally no longer being a top contributor in 2018 after relative abundance became more similar between sites. Between US & DSF, there was a pattern observed for the collector-filterer genus *Hydropsyche spp.* who was a top contributor to dissimilarity in 2009 (11%) and 2012 (7%), having greater relative abundance upstream, but ceased being a top contributor thereafter concurrent with upgrades and an increase in their relative abundance at DSF, though US relative abundance spiked in 2009. Finally, between US and DSF, Simuliidae was a top contributor in 2009 (11%), 2012 (7%), and 2015 (7%), but not in 2018 following upgrades, becoming less common at DSF and therefore closer to abundances measured at US, though US and DSF relative abundances both spiked in 2012.

The GLMs for MFBI and family richness yielded a significant interaction term ($p < 0.001$ and $p = 0.013$, respectively; **Figure 8**). Within year, pairwise comparisons based on estimated marginal means of MFBI values between US and DSN sites showed non-significant differences for 2009, 2012, and 2018 ($p \geq 0.220$), but revealed that MFBI value was significantly greater at DSN than US in 2015 ($p = 0.003$). On the other hand, pairwise comparisons between US and DSF indicated that MFBI value was significantly greater at DSF for 2009 ($p < 0.001$) but was not significantly different in 2012 ($p = 0.119$), 2015 ($p = 0.572$), or 2018 ($p = 0.112$). Regarding family richness, pairwise comparisons between US and DSN sites within years revealed that family richness was significantly greater US for 2009 ($p = 0.047$) and 2012 ($p = 0.030$), but not different for 2015 ($p = 0.161$) and 2018 ($p = 0.192$). Additionally, pairwise comparisons based on estimated marginal means for family richness between US and DSF indicated that family richness was significantly greater US for 2009 ($p < 0.001$), 2012 ($p < 0.001$), and 2015 ($p = 0.047$), but no significant difference was shown for 2018 ($p = 0.420$). GLMs for both the % scrapers and % filterers yielded non-significant site by year interactions ($p = 0.156$ and 0.110, respectively).

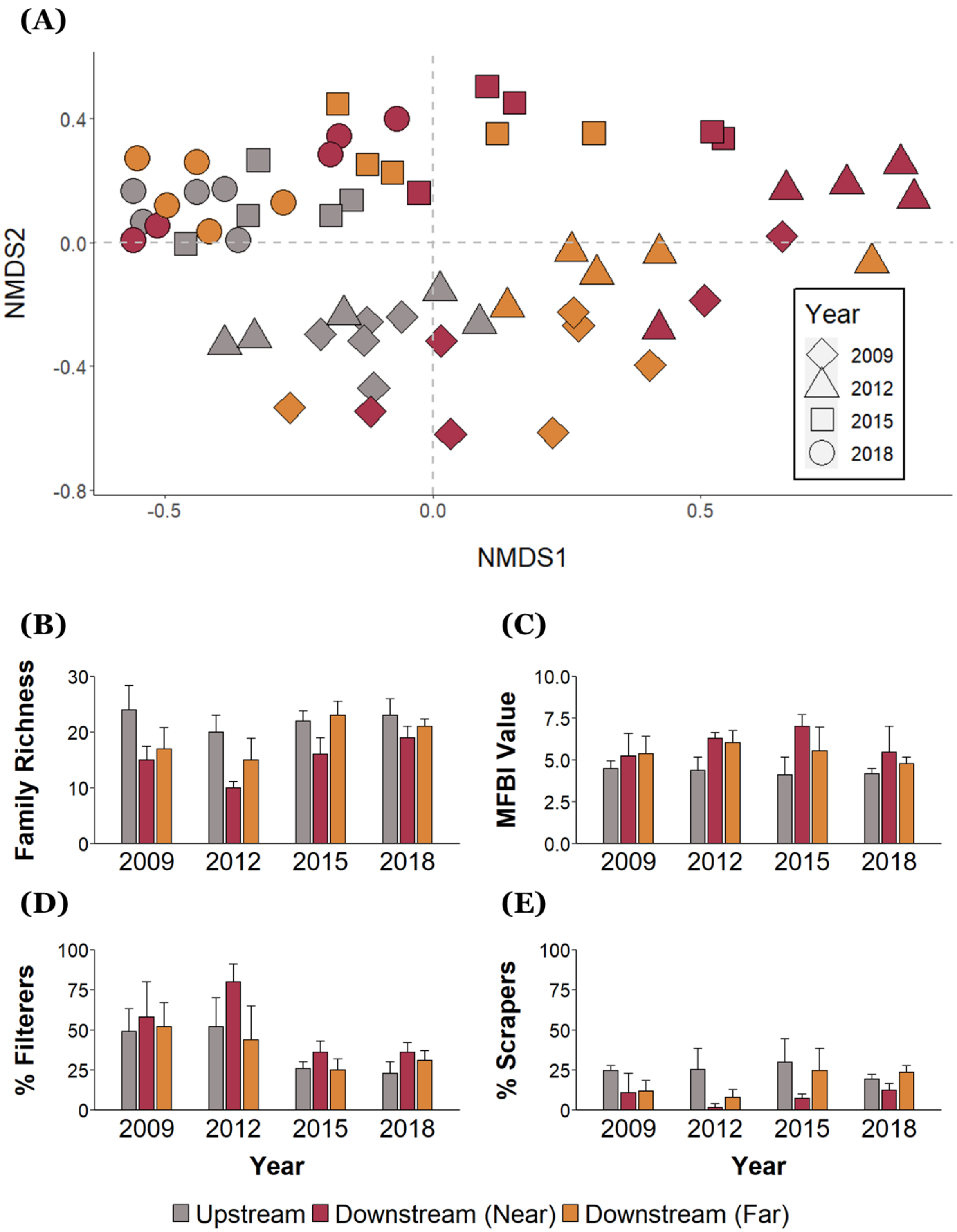

Figure 7. Ordination (NMDS) depicting dissimilarity of benthic macroinvertebrate assemblages (A) and bar plots describing average family richness (B), modified family biotic index (MFBI) value (C), % filterers (D), and % scrapers (E) in reaches located upstream, downstream (near) and downstream (far) of the Waterloo wastewater treatment plant outfall to the Grand River in the autumn of 2009, 2012, 2015, and 2018.

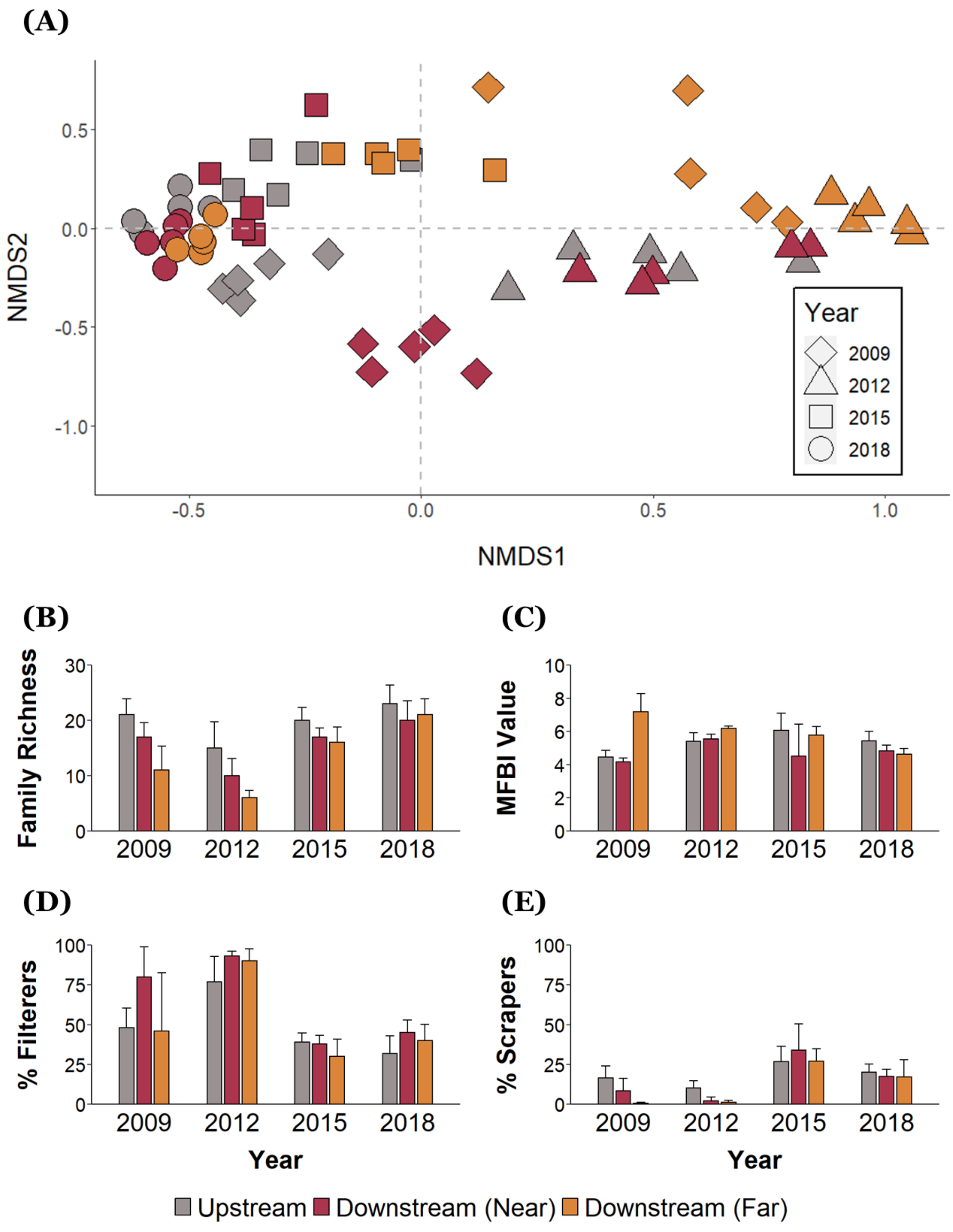

Figure 8. Ordination (NMDS) depicting dissimilarity of benthic macroinvertebrate assemblages (A) and bar plots describing average family richness (B), modified family biotic index (MFBI) value (C), % filterers (D), and % scrapers (E) in reaches located upstream, downstream (near) and downstream (far) of the Kitchener wastewater treatment plant outfall to the Grand River in the autumn of 2009, 2012, 2015, and 2018.

3.4 Assessment of Wastewater Impacts in the Region of Waterloo

Preston & Wellesley WWTPs

The NMDS ordination depicting BMI samples collected at the Preston (**Figure 9A)** and Wellesley (**Figure 10A**) WWTPs showed BMI assemblages of the upstream and downstream sites grouped together within years, but separate from sites among years. For both the Preston and Wellesley WWTPs, the two-way PERMANOVA assessing site and year effects of effluent release on assemblage composition yielded non-significant interactions ($p = 0.107$ and 0.510, respectively) and site terms ($p = 0.281$ and 0.158), but significant year terms (all p-values < 0.001). For the Preston WWTP, the GLMs for MFBI, % scrapers, family richness, and % filterers, had non-significant site terms (all p-values ≥ 0.188; **Figure 9**). The same lack of significant site term was revealed for the Wellesley WWTP for MFBI, % scrapers, and % filterers (all p-values ≥ 0.235), but not for family richness which was significantly greater upstream ($p = 0.003$; **Figure 10**).

Hespeler WWTP

The NMDS ordination for the Hespeler WWTP showed separation between upstream and downstream sites for all years, with a high degree of similarity among sites within years, but substantial separation between years regardless of site (**Figure 11A**). The two-way PERMANOVA assessing site and year effects of effluent release from the Hespeler WWTP on assemblage composition indicated a lack of a significant interaction term ($p = 0.306$), but significant site and year effects (both p-values < 0.001). It was found that there were no significant interactive effects on MFBI ($p = 0.419$), though MFBI values were observed to be significantly greater downstream ($p = 0.013$). Conversely, significant interactive effects were seen for % scrapers ($p = 0.048$), % filterers ($p = 0.002$), and family richness ($p = 0.023$; **Figure 11**). Pairwise comparisons based on estimated marginal means between upstream and downstream sites indicated that % scrapers were significantly greater downstream in 2009 ($p = 0.004$), but not different in all other years ($p \geq 0.418$). On the other hand, the same pairwise comparisons for % filterers showed no significant differences in 2009 and 2015 ($p \geq 0.210$), while % filterers were significantly greater upstream in 2012 ($p = 0.005$), and significantly greater downstream in 2018 ($p = 0.024$). Finally, pairwise comparisons for family richness featured an array of differing results, indicating that upstream family richness was significantly greater in 2009 ($p = 0.048$), not different in 2012 or 2015 ($p = 0.191$ and $p = 0.541$, respectively), and significantly greater downstream in 2018 ($p = 0.038$).

New Hamburg & Ayr WWTPs

The NMDS ordination depicting BMI samples collected at the New Hamburg WWTP showed BMI assemblages with a large degree of separation between sites within years for 2009 and 2015, but more tightly packed in 2012 and 2018 (**Figure 12A**). The NMDS ordination for the Ayr WWTP showed separation between sites within years (especially in 2012), but more pronounced separation among sites between years (**Figure 13A**). The two-way PERMANOVAs assessing site and year effects of effluent release from the New Hamburg and Ayr WWTPs on assemblage composition detected significant site ($p < 0.001$ and $p = 0.011$, respectively), year (both p-values < 0.001), and interaction effects ($p < 0.056$ and $p = 0.015$, respectively). Post-hoc, pairwise PERMANOVAs between upstream & downstream sites revealed significant site effects for all years at the New Hamburg (all p-values ≤ 0.010) and Ayr WWTPs (all p-values ≤ 0.049). The average dissimilarity between upstream and downstream sites at the New Hamburg WWTP did not follow a consistent trend, decreasing from 0.57 (± 0.07) to 0.35 (± 0.04) between 2009 & 2012, then increasing to 0.56 (± 0.05) by 2015, and finally decreasing once more to 0.29 (± 0.05) in 2018. In contrast, for the Ayr WWTP, the average dissimilarities remained somewhat consistent through time, ranging from 0.30 – 0.40 from 2012 to 2018.

A pattern similar to that seen in the average dissimilarity for the New Hamburg WWTP was also observed in its % filterers and MFBI, whose GLMs revealed significant interaction terms (both p-values < 0.001; **Figure 12**). Pairwise comparisons based on estimated marginal means between upstream and downstream sites within years for % filterers showed that % filterers was significantly greater downstream in 2009 and 2015 ($p < 0.001$ and $p < 0.001$, respectively), significantly greater upstream in 2012 ($p = 0.011$), but not different in 2018 ($p = 0.388$). Likewise, pairwise comparisons for MFBI values between upstream and downstream sites within years revealed significantly greater values upstream in 2009 and 2015 (both p-values < 0.001), significantly greater values downstream in 2012 ($p = 0.007$), and no difference in 2018

($p = 0.102$). No significant interactive or site effect was observed for % scrapers ($p = 0.628$ and $p = 0.769$, respectively) or family richness ($p = 0.876$ and $p = 0.250$, respectively) for sites upstream and downstream of the New Hamburg WWTP.

GLMs for the Ayr WWTP yielded a significant site term for MFBI, % filterers, % scrapers, and family richness (all p-values ≤ 0.043), but only % scrapers featured significant interactive effects ($p = 0.009$; **Figure 13**). Pairwise comparisons based on estimated marginal means between upstream and downstream sites within years indicated that % scrapers were significantly greater at downstream sites compared to upstream 2012 and 2018 ($p < 0.001$ and $p = 0.070$, respectively), while no significant differences were found in 2015 ($p = 0.334$).

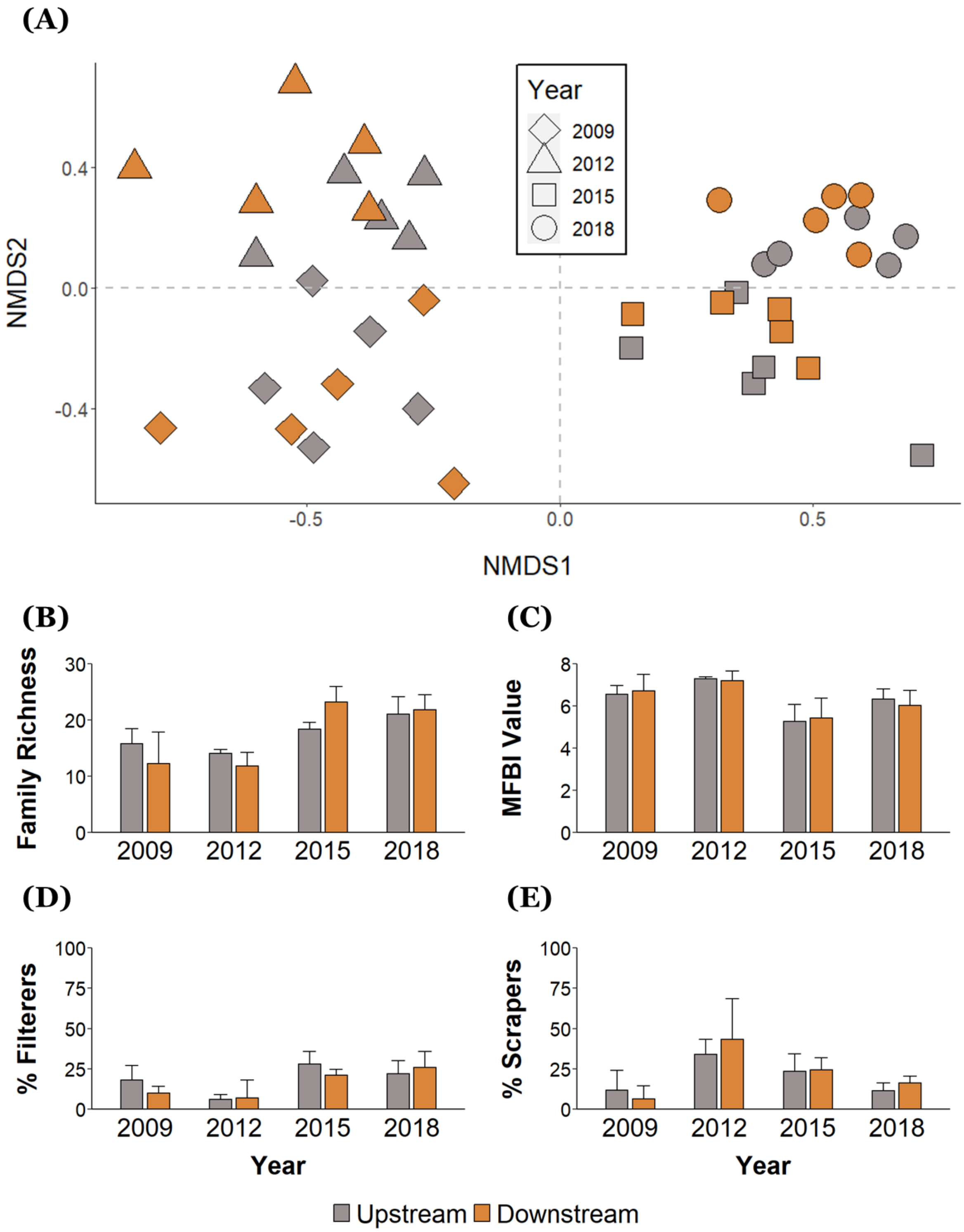

Figure 9. Ordination (NMDS) depicting dissimilarity of benthic macroinvertebrate assemblages (A) and bar plots describing average family richness (B), modified family biotic index (MFBI) value (C), % filterers (D), and % scrapers (E) in reaches located upstream and downstream of the Preston wastewater treatment plant outfall to the Grand River in the autumn of 2009, 2012, 2015, and 2018.

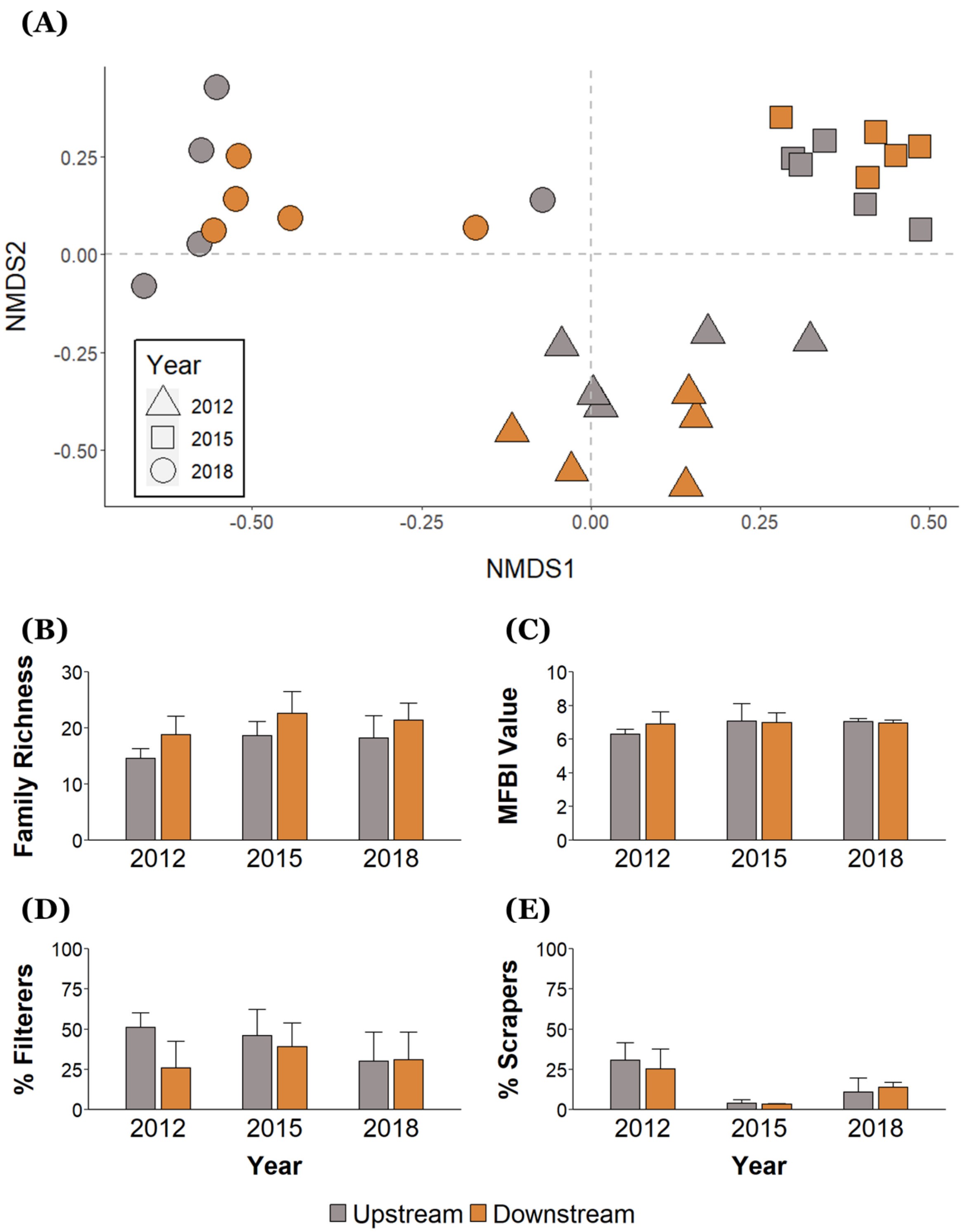

Figure 10. Ordination (NMDS) depicting dissimilarity of benthic macroinvertebrate assemblages (A) and bar plots describing average family richness (B), modified family biotic index (MFBI) value (C), % filterers (D), and % scrapers (E) in reaches located upstream and downstream of the Wellesley wastewater treatment plant outfall to the Nith River in the autumn of 2012, 2015, and 2018.

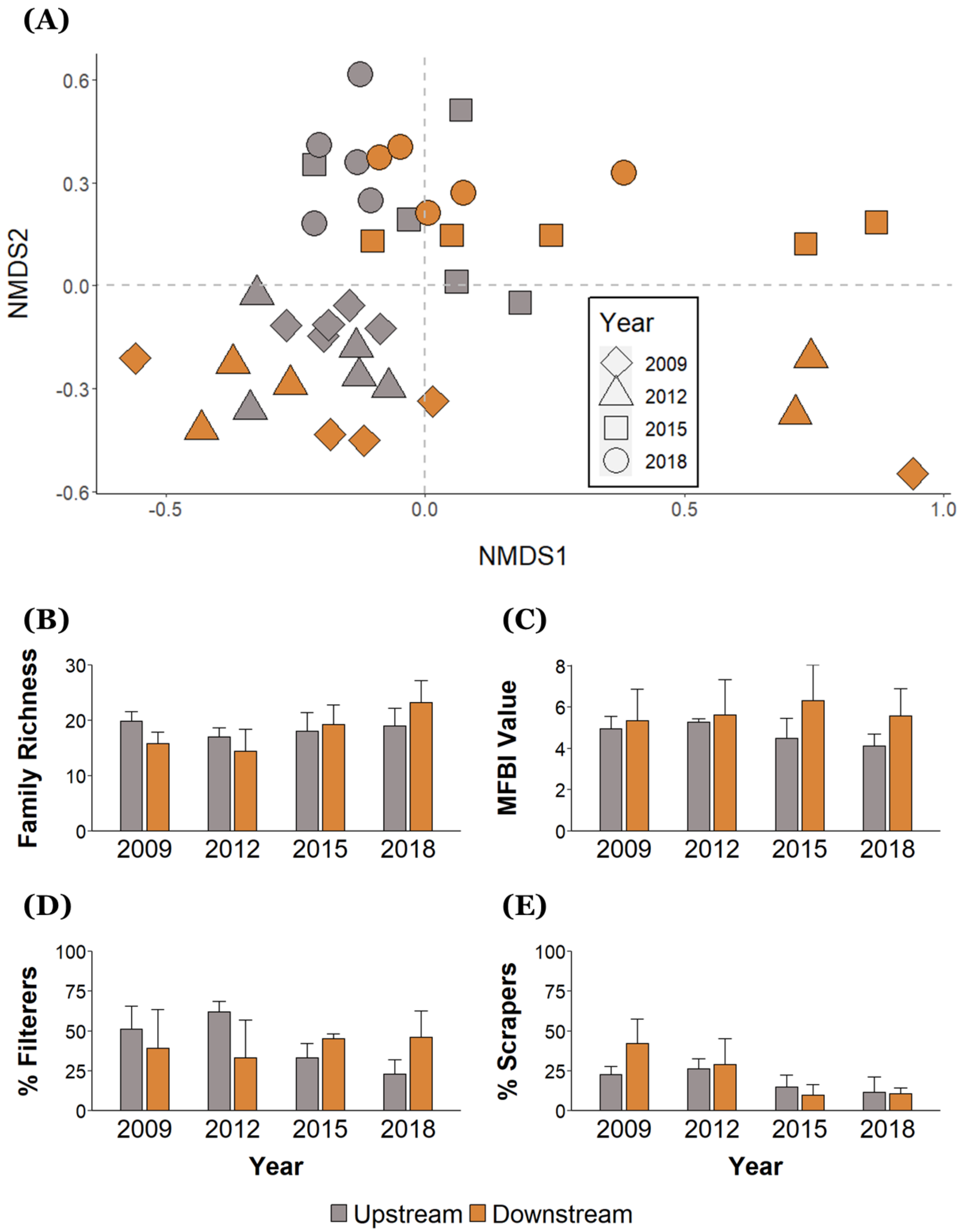

Figure 11. Ordination (NMDS) depicting dissimilarity of benthic macroinvertebrate assemblages (A) and bar plots describing average family richness (B), modified family biotic index (MFBI) value (C), % filterers (D), and % scrapers (E) in reaches located upstream and downstream of the Hespeler wastewater treatment plant outfall to the Speed River in the autumn of 2009, 2012, 2015, and 2018.

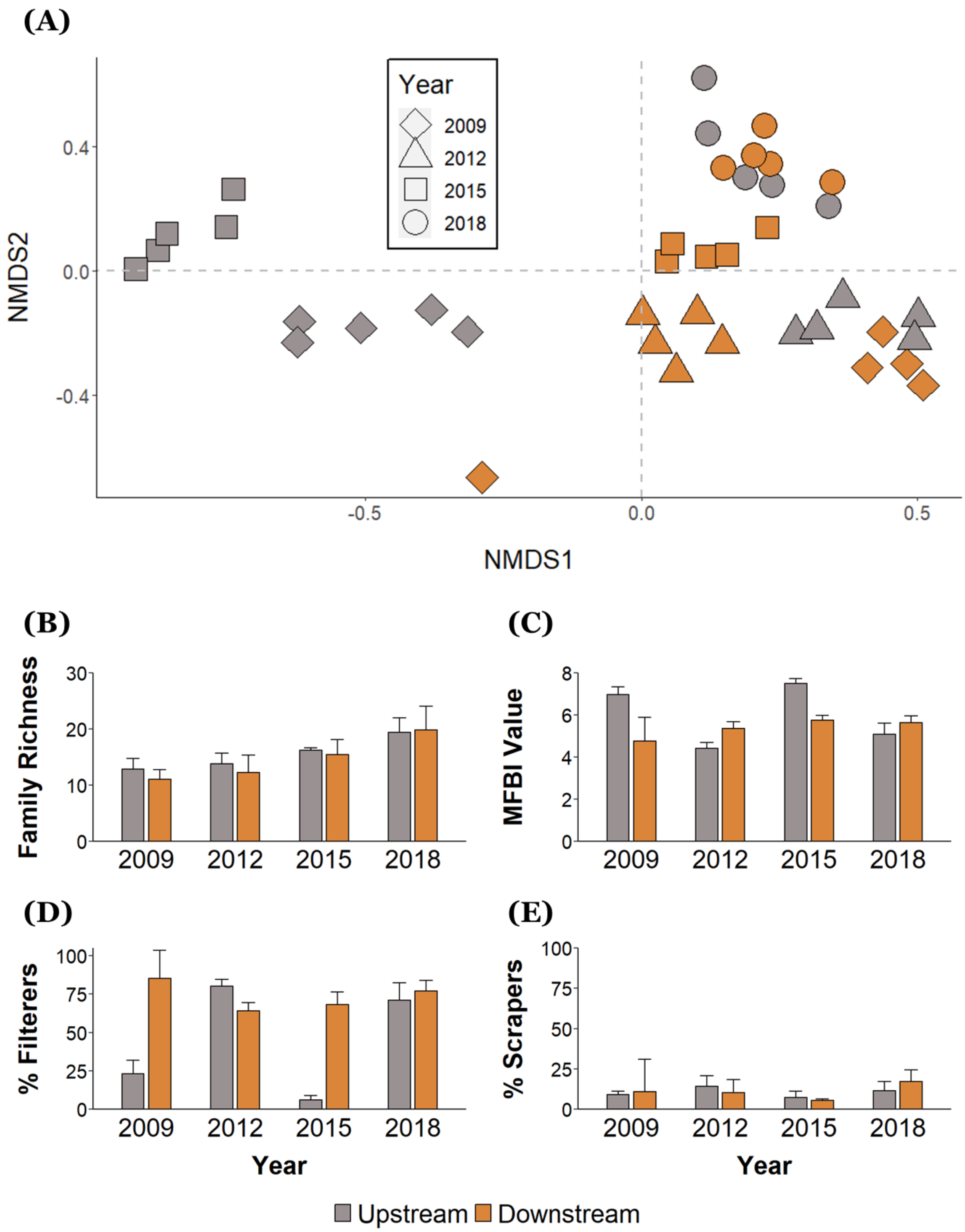

Figure 12. Ordination (NMDS) depicting dissimilarity of benthic macroinvertebrate assemblages (A) and bar plots describing average family richness (B), modified family biotic index (MFBI) value (C), % filterers (D), and % scrapers (E) in reaches located upstream and downstream of the New Hamburg wastewater treatment plant outfall to the Nith River in the autumn of 2009, 2012, 2015, and 2018.

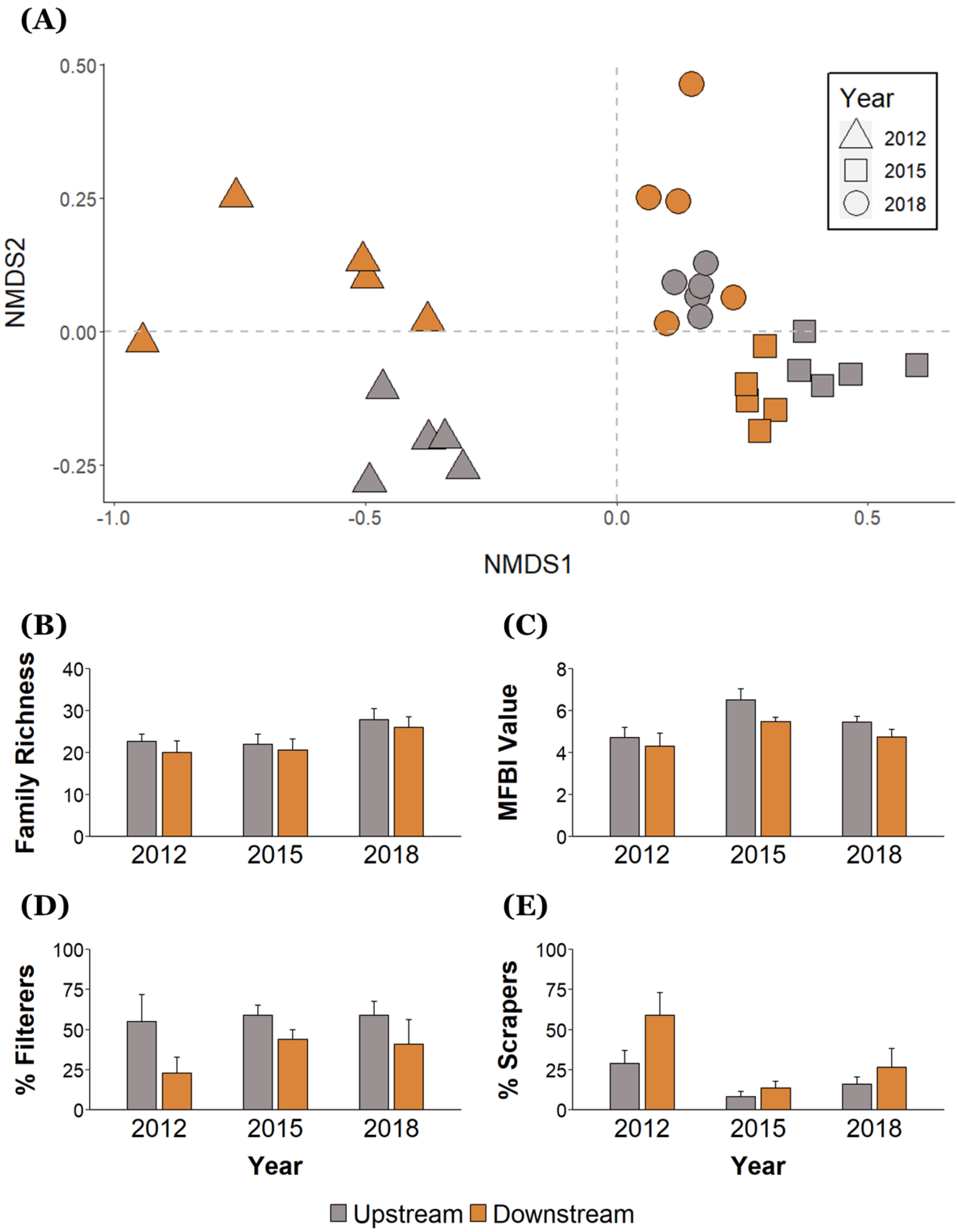

Figure 13. Ordination (NMDS) depicting dissimilarity of benthic macroinvertebrate assemblages (A) and bar plots describing average family richness (B), modified family biotic index (MFBI) value (C), % filterers (D), and % scrapers (E) in reaches located upstream and downstream of the Ayr wastewater treatment plant outfall to the Nith River in the autumn of 2012, 2015, and 2018.

4.0 Discussion

Over the past decade, upgrades have been implemented to the wastewater treatment processes of two of the largest WWTPs in the Region of Waterloo, with major upgrades occurring at the Kitchener WWTP in 2012/2013 and at the Waterloo WWTP in 2017/2018. These upgrades led to sizeable reductions in effluent nutrient concentrations (ammonia/nitrate) and intersex incidence in male rainbow darters (Hicks et al., 2017a; Nikel, 2020). However, previous to this study, a long-term, community-level investigation into biotic effects had yet to be completed. The current assessment used data provided by the Region of Waterloo, collected by LGL Limited as part of the Surface Water Quality Monitoring Program, to investigate if improvements in water quality resulting from WWTP upgrades led to the rehabilitation of downstream BMI assemblages. Based on these data, upstream and downstream BMI assemblages became more similar following the implementation of upgrades at the Kitchener and Waterloo WWTPs, indicating that upgrades were effective in rehabilitation. Another facet of the study was to investigate if other WWTPs in the Region of Waterloo impacted BMI assemblages downstream of their effluent outfall. The findings show that BMI assemblages downstream of other WWTPs were sometimes dissimilar to their upstream counterparts, with the Ayr, New Hamburg, and Hespeler WWTPs appearing to affect BMI downstream. However, further study would be necessary to determine if additional remedial actions are appropriate and would have the potential to shift downstream BMI assemblages to become more similar to upstream assemblages. Additionally, the results highlight the need to maintain the robust BACI study design currently employed to better isolate effluent effects given the clear influence of interannual variability independent of WWTP effluent exposure evident in all analyses of BMI assemblages. Furthermore, analyses of the entire assemblage appear to be more consistently able to reveal patterns in the BMI assemblage compared to metrics that only use a subset of the assemblage, like % filterers and % scrapers, though these metrics were valuable. Going forward assessments should account for the apparent effect that sampling in or out of the effluent plume has, given that variation arising from sample location may be masking the extent of effects or benefits resulting from upgrades.

4.1 Ecological Recovery in BMI Associated with WWTP Upgrades

Benthic macroinvertebrates responded to major changes in the wastewater treatment process, whether that be process upgrades or treatment interruptions. Indeed, assemblages became less similar between upstream and downstream sites at the Waterloo WWTP following increases in ammonia concentrations resulting from construction. However, with the completion of construction and implementation of upgrades that improved nitrification and therefore reduced ammonia, BMI assemblages below the outfall became more similar to the upstream control sites. The same pattern was observed at the Kitchener WWTP, with assemblages becoming more similar following upgrades that introduced full nitrification and reduced ammonia loading (i.e., 2012). These findings are similar to previous studies on fish that have shown a reduction in key adverse effects in fish populations following WWTP upgrades at the Kitchener (Hicks et al., 2017a; Marjan et al., 2018), and Waterloo (Nikel, 2020) WWTPs, indicating that upgrades have been effective in remediating ecosystem conditions in the central Grand River.

Increased similarity in BMI assemblage composition has also been observed in past studies assessing benefits of wastewater treatment upgrades of wastewater effluent release occurring in other river systems. For example, in the River Ray, values of biotic indices downstream of the Swindon WWTP generally became more similar to values observed upstream following reductions of instream ammonia concentrations (Johnson et al., 2019). Similar findings were also observed in the White River, Indiana, U.S., where upgrades significantly reduced ammonia and BOD, and significantly increased nitrate in-stream, leading to the ratio of pollution-tolerant to pollution-sensitive taxa downstream of the WWTP outfall increasing to levels seen upstream (Crawford et al., 1992). Likewise, following cessation of effluent discharge to small streams in the Australian Blue Mountains, upstream and downstream assemblages showed increasing similarity (Besley & Chessman, 2008). The current study in the Grand River thus adds to this growing body of literature indicating that WWTP upgrades are an effective approach to improving benthic community health.

While similar responses to upgrades were seen downstream of both the Waterloo and Kitchener WWTPs, impacts were more pronounced at the near-field (DSN) site for Waterloo, but the far-field site (DSF) for Kitchener. At the Waterloo WWTP, the pattern was more pronounced at the DSN site, likely because DSN is only about 1 km downstream of the WWTP outfall, whereas DSF is about 5 km downstream. Given that the effluent tends to remain on the west bank after release (Nikel, 2020), the effluent contaminants likely had more opportunity to degrade, disperse, and dilute across the channel prior to reaching DSF, leading to the observed lesser effects. Less pronounced impacts at sites further downstream from WWTP outfalls in the River Ray, England were also linked to greater dilution in a study by Johnson et al. (2019). In contrast, differences with the Kitchener US site were more pronounced at the DSF site. Greater dissimilarity in the assemblages at the Kitchener DSF site may be due to the longstanding dissolved oxygen depletion that occurred at the DSF (AKA Blair) sampling site (Loomer & Cooke, 2011). This phenomenon for oxygen downstream of the Kitchener WWTP was likely a result of cumulative depletion of dissolved oxygen, perhaps from in-stream oxidation of ammonia (Eklöv et al., 1998) or resulting from excessive, effluent-induced aquatic plant and macroalgae growth (Loomer & Cooke, 2011). Prior to the upgrades, a daily low in dissolved oxygen occurred at the Kitchener DSF site in the summer months, generally occurring between 5:00 and 11:00 am with hourly medians of 4.37 – 5.14 mg/L (Loomer & Cooke, 2011). Following the implementation of full nitrification at the Kitchener wastewater (i.e. completion of Kitchener plant 2) in early 2013, dissolved oxygen increased and daily lows were less pronounced, with hourly medians of 7.67 – 7.88 mg/L from 5:00 and 11:00 am (**Figure A1**). This may have been a contributing factor to the recovery seen in BMI assemblages at the Kitchener DSF site (Johnson et al., 2019). Given that the observed effects of effluent discharge were not always most impactful close to the WWTP outfall, this study's findings demonstrate the importance of monitoring at multiple downstream sites.

Although the assemblages upstream and downstream of the two upgraded WWTPs have gotten more similar over time, some differences persist. While it is possible that the current suite of WWTP upgrades is not sufficient to shift BMI assemblages further, assemblages may not have become indistinguishable due to the inherent natural variation between any two habitats. Additionally, the latest upgrades may have been completed too close to the last sampling period (2018) for effects to have fully manifested.

Between any two riffles, there is likely to be some degree of natural variation that will result in differing habitat conditions for the BMI assemblage therein. Despite a priori efforts to reduce variability by sampling similar riffle habitats, there will always be some dissimilarity between sample sites. Moreover, deterministic factors (e.g. water chemistry, flow, and substrate) are not the sole drivers of BMI assemblage composition, with stochastic factors such as dispersal and disturbance also playing roles in streams impacted by anthropogenic activities (Larsen & Ormerod, 2014). With this natural variation present, upstream and downstream assemblages may always differ regardless of circumstance and very well may explain the consistent dissimilarity between upstream and downstream assemblages.

Some of the upgrades at the Kitchener and Waterloo WWTPs may have been implemented too close to the 2018 sampling for their effects to be observed in the Region of Waterloo collections. For example, tertiary filtering was added to the Kitchener WWTP to remove phosphorus in late 2017, and the latest upgrade to the Waterloo WWTP occurred only about eight months prior to the 2018 sample collection. Recovery of BMI assemblages is not always immediate, with the return to an unimpacted state following a disturbance often requiring an extended period of time (Besley & Chessman, 2008; Clements et al., 2021; Crawford et al., 1992; Johnson et al., 2019). A lag in recovery for BMI assemblages can be influenced by: 1) availability and distance from a source of colonizers (Downes et al., 2002); 2) the ability of taxa to disperse (Arce et al., 2014); 3) the generation time of taxa present (Rader et al., 2008); 4) and the presence of a resistant assemblage of effluent-tolerant taxa preventing the establishment of sensitive taxa (Frost et al., 2006).

Availability and distance from a source of colonizers can determine if colonization is possible following WWTP upgrades (Arce et al., 2014). Indeed, delays in BMI assemblage recovery have been attributed to lack of access to colonizers in a number of studies (Chadwick et al., 1986; Clements et al., 2021; Lydy et al., 2000; Smith et al., 2011). Moreover, barriers, such as dams, have been posited to limit colonizer access to reaches downstream of a WWTP outfall following implementation of upgrades (Crawford et al., 1992). However, the connectivity between upstream and downstream sites at the Kitchener and Waterloo WWTPs is likely substantial, with a maximum distance of about 6 km between sites and no major physical obstructions. Consequently, availability and distance to colonizers are unlikely to prevent the downstream BMI assemblage to respond to upgrades, however more time may be required for the colonization to occur, especially if an accumulation of organic matter (e.g. macrophytes) resulting from effluent release persists after upgrades.

The dispersal ability of BMI present can also limit whether organisms colonize. Taxa that exhibit high dispersal ability often become more common or dominant in the assemblage during initial recovery (Foster et al., 2020; Kotalik et al., 2021; Oliver et al., 2012), with taxa having low dispersal ability lagging behind. Dispersal ability has been posited to be an issue for the recovery of freshwater (Unionid) mussels downstream of the Kitchener WWTP given their sedentary lifestyle reliant on host fish for migration (Gillis et al., 2017b). While Unionids were not well-reflected in the current study, many of the taxa that were rare or missing downstream in 2018 but abundant upstream were those with a high dispersal ability, indicating that this trait likely did not contribute to the lag seen in BMI assemblage recovery. However, for the Waterloo WWTP, most of the taxa most responsible for dissimilarity in 2018 had low dispersal ability (Ephemerellidae, Simuliidae), thereby reducing their capacity to migrate downstream and indicating that this may have played a role in delayed recovery. While there does appear to be some limitation in the dispersal ability of taxa present upstream but absent or rare downstream at the Kitchener WWTP, this likely isn't causing the lag seen in BMI assemblage recovery at either WWTP given the high connectivity between upstream and downstream sites.

The life cycle of the BMI present at a given site is an important factor to consider as BMI often have generation times of several months to a year (Hilsenhoff, 1977). These long generation times mean it may take an extended period for the consequences of upgrades to manifest in the BMI assemblage. It was observed that most taxa responsible for dissimilarity between upstream and downstream sites in 2018 were univoltine and therefore produce a single generation per year. Although the major upgrades for the Kitchener WWTP occurred in 2012/2013, the Waterloo WWTP achieved full nitrification in March of 2018, only several months prior to the 2018 collection. Several months is likely too short for univoltine taxa to have had the opportunity to produce a new generation exposed to post-upgrade conditions and thereby exhibit effects of upgrades. Considering the differences in upgrade timeline between the two WWTPs, life cycle likely does not fully explain delays in BMI assemblage recovery seen at the Kitchener WWTP but may play a role for assemblages downstream of the Waterloo WWTP.

The presence of a resistant assemblage of effluent-tolerant taxa can prevent the colonization of sensitive species in aquatic communities (Frost et al., 2006). For example, muted recovery of BMI assemblages following amelioration of acid-impacted streams has been attributed to a biotic resistance of tolerant taxa from being displaced by sensitive taxa (Layer et al., 2013; Ledger & Hildrew, 2005). With anthropogenic disturbances such as wastewater effluent having the ability to shift BMI assemblages to alternate states (Folke et al., 2004), another disturbance (e.g. flood) may be required to displace the tolerant taxa who thrived in impaired conditions before more sensitive species can colonize following WWTP upgrades in the Grand River. Considering the many factors that can delay BMI response, another BMI sample collection may be crucial in revealing any further benefits resulting from both recent and earlier upgrades.

4.2 Differences Detected Between BMI Assemblages Upstream and Downstream of Effluent Outfall at Other WWTPs in the Region of Waterloo

Beyond the Kitchener and Waterloo WWTPs, BMI assemblages were also sampled at several additional plants in the Region of Waterloo that received little to no upgrades during the study period. Not only do each of these plants differ considerably in infrastructure (process, size, receiving environment, etc.), there are also many other anthropogenic stressors like agriculture and urban runoff whose influence vary throughout Region of Waterloo, along with differences in river morphology. The impact these factors have on BMI assemblages will likely differ for each WWTP, and may potentially confound the effects of effluent exposure.

Of the five other WWTPs included in this Region of Waterloo study, only assemblages and metrics at the Wellesley and Preston WWTPs indicated that upstream and downstream sites were similar, and therefore provided no evidence of effluent effects. The lack of difference observed between assemblages at these two WWTPs may be due to minimal impact of the effluent or may be a result of other anthropogenic stressors present in the reach obscuring the potential impact of the effluent. The Kitchener and Waterloo WWTPs are located upstream of the Preston WWTP, both of which discharge much more effluent (average effluent discharge of 9,239 m^3/day in 2018 for Preston relative to Waterloo's 41,805 m^3/day and Kitchener's 67,902

m^3/day), which may exert too strong an influence on the assemblage for the additional effluent from Preston to have an observable effect. Furthermore, there are upstream agricultural and urban inputs from the Grand and Speed Rivers that may affect the BMI assemblage at Preston WWTP. As for the Wellesley WWTP, there are inputs from further upstream in the highly agricultural Nith River watershed that may also affect BMI assemblages. These other stressors may influence both upstream and downstream assemblages, confounding or masking the potential effects of effluent exposure. However, there was a detectable effect on BMI assemblages associated with effluent release at the New Hamburg, Ayr, and Hespeler WWTPs. While there do appear to be effects on BMI assemblages downstream of some WWTP outfalls, further studies at these sites would be required to determine if upgrades would result in benefits, especially in the context of cumulative effects of multiple stressors.

4.3 Effectiveness of Study Design

This study has shown the effectiveness of using BMI assemblages as an indicator of effluent impacts in the Region of Waterloo with assemblages appearing to respond to WWTP upgrades. The effectiveness is likely due to the many characteristics BMI possess that facilitate their use. Benthic macroinvertebrates seemed to respond quicker to WWTP upgrades than Unionid mussels (Gillis et al., 2017b), changing within two years of upgrade implementation whereas mussels did not within a year, with the authors positing that it may require decades for the Unionids to respond. This highlights a strength of using BMI, with the relatively shorter life cycles of most BMI compared to Unionids potentially allowing for shifts to occur sooner, although the more complex life cycle of Unionid mussels may have also delayed response (Gillis et al., 2017b). Furthermore, BMI are of a moderate trophic level, meaning they are an important resource for higher trophic levels (Webster et al., 2022) such as fish, while consuming lower trophic levels such as algae. As a result, that they can integrate impacts from both bottom-up and top-down forces (Wallace & Webster, 1996). Although BMI are able to integrate impacts from a variety of sources and provide insight on general water quality, it can be difficult to isolate the effects of any particular stressor on the assemblage (Hall et al., 2016). In addition, BMI may also not respond directly to specific contaminants, such as estrogens, that can cause reproductive effects in fish or other vertebrates (Hicks et al., 2017a; Marjan et al., 2018; Tetreault et al., 2011). While BMI are an effective indicator of effluent impacts, it is important to recognize the capabilities and limitations of any biota studied in addressing specific questions.

The need to account for interannual variation was evident at all WWTPs studied in the Region of Waterloo, with assemblages at the same site featuring marked dissimilarity between years. This dissimilarity isn't surprising as even streams with no known impairment have some degree of natural interannual variability within BMI assemblages (Milner et al., 2006). Interannual variation can arise from a variety of sources such as droughts, snowfalls, and floods (Bradt et al., 1999). Indeed, some of the sampling years included in the current study were exceptionally dry or wet and may have influenced BMI assemblage composition. For example, in conjunction with the exceptionally dry conditions in the summer and fall of 2012, there was a strong divergence in the 2012 Kitchener WWTP assemblage with both upstream and downstream assemblages markedly different than all other years. Furthermore, there were stark dissimilarities in the assemblages among years for almost all WWTPs studied, especially in the New Hamburg, Preston, and Wellesley BMI assemblages. Interannual variation was also observed in the BMI assemblages of the central Grand River, with pronounced dissimilarity between assemblages of different years, as well as the degree of similarity between sites within years changing over time. Considering that there was marked dissimilarity even between assemblages sampled at the same upstream site among years (e.g., Wellesley and New Hamburg WWTPs), this interannual variation is likely attributable to factors other than effluent exposure. Therefore, it is necessary to account for this interannual variation when investigating effluent effects.

It is also imperative to account for the influence of other anthropogenic stressors in the system when investigating the effectiveness of WWTP upgrades. This need is evident in the 2020 longitudinal survey of BMI in the central Grand River, with the effect of effluent difficult to discern given the lack of pattern based on network position. A lack of pattern is not unexpected as there is a bottom-draw dam (Shand Dam) upstream of the central Grand River, a confluence with the agricultural- and wastewater-impacted Conestogo River, as well as inputs from the Kitchener and Waterloo wastewater treatment plants and urban areas. Agriculture (Delong & Brusven, 1998), municipal wastewater (Manfrin et al., 2013), and dams (Ellis &

Jones, 2013) have been shown to disrupt natural longitudinal gradients in BMI assemblages, potentially confounding and masking the signal from effluent effects. As a result of these other anthropogenic stressors, as well as natural variation, assemblages at a given site may very well shift from year to year for reasons other than WWTP upgrades, masking the influence of upgrades. However, BMI assemblages upstream and downstream of the effluent outfall will theoretically experience these other stressors and natural variation similarly, with comparisons between these assemblages accounting for their influence and better isolating for the effects of upgrades over time. Thus, usage of the BACI design is critical as it enables an isolated comparison of the differences between upstream and downstream sites, accounting for confounding factors.

4.4 Effectiveness of Data Descriptors

In the Region of Waterloo data set, utilizing the entire BMI assemblage in analysis generally revealed clearer patterns than analysis of specific metrics, although some metrics appeared to be more effective than others (patterns summarized in **Table 6** and **Table 7**). In this study, multivariate analysis detected effluent effects at five of the seven WWTPs, while univariate metrics were inconsistent in their detection, often providing conflicting results regarding effluent effects for the same WWTP. Past studies have also found multivariate descriptions of the BMI assemblage to be more sensitive to anthropogenic stressors than univariates (Cao et al., 1996) or have higher strength in detection (Milner & Oswood, 2000; Yates & Bailey, 2011). This may be due to multivariate approaches capturing effects in the entire assemblage, whereas univariates focus on a single aspect of the assemblage that may not be subject to change.

Some metrics showed more consistency than others regarding the detection of effluent effects. Patterns in family richness aligned with multivariate analysis findings for five out of the seven WWTPs, even changing in relation to the upgrade timeline for both the Kitchener and Waterloo WWTPs. On the other hand, % scrapers and MFBI were more inconsistent, only aligning with patterns in multivariate analysis for four out of seven WWTPs, and only changing in association with the upgrade timeline for one upgraded WWTP each. Least sensitive of all was

% filterers, which only corroborated multivariate analysis findings at the Preston and Wellesley WWTPs.

However, it should be noted that all metrics had some results that were contrary to ecological expectations. For example, an increase in % scrapers would be expected following upgrades (Shieh et al., 1999), but this did not occur at the Kitchener WWTP following improved treatment. There are many potential causes for this observation, such as many scraper taxa also being sensitive to pollution (Barbour et al., 1996), or that eutrophication favored filamentous algae (Cattaneo, 1987), which are generally resistant to predation by scrapers (Wellnitz & Ward, 1998), although it is difficult to say which mechanism was truly responsible. Additionally, reductions in scrapers can be coupled with a dominance in collector-filterers (Shieh et al., 1999), which wasn't observed for either WWTP during the initial pre-upgrade collection (except for Kitchener's DSF). This may be due to other factors in the system exerting more influence on collector-filterers than effluent discharge (Arce et al., 2014). Furthermore, MFBI values were higher upstream than at the downstream site at both the New Hamburg and Ayr WWTPs, and family richness was greater downstream than upstream, indicating that water quality or habitat was worse upstream than downstream. Among univariates, family richness appears to consistently capture changes in wastewater treatment in BMI. However, % scrapers and the MFBI are somewhat inconsistent, while % collector-filterers seem to not be sensitive in this system. Although univariate analyses in many instances were inconsistent in the patterns they revealed, they do have their use supplementing multivariate approaches, which better integrated the information and allowed for greater interpretation of trends in BMI response across the Region of Waterloo.

4.5 Considerations for Future BMI Biomonitoring of WWTP Impacts

Although the sites chosen for sampling have proven effective thus far, the locations of replicate sampling in relation to the effluent plume appear to affect similarity among BMI assemblages (coordinates and maps of replicate locations can be found in LGL Limited, 2010, 2013, 2016, 2019). For example, at the Waterloo DSN site, there was notable variability among the replicate samples collected within a given year. After further investigation, the variability appeared to be due to the bank of collection, with marked dissimilarity between assemblages collected on the west versus east bank. This dissimilarity was likely a result of the effluent being discharged with minimal mixing, making the distribution of the plume uneven and concentrated on the west bank of the Grand River (Nikel, 2020). Consequently, assemblages sampled on the west bank would be under direct influence of the effluent, while those sampled on the east bank would be virtually unaffected. A similar situation was observed for the Kitchener DSN site, with effluent prior to 2016 being released in the centre of the river, but running along the east bank, (LGL Limited, 2016, 2019) before a diffuser system was commissioned releasing effluent more evenly in the middle of the channel thereafter. Additionally, sampling in and out of the effluent plume may have contributed to the less pronounced effects seen at the Kitchener DSN site relative to DSF, as some BMI assemblages sampled may not have been exposed to effluent which therefore muted responses. This was also observed downstream of the Hespeler WWTP, with samples taken on either side of a fork in the river exhibiting stark differences, potentially as a result of the effluent plume flowing primarily down the east channel (LGL Limited, 2019). While changing the sampling location to the section of river where effluent is released may disrupt the study design as the entire riffle would still be sampled upstream, it would better align with the goals of the Surface Water Quality Monitoring Program in characterizing WWTP impacts and evaluating the efficacy of WWTP upgrades. Should sampling continue across the entire habitat regardless of effluent exposure, it should be noted that the signal of effluent impact appears to be affected by replicates collected outside of the effluent plume.

Changes in site location between years can complicate analysis, potentially affecting confidence in conclusions yielded. For Hespeler, the impact of effluent release was still detectable despite a change in the US sampling site location, moving approximately 150 m downstream between 2012 and 2015. While site effects were detected, there was also a significant interaction term for three out of four univariate analyses on metrics. Upon further inspection of the data, a distinct change occurred between 2012 and 2015 following the changing of the upstream site location, where one site featured higher relative values for metrics compared to the other site previous to location movement, with this relationship switching thereafter. This may have caused the detection of an interaction term which would hypothetically confound any evidence for WWTP upgrade effects. It appears that while movements in site location of this magnitude may not preclude the detection of effluent effects, it may preclude comparisons to samples collected prior to the site change if differences between the two assemblages are large enough.

Table 6. Summary of results for the univariate (metrics) and multivariate (assemblage composition) analyses completed assessing the site, year, and interactive effects at the Kitchener and Waterloo wastewater treatment plants (WWTPs). A blue check mark denotes that this analysis or metric indicated that the benthic macroinvertebrate assemblage trended in association with the implementation of upgrades, while red dashes indicate that they did not trend in association. Note: MFBI = Modified Family Biotic Index.

WWTP	Assemblage Composition	MFBI	% Scrapers	% Filterers	Family Richness	# Consistent with Upgrades
Waterloo	✓	—	✓	—	✓	3/5
Kitchener	✓	✓	—	—	✓	3/5

Table 7. Summary of results for the univariate (metrics) and multivariate (assemblage composition) analyses completed assessing the site, year, and interactive effects of five wastewater treatment plants (WWTPs) in the Region of Waterloo. A blue check mark denotes that this analysis or metric indicated the presence of effluent effects (i.e. assemblages upstream and downstream of effluent outfalls differed throughout the sampling period). A red dash denotes assemblages did not indicate the presence of effluent effects. Note: MFBI = Modified Family Biotic Index.

WWTP	Assemblage Composition	MFBI	% Scrapers	% Filterers	Family Richness	# of Trends Indicating Effects
Wellesley	—	—	—	—	—	0/5
New Hamburg	✓	—	—	—	—	1/5
Ayr	✓	—	✓	—	✓	3/5
Preston	—	—	—	—	—	0/5
Hespeler	✓	✓	—	—	—	2/5

5.0 Recommendations for Future Studies

This study has demonstrated the effectiveness of WWTP upgrades, with BMI assemblages downstream of the Kitchener and Waterloo WWTP outfalls becoming more similar to upstream assemblages. However, it is highly recommended that data from future BMI collections as a part of the Surface Water Quality Monitoring Program are brought into these analyses to capture any further recovery resulting from upgrades. Not all impacts can be seen in the BMI assemblage immediately, and with some upgrades implemented after the final collection in 2018 and more yet to come, the upcoming 2022 collection could reveal not only further impacts from previous upgrades, but also the impacts of upgrades implemented since. Collections beyond 2022 will also be helpful in further elucidating the impacts of any subsequent upgrades, with BMI also acting as a continual bioindicator to monitor effluent effects regardless of modifications to the wastewater treatment process.

An important part of elucidating the effect of upgrades on BMI was the application of the BACI design. Benthic macroinvertebrate data collected before, during, and after upgrades were crucial in assessing the effectiveness of these upgrades as the design isolated the effects of treatment processes from those of natural variation and other human activities in the watershed apparent in the longitudinal survey and assessments of WWTPs. It is imperative that regular collections retain the BACI design to allow for future analysis to continue to make robust assessments of the quality of receiving waters.

Looking holistically at the patterns that analyses based on the entire assemblage versus metrics (MFBI, family richness, etc.) revealed, the entire assemblage appears to yield more confident conclusions on effluent effects and WWTP upgrades. Patterns revealed by % filterers were sometimes contrary to those revealed by other metrics and as such, likely won’t have much use in the future. However, family richness, MFBI, and % scrapers may have their use as supplementary analyses as the patterns they revealed often aligned with analyses on the entire assemblage, as well as diversify methods to confirm the effectiveness of upgrades.

A potential limitation of this study is the location of the effluent plume relative to the replicate sampling locations. At the Waterloo, Kitchener, and Hespeler WWTPs, the effluent stays in a particular section of the channel based on where the outfall discharges and remains unmixed by the time it reaches the DSN site. This means that BMI samples taken outside of the effluent plume are either minimally or completely unaffected by the effluent, and therefore not reflective of effluent impact. This may have muted the response from BMI assemblages in some analyses, making conclusions more difficult to reach. There are three possible ways to address this. Firstly, continue with random-sampling within the riffle downstream, but acknowledge the effect that replicates collected outside of the effluent plume may have on analysis. Secondly, continue randomly sampling within the riffle downstream, but add enough replicates so that there are five taken directly within the plume. This will require additional replicates to be collected, but allow for the replicates taken to be split into two "sub-sites" with different advantages. While not an immediate solution, the five replicates sampled as before (in and out of the plume) can be compared to the replicates collected directly in the plume, providing the opportunity to compare the two and calibrate the differences between. Eventually, once the difference between the two sub-sites have been confidently characterized, it will allow for future analyses to better address questions of effluent effects and results of upgrades. Lastly, continue with sampling as before, but add an entirely new site near the downstream-near site whose riffle is completely within the effluent plume to better capture effluent effects and results of upgrades.

It should be noted that all recommendations are regarding benthic macroinvertebrates in the Region of Waterloo study. While BMI have been shown to be a useful bioindicator capable of capturing the effects of WWTP upgrades, there are other aspects of the ecosystem within and beyond the Region of Waterloo to account for. For instance, while further reductions in nutrient emissions from WWTPs may not elicit change in BMI assemblages, they will still reduce nutrient loading to Lake Erie. Therefore, all recommendations pertain to the BMI monitoring program and should be weighed against other priorities and aspects of the ecosystem when being considered for implementation.

Summary of Recommendations

1. Continue with regular sample collections to monitor BMI assemblages.
2. Incorporate the next sample collection into these analyses to track the impact of previous and subsequent upgrades on water quality.

3. Maintain overall study design, sampling at both the upstream site as a reference and downstream sites as impacts sites.
4. Further study should be conducted at the Hespeler, New Hamburg, and Ayr WWTPs to confirm the potential of upgrades to address apparent effluent impacts on downstream BMI assemblages.
5. Analysis should focus more on the entire assemblage using multivariates as they appear to be more effective at detecting differences compared to metrics such as % filterers, although family richness, % scrapers, and MFBI appear to have some supplementary use.
6. Consider options to address the variation seen in BMI samples collected in and outside of effluent plumes downstream of the Kitchener, Waterloo, and Hespeler WWTPs.

6.0 Conclusions

Municipal wastewater is a well-established threat to the aquatic environment that will only be exacerbated as the human population continues to grow. With this inevitable population growth is the need for more effective wastewater treatment to ensure that effects are minimized as efficiently as possible. The current study used BMI to assess the effectiveness of upgrades implemented to the Kitchener and Waterloo WWTPs and found that assemblages upstream and downstream of effluent outfall generally became more similar with upgrades in treatment. Additionally, BMI assemblages at the Hespeler, New Hamburg, and Ayr WWTPs in the Region of Waterloo featured dissimilarity between their upstream and downstream assemblages, indicating that these WWTPs are affecting the biotic community downstream of effluent outfall. However, further study should be completed to confirm whether there is potential for upgrades to improve conditions for BMI downstream of these three WWTPs. Furthermore, all investigations of BMI in both the central Grand River and adjacent to WWTPs exhibited marked interannual variation, demonstrating the need for a robust study design (i.e. BACI) to better isolate for effluent effects and by extension, WWTP upgrades. These findings will help inform wastewater managers regarding upgrade implementation, as well as provide insights into designing and maintaining an effective biomonitoring program. It is recommended that monitoring is continued, the data be re-analyzed once new BMI samples are collected to detect delayed effects, analyses focus on the entire BMI assemblage, and that either future sampling or interpretations of results are modified to account for the potential effects of sampling BMI in and out of effluent plumes.

References

Allan, J. D. (2004). Landscapes and riverscapes: The influence of land use on stream ecosystems. *Annual Review of Ecology, Evolution, and Systematics*, *35*(2002), 257–284. https://doi.org/10.1146/annurev.ecolsys.35.120202.110122

Allan, J. D., Castillo, M. M., & Capps, K. A. (2021). Stream ecology. In *Springer Nature Switzerland*. https://doi.org/10.1201/9781420031676.ch8

Arce, E., Archaimbault, V., Mondy, C. P., & Usseglio-Polatera, P. (2014). Recovery dynamics in invertebrate communities following water-quality improvement: Taxonomy-vs trait-based assessment. *Freshwater Science*, *33*(4), 1060–1073. https://doi.org/10.1086/678673

Ashauer, R. (2016). Post-ozonation in a municipal wastewater treatment plant improves water quality in the receiving stream. *Environmental Sciences Europe*, *28*(1), 1–7. https://doi.org/10.1186/s12302-015-0068-z

Bahamonde, P. A., Tetreault, G. R., McMaster, M. E., Servos, M. R., Martyniuk, C. J., & Munkittrick, K. R. (2014). Molecular signatures in rainbow darter (*Etheostoma caeruleum*) inhabiting an urbanized river reach receiving wastewater effluents. *Aquatic Toxicology*, *148*, 211–220. https://doi.org/10.1016/j.aquatox.2014.01.010

Barbour, M. T., Faulkner, C., Gerritsen, J., Snyder, B. D., & Stribling, J. B. (1999). Rapid bioassessment protocols for use in streams and wadeable rivers: Periphyton, benthic macroinvertebrates, and fish. In *Rapid Bioassessment Protocols for Use in Streams and Rivers*. http://www.epa.gov/OWOW/monitoring/techmon.html

Barbour, M. T., Gerritsen, J., Griffith, G. E., Frydenborg, R., McCarron, E., White, J. S., & Bastian, M. L. (1996). A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society*, *15*(2), 185–211. https://doi.org/10.2307/1467948

Besley, C. H., & Chessman, B. C. (2008). Rapid biological assessment charts the recovery of stream macroinvertebrate assemblages after sewage discharges cease. *Ecological Indicators*, *8*(5), 625–638. https://doi.org/10.1016/j.ecolind.2007.09.003

Blair, B. D., Crago, J. P., Hedman, C. J., & Klaper, R. D. (2013). Pharmaceuticals and personal care products found in the Great Lakes above concentrations of environmental concern. *Chemosphere*, *93*(9), 2116–2123. https://doi.org/10.1016/j.chemosphere.2013.07.057

Bode, R. W., Novak, M. A., & Abele, L. E. (1991). Work plan for biological stream monitoring in new york state. In *NYS Department of Environmental Conservation*.

Bradt, P., Urban, M., Goodman, N., Bissell, S., & Spiegel, I. (1999). Stability and resilience in benthic macroinvertebrate assemblages. Impact of physical disturbance over twenty-five years. *Hydrobiologia*, *403*, 123–133. https://doi.org/10.1023/A:1003711928980

Cao, Y., Bark, A. W., & Williams, W. P. (1996). Measuring the responses of macroinvertebrate communities to water pollution: A comparison of multivariate approaches, biotic and diversity indices. *Hydrobiologia*, *341*(1), 1–19. https://doi.org/10.1007/BF00012298

Carey, R. O., & Migliaccio, K. W. (2009). Contribution of wastewater treatment plant effluents to nutrient dynamics in aquatic systems. *Environmental Management*, *44*(2), 205–217. https://doi.org/10.1007/s00267-009-9309-5

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications*, *8*(3), 559–568. https://doi.org/10.1890/1051-0761(1998)008[0559:NPOSWW]2.0.CO;2

Cattaneo, A. (1987). Periphyton in lakes of different trophy. *Canadian Journal of Fisheries and Aquatic Sciences*, *44*(2), 296–303. https://doi.org/10.1139/f87-038

Chadwick, J. W., Canton, S. P., & Dent, R. L. (1986). Recovery of benthic invertebrate communities in silver bow creek, montana, following improved metal mine wastewater treatment. *Water, Air, & Soil Pollution*, *28*(3–4), 427–438. https://doi.org/10.1007/BF00583505

Clements, W. H., Herbst, D. B., Hornberger, M. I., Mebane, C. A., & Short, T. M. (2021). Long-term monitoring reveals convergent patterns of recovery from mining contamination across 4 western US watersheds. *Freshwater Science*, *40*(2), 407–426. https://doi.org/10.1086/714575

Commission for Environmental Cooperation (CEC). (1997). *Ecological regions of North America: toward a common perspective*. Communications and Public Outreach Department of the CEC Secretariat. http://www.cec.org/publications/ecological-regions-of-north-america/#:~:text=Toward a Common Perspective,-November 1%2C 1997&text=Climatic types range from the,variety of scenic natural beauty.

Crawford, C. G., Wangsness, D. J., & Martin, J. D. (1992). Recovery of benthic-invertebrate communities in the White River near Indianapolis, Indiana, USA, following implementation of advanced treatment of municipal wastewater. *Archiv Für Hydrobiologie*, *126*(1), 67–84.

Delong, M. D., & Brusven, M. A. (1998). Macroinvertebrate community structure along the longitudinal gradient of an agriculturally impacted stream. *Environmental Management*, *22*(3), 445–457. https://doi.org/10.1007/s002679900118

Downes, B. J., Barmuta, L. A., Fairweather, P. G., Faith, D. P., Keough, M. J., Lake, P. S., Mapstone, B. D., & Quinn, G. P. (2002). *Monitoring Ecological Impacts* (1st ed.). Cambridge University Press.

Eklöv, A. ., Greenberg, L. ., Brönmark, C., Larsson, P., & Berglund, O. (1998). Response of stream fish to improved water quality: a comparison Between the 1960S and 1990S. *Freshwater Biology*, *40*, 771–782.

Ellis, L. E., & Jones, N. E. (2013). Longitudinal trends in regulated rivers: A review and synthesis within the context of the serial discontinuity concept. *Environmental Reviews, 21*(3), 136–148. https://doi.org/10.1139/er-2012-0064

Environment and Climate Change Canada (ECCC). (2012a). *Canadian Aquatic Biomonitoring Network field manual: Wadeable streams. Retrieved September 2020*. https://mspace.lib.umanitoba.ca/xmlui/bitstream/handle/1993/30905/CABINFieldManual_EN_2012.pdf?sequence=1&isAllowed=y

Environment and Climate Change Canada (ECCC). (2012b). *Metal mining technical guidance for EEM. Retrieved September 2020*. https://www.ec.gc.ca/esee-eem/AEC7C481-D66F-4B9B-BA08-A5DC960CDE5E/COM-1434---Tec-Guide-for-Metal-Mining-Env-Effects-Monitoring_En_02[1].pdf

Environment and Climate Change Canada (ECCC). (2020). *CABIN laboratory methods - processing, taxonomy, and quality control of benthic macroinvertebrate samples. Retrieved September 2020*. publications.gc.ca/pub?id=9.895039&sl=0

Folke, C., Carpenter, S., Walker, B., Scheffer, M., Elmqvist, T., Gunderson, L., & Holling, C. S. (2004). Regime shifts, resilience, and biodiversity in ecosystem management. *Annual Review of Ecology, Evolution, and Systematics*, *35*, 557–581. https://doi.org/10.1146/annurev.ecolsys.35.021103.105711

Foster, A. D., Claeson, S. M., Bisson, P. A., & Heimburg, J. (2020). Aquatic and riparian ecosystem recovery from debris flows in two western Washington streams, USA. *Ecology and Evolution*, *10*(6), 2749–2777. https://doi.org/10.1002/ece3.5919

Frost, T. M., Fischer, J. M., Klug, J. L., Arnott, S. E., & Montz, P. K. (2006). Trajectories of zooplankton recovery in the Little Rock Lake whole-lake acidification experiment. *Ecological Applications*, *16*(1), 353–367. https://doi.org/10.1890/04-1800

Fuzzen, M. L. M., Bragg, L. M., Tetreault, G. R., Bahamonde, P. A., Tanna, R. N., Bennett, C. J., McMaster, M. E., & Servos, M. R. (2016). An assessment of the spatial and temporal variability of biological responses to municipal wastewater effluent in rainbow darter (Etheostoma caeruleum) collected along an urban gradient. *PLoS ONE*, *11*(10), 1–31. https://doi.org/10.1371/journal.pone.0164879

Gillis, P. L., Gagné, F., Mcinnis, R., Hooey, T. M., Choy, E. S., André, C., Hoque, M. E., & Metcalfe, C. D. (2014). The impact of municipal wastewater effluent on field-deployed freshwater mussels in the Grand River (Ontario, Canada). *Environmental Toxicology and Chemistry*, *33*(1), 134–143. https://doi.org/10.1002/etc.2401

Gillis, P. L., McInnis, R., Salerno, J., de Solla, S. R., Servos, M. R., & Leonard, E. M. (2017a). Freshwater mussels in an urban watershed: Impacts of anthropogenic inputs and habitat alterations on populations. *Science of the Total Environment*, *574*, 671–679. https://doi.org/10.1016/j.scitotenv.2016.09.110

Gillis, P. L., McInnis, R., Salerno, J., de Solla, S. R., Servos, M. R., & Leonard, E. M. (2017b). Municipal wastewater treatment plant effluent-induced effects on freshwater mussel populations and the role of mussel refugia in recolonizing an extirpated reach. *Environmental Pollution*, *225*, 460–468. https://doi.org/10.1016/j.envpol.2017.03.010

Grand River Conservation Authority (GRCA). (2018). *Geology of the Grand River Watershed.* Retrieved April 2021. https://www.grandriver.ca/en/our-watershed/resources/Documents/Groundwater/Watershed-Geology_March272019.pdf

Grand River Conservation Authority (GRCA). (2022). *GRCA dams*. https://www.grandriver.ca/en/our-watershed/GRCA-dams.aspx

Grantham, T. E., Cañedo-Argüelles, M., Perrée, I., Rieradevall, M., & Prat, N. (2012). A mesocosm approach for detecting stream invertebrate community responses to treated wastewater effluent. *Environmental Pollution*, *160*(1), 95–102. https://doi.org/10.1016/j.envpol.2011.09.014

Green, R. H. (1979). *Sampling design and statistical methods for environmental biologists*. John Wiley & Sons, Inc.

Gücker, B., Brauns, M., Pusch, M. T., Journal, S., American, N., Society, B., & June, N. (2006). Effects of wastewater treatment plant discharge on ecosystem structure and function of lowland streams. *Journal of the North American Benthological Society*, *25*(2), 313–329.

Gulis, V., & Suberkropp, K. (2003). Leaf litter decomposition and microbial activity in nutrient-enriched and unaltered reaches of a headwater stream. *Freshwater Biology*, *48*(1), 123–134. https://doi.org/10.1046/j.1365-2427.2003.00985.x

Hall, L. W., Anderson, R. D., & Killen, W. D. (2016). Spatiotemporal trends analysis of benthic communities and physical habitat during non-severe drought and severe-drought years in a residential creek in California. *California Fish and Game*, *102*(2), 55–70.

Hauer, F. R., & Lamberti, G. A. (1996). *Methods in Stream Ecology*. Academic Press.

Hicks, K. A. (2017). *Response of wild fish to municipal wastewater treatment plant upgrades*. Doctoral thesis. University of Waterloo, Waterloo, Canada.

Hicks, K. A., Fuzzen, M. L. M., McCann, E. K., Arlos, M. J., Bragg, L. M., Kleywegt, S., Tetreault, G. R., McMaster, M. E., & Servos, M. R. (2017). Reduction of intersex in a wild fish population in response to major municipal wastewater treatment plant upgrades. *Environmental Science and Technology*, *51*(3), 1811–1819. https://doi.org/10.1021/acs.est.6b05370

Hicks, K. A., Loomer, H. A., Fuzzen, M. L. M., Kleywegt, S., Tetreault, G. R., McMaster, M. E., & Servos, M. R. (2017). δ15N tracks changes in the assimilation of sewage-derived nutrients into a riverine food web before and after major process alterations at two municipal wastewater treatment plants. *Ecological Indicators*, *72*, 747–758. https://doi.org/10.1016/j.ecolind.2016.09.011

Hilsenhoff, W. L. (1977). Use of arthropods to evaluate water quality of streams. *Technical Bulletin*, *100*(100), 1–40.

Hilsenhoff, W. L. (1988). Seasonal correction factors for the biotic index. *Great Lakes Entomologist*, *21*(1), 9–14.

Hood, J. L. A., Taylor, W. D., & Schiff, S. L. (2014). Examining the fate of WWTP effluent nitrogen using δ15N-NH4+, δ15N-NO3-and δ15N of submersed macrophytes. *Aquatic Sciences*, *76*(2), 243–258. https://doi.org/10.1007/s00027-013-0333-4

Hynes, H. B. N. (1970). *The ecology of running waters*. Blackburn Press.

Johnson, A. C. (2019). Is freshwater macroinvertebrate biodiversity being harmed by synthetic chemicals in municipal wastewater? *Current Opinion in Environmental Science and Health*, *11*, 8–12. https://doi.org/10.1016/j.coesh.2019.05.005

Johnson, A. C., Jürgens, M. D., Edwards, F. K., Scarlett, P. M., Vincent, H. M., & von der Ohe, P. (2019). What works? The influence of changing wastewater treatment type, including tertiary granular activated charcoal, on downstream macroinvertebrate biodiversity over time. *Environmental Toxicology and Chemistry*, *38*(8), 1820–1832. https://doi.org/10.1002/etc.4460

Kotalik, C. J., Cadmus, P., & Clements, W. H. (2021). Before-After Control-Impact field surveys and novel experimental approaches provide valuable insights for characterizing stream recovery from acid mine drainage. *Science of the Total Environment*, *771*, 145419. https://doi.org/10.1016/j.scitotenv.2021.145419

Krynak, E. M., & Yates, A. G. (2018). Benthic invertebrate taxonomic and trait associations with land use in an intensively managed watershed: Implications for indicator identification. *Ecological Indicators*, *93*(June), 1050–1059. https://doi.org/10.1016/j.ecolind.2018.06.002

Lake Erie Source Protection Region Technical Team. (2008). *Grand River Watershed Characterization Report*. https://www.sourcewater.ca/en/source-protection-areas/resources/Documents/Grand/Grand_Reports_Characterization.pdf

Larsen, S., & Ormerod, S. J. (2014). Anthropogenic modification disrupts species co-occurrence in stream invertebrates. *Global Change Biology*, *20*(1), 51–60. https://doi.org/10.1111/gcb.12355

Layer, K., Hildrew, A. G., & Woodward, G. (2013). Grazing and detritivory in 20 stream food webs across a broad pH gradient. *Oecologia*, *171*(2), 459–471. https://doi.org/10.1007/s00442-012-2421-x

Lecerf, A., Usseglio-Polatera, P., Charcosset, J. Y., Lambrigot, D., Bracht, B., & Chauvet, E. (2006). Assessment of functional integrity of eutrophic streams using litter breakdown and benthic macroinvertebrates. *Archiv Fur Hydrobiologie*, *165*(1), 105–126. https://doi.org/10.1127/0003-9136/2006/0165-0105

Ledger, M. E., & Hildrew, A. G. (2005). The ecology of acidification and recovery: Changes in herbivore-algal food web linkages across a stream pH gradient. *Environmental Pollution*, *137*(1), 103–118. https://doi.org/10.1016/j.envpol.2004.12.024

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, *129*(2), 271–280. https://doi.org/10.1007/s004420100716

LGL Limited. (2010). *Surface water quality monitoring program for the grand river: 2009 year end report*. Prepared for the Region of Waterloo. LGL File TA4810.

LGL Limited. (2013). *Surface water quality monitoring program for the grand river: 2012 year end report*. Prepared for the Region of Waterloo. LGL File TA4813.

LGL Limited. (2016). *Surface water quality monitoring program for the grand river: 2015 year end report*. Prepared for the Region of Waterloo. LGL File TA8502.

LGL Limited. (2019). *Surface water quality monitoring program for the Grand and Nith rivers: 2018 year end report*. Prepared for the Region of Waterloo. LGL File TA8688-18.

Loomer, H. A., & Cooke, S. E. (2011). *Water quality in the Grand River Watershed: Current conditions & trends (2003-2008). Grand River Conservation Authority. Retrieved January 2021*. https://www.sourcewater.ca/en/source-protection-areas/resources/Documents/Grand/Grand_Reports_WaterQuality_2011.pdf

Loomer, H. A., Oakes, K. D., Schiff, S. L., Taylor, W. D., & Servos, M. R. (2015). Use of stable isotopes to trace municipal wastewater effluents into food webs within a highly developed river system. *River Research and Applications*, *31*, 1093–1100. https://doi.org/10.1002/rra

Loos, R., Carvalho, R., António, D. C., Comero, S., Locoro, G., Tavazzi, S., Paracchini, B., Ghiani, M., Lettieri, T., Blaha, L., Jarosova, B., Voorspoels, S., Servaes, K., Haglund, P., Fick, J., Lindberg, R. H., Schwesig, D., & Gawlik, B. M. (2013). EU-wide monitoring survey on emerging polar organic contaminants in wastewater treatment plant effluents. *Water Research*, *47*(17), 6475–6487. https://doi.org/10.1016/j.watres.2013.08.024

Lydy, M. J., Crawford, C. G., & Frey, J. W. (2000). A comparison of selected diversity, similarity, and biotic indices for detecting changes in benthic-invertebrate community structure and stream quality. *Archives of Environmental Contamination and Toxicology*, *39*(4), 469–479. https://doi.org/10.1007/s002440010129

Manfrin, A., Larsen, S., Traversetti, L., Pace, G., & Scalici, M. (2013). Longitudinal variation of macroinvertebrate communities in a Mediterranean river subjected to multiple anthropogenic stressors. *International Review of Hydrobiology*, *98*(3), 155–164. https://doi.org/10.1002/iroh.201201605

Marjan, P., Van Der Kraak, G. J., MacLatchy, D. L., Fuzzen, M. L. M., Bragg, L. M., McMaster, M. E., Tetreault, G. R., & Servos, M. R. (2018). Assessing recovery of in vitro steroid production in male rainbow darter (*Etheostoma caeruleum*) in response to municipal wastewater treatment plant infrastructure changes. *Environmental Toxicology and Chemistry*, *37*(2), 501–514. https://doi.org/10.1002/etc.3986

Marti, E., Aumatell, J., Godé, L., Poch, M., & Sabater, F. (2004). Nutrient retention efficiency in streams receiving inputs from wastewater treatment plants. *Journal of Environmental Quality*, *33*(1), 285–293. https://doi.org/10.2134/jeq2004.2850

Milner, A. M., & Oswood, M. W. (2000). Urbanization gradients in streams of Anchorage, Alaska: A comparison of multivariate and multimetric approaches to classification. *Hydrobiologia*, *422–423*, 209–223. https://doi.org/10.1007/978-94-011-4164-2_17

Milner, Alexander M., Conn, S. C., & Brown, L. E. (2006). Persistence and stability of macroinvertebrate communities in streams of Denali National Park, Alaska: Implications for biological monitoring. *Freshwater Biology*, *51*(2), 373–387. https://doi.org/10.1111/j.1365-2427.2005.01488.x

Mor, J. R., Dolédec, S., Acuña, V., Sabater, S., & Muñoz, I. (2019). Invertebrate community responses to urban wastewater effluent pollution under different hydro-morphological conditions. *Environmental Pollution*, *252*, 483–492. https://doi.org/10.1016/j.envpol.2019.05.114

Nikel, K. E. (2020). *Wild fish responses to wastewater treatment plant upgrades in the Grand River, Ontario*. Master's thesis. University of Waterloo.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2020). *vegan: Community ecology package*. https://cran.r-project.org/package=vegan

Oliver, A. A., Bogan, M. T., Herbst, D. B., & Dahlgren, R. A. (2012). Short-term changes in-stream macroinvertebrate communities following a severe fire in the Lake Tahoe basin, California. *Hydrobiologia*, *694*(1), 117–130. https://doi.org/10.1007/s10750-012-1136-7

Ortiz, J. D., & Puig, M. A. (2007). Point source effects on density, biomass and diversity of benthic macroinvertevrates in a Mediterranean stream. *River Research and Applications*, *23*(2), 155–170. https://doi.org/https://doi.org/10.1002/rra.971

Paul, M. J., & Meyer, J. L. (2001). Streams in the urban landscape. *Annual Review of Ecology and Systematics*, *32*(1), 333–365.

Poulton, B. C., Graham, J. L., Rasmussen, T. J., & Stone, M. L. (2015). Responses of macroinvertebrate community metrics to a wastewater discharge in the Upper Blue River of

Kansas and Missouri, USA. *Journal of Water Resource and Protection*, *07*(15), 1195–1220. https://doi.org/10.4236/jwarp.2015.715098

Rader, R. B., Voelz, N. J., & Ward, J. V. (2008). Post-Flood recovery of a macroinvertebrate community in a regulated River: Resilience of an anthropogenically altered ecosystem. *Restoration Ecology*, *16*(1), 24–33. https://doi.org/10.1111/j.1526-100X.2007.00258.x

Region of Waterloo. (2007). *Wastewater treatment master plan, final report.* Retrieved January 2021 . https://www.regionofwaterloo.ca/en/living-here/resources/Documents/water/plans/WS2007AugWWTPMasterPlanFinalReport.pdf

Region of Waterloo. (2018). *Region of Waterloo 2018 wastewater treatment master plan. Retrieved December 2020*. https://www.regionofwaterloo.ca/en/living-here/resources/Documents/water/projects/wastewater/plan/WS2018Wastewater_Treatment_Master_Plan_WWTMP_Final_Report.pdf

Region of Waterloo. (2020). *Water and wastewater monitoring report. Retrieved December 2020*. https://www.regionofwaterloo.ca/en/living-here/resources/Documents/water/reports/WS2020WaterAndWastewaterMonitoringReport.PDF

Sánchez-Murillo, R., Brooks, E. S., Sampson, L., Boll, J., & Wilhelm, F. (2014). Ecohydrological analysis of Steelhead (*Oncorhynchus mykiss*) habitat in an effluent dependent stream in the Pacific Northwest, USA. *Ecohydrology*, *7*(2), 557–568. https://doi.org/10.1002/eco.1376

Servos, M., Dhiyebi, H., Tetreault, G. R., Schiff, S. L., Venkiteswaran, J. J., Yates, A. G., & Van Der Kraak, G. (2015). *A regional monitoring framework for supporting cumulative effects assessment and adaptive management in the Grand River*. 1–77.

Shieh, S. H., Kondratieff, B. C., & Ward, J. V. (1999). Longitudinal changes in benthic organic matter and macroinvertebrates in a polluted Colorado plains stream. *Hydrobiologia*, *411*(1989), 191–209. https://doi.org/10.1023/A:1003835929331

Smith, J. G., Brandt, C. C., & Christensen, S. W. (2011). Long-term benthic macroinvertebrate community monitoring to assess pollution abatement effectiveness. *Environmental*

Management, *47*(6), 1077–1095. https://doi.org/10.1007/s00267-010-9610-3

Sonthiphand, P., Cejudo, E., Schiff, S. L., & Neufeld, J. D. (2013). Wastewater effluent impacts ammonia-oxidizing prokaryotes of the Grand River, Canada. *Applied and Environmental Microbiology*, *79*(23), 7454–7465. https://doi.org/10.1128/AEM.02202-13

Spänhoff, B., Bischof, R., Böhme, A., Lorenz, S., Neumeister, K., Nöthlich, A., & Küsel, K. (2007). Assessing the impact of effluents from a modern wastewater treatment plant on breakdown of coarse particulate organic matter and benthic macroinvertebrates in a lowland river. *Water, Air, and Soil Pollution*, *180*(1–4), 119–129. https://doi.org/10.1007/s11270-006-9255-2

Srikanthan, N. (2019). *Analysis of temporal changes in estrogenic compounds released from municipal wastewater treatment plants.* Master's thesis. University of Waterloo.

Suckling, D. M. (1982). Organic wastewater effects on benthic invertebrates in the Manawatu River. *New Zealand Journal of Marine and Freshwater Research*, *16*(3–4), 263–270. https://doi.org/10.1080/00288330.1982.9515969

Sundermann, A., Gerhardt, M., Kappes, H., & Haase, P. (2013). Stressor prioritisation in riverine ecosystems: Which environmental factors shape benthic invertebrate assemblage metrics? *Ecological Indicators*, *27*, 83–96. https://doi.org/10.1016/j.ecolind.2012.12.003

Tetreault, G. R., Bennett, C. J., Shires, K., Knight, B., Servos, M. R., & McMaster, M. E. (2011). Intersex and reproductive impairment of wild fish exposed to multiple municipal wastewater discharges. *Aquatic Toxicology*, *104*(3–4), 278–290. https://doi.org/10.1016/j.aquatox.2011.05.008

Tetreault, G. R., Brown, C. J. M., Bennett, C. J., Oakes, K. D., McMaster, M. E., & Servos, M. R. (2013). Fish community responses to multiple municipal wastewater inputs in a watershed. *Integrated Environmental Assessment and Management*, *9*(3), 456–468. https://doi.org/10.1002/ieam.1364

UN-Water. (2015). *Wastewater management: A UN-Water analytical brief. Retrieved March 2021*. https://www.unwater.org/publications/wastewater-management-un-water-analytical-brief/

US Environmental Protection Agency (EPA). (2012). *Freshwater biological traits database, final report 2012* (Issue June). https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=241813&Lab=NCEA

Venkiteswaran, J. J., Schiff, S. L., & Taylor, W. D. (2015). Linking aquatic metabolism, gas exchange, and hypoxia to impacts along the 300-km Grand River, Canada. *Freshwater Science*, *34*(4), 1216–1232. https://doi.org/10.1086/683241

Verdonschot, P. F. M. (2006). Data composition and taxonomic resolution in macroinvertebrate stream typology. *Hydrobiologia*, *566*(1), 59–74. https://doi.org/10.1007/s10750-006-0070-y

Wallace, J. B., & Webster, J. R. (1996). The role of macroinvertebrates in stream ecosystem function. *Annual Review of Entomology*, *41*(1), 115–139. https://doi.org/10.1146/annurev.en.41.010196.000555

Walsh, C. J., Fletcher, T. D., & Ladson, A. R. (2005). Stream restoration in urban catchments through redesigning stormwater systems: Looking to the catchment to save the stream. *Journal of the North American Benthological Society*, *24*(3), 690–705. https://doi.org/10.1899/04-020.1

Webster, J. S., Gido, K. B., Hedden, S. C., Propst, D. L., & Whitney, J. E. (2022). Response of arid-land macroinvertebrate communities to extremes of drought, wildfire, and monsoonal flooding. *River Research and Applications*, *January*, 1–14. https://doi.org/10.1002/rra.3949

Wellnitz, T. A., & Ward, J. V. (1998). Does light intensity modify the effect mayfly grazers have on periphyton? *Freshwater Biology*, *39*(1), 135–149. https://doi.org/10.1046/j.1365-2427.1998.00270.x

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Yates, A. G., & Bailey, R. C. (2011). Effects of taxonomic group, spatial scale and descriptor on the relationship between human activity and stream biota. *Ecological Indicators*, *11*(3), 759–771. https://doi.org/10.1016/j.ecolind.2010.09.003

Appendices

Appendix A

Supplemental Information

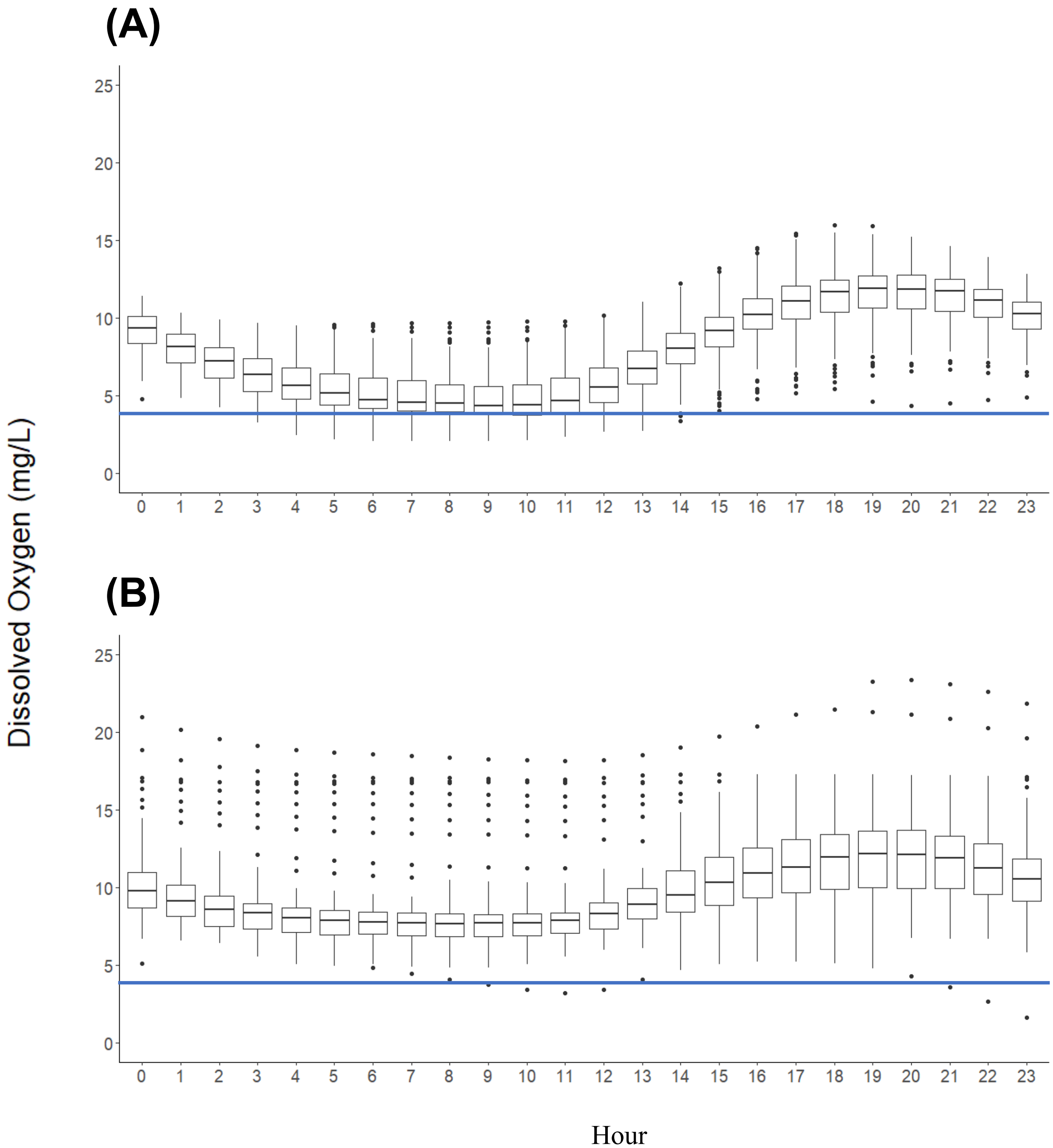

Figure A1. Boxplot of dissolved oxygen concentrations for each hour in 2012 (A) and 2013 (B) from Jun – September at the Blair water quality station (Kitchener wastewater treatment plant at the DSF site) sourced from the Grand River Conservation Authority's Monitoring Data (https://data.grandriver.ca/downloads-monitoring.html).

Table A1. Population served for each year at each wastewater treatment plant.

Year	Ayr	Hespeler	Kitchener	New Hamburg	Preston	Waterloo	Wellesley
2006	3,989	21,972	210,854	9,804	20,357	119,406	2,279
2007	4,018	22,117	215,247	10,319	20,559	120,265	2,420
2008	4,088	22,166	219,596	10,742	20,646	121,413	2,556
2009	4,195	23,163	221,223	11,016	20,682	124,006	2,700
2010	4,209	24,333	226,106	11,467	20,257	126,029	2,849
2011	4,255	24,646	227,761	11,773	20,409	127,688	2,921
2012	4,658	25,239	231,488	12,268	20,174	131,776	3,191
2013	4,822	25,595	230,922	12,575	20,415	134,851	3,211
2014	4,879	25,737	234,466	12,787	20,656	136,179	3,270
2015	4,952	25,759	237,417	12,978	20,722	137,322	3,353
2016	5,175	25,845	240,669	13,252	21,079	138,464	3,408
2017	5,339	25,699	248,481	13,657	21,469	145,381	3,495
2018	5,524	25,991	253,621	14,043	22,517	153,271	3,523
2019	5,724	25,821	258,675	14,239	23,466	150,283	3,604
2020	5,784	25,970	260,556	14,543	24,209	141,902	3,652

Table A2. Average measured flow (m^3/day) each year at each wastewater treatment plant.

Year	Ayr	Hespeler	Kitchener	New Hamburg	Preston	Waterloo	Wellesley
2006	1,262	8,013	74,344	3,620	12,234	46,012	758
2007	1,216	7,252	70,051	3,252	11,015	41,358	539
2008	1,315	8,056	74,935	4,066	12,767	47,562	710
2009	1,395	7,929	73,002	3,763	11,945	45,940	648
2010	1,258	7,462	64,329	3,235	8,754	42,007	595
2011	1,277	7,666	70,443	3,844	9,109	45,540	771
2012	1,235	6,660	65,858	3,367	8,463	42,104	608
2013	1,350	7,337	72,433	3,953	9,107	48,570	812
2014	1,313	6,808	70,988	3,736	9,168	48,242	831
2015	1,269	6,435	64,136	3,320	8,450	38,391	686
2016	1,322	6,500	65,247	3,532	8,646	39,750	806
2017	1,505	6,692	68,684	4,039	9,109	42,473	923
2018	1,319	6,320	67,902	4,039	9,239	41,805	800
2019	1,247	6,259	68,080	3,889	9,342	39,331	862
2020	1,394	6,077	65,604	3,747	9,005	38,242	856

Table A3. Description of the treatment process at the Ayr wastewater treatment plant, as well as any changes incurred from 2009-2018, the period during which benthic macroinvertebrate samples occurred (RoW, 2018).

Treatment	**Description**	**Changes from 2009-2018**
Process	Extended aeration package	N/A
Primary	Screening; aerated grit removal	N/A
Secondary	Aeration; secondary clarification	N/A
Tertiary	Tertiary filtration	N/A
Disinfection	UV	N/A
Biosolids	Aerobic digestion; dewatered; liquid-digested sludge stored on-stie in biosolid lagoons before land application	N/A

Table A4. Description of the treatment process at the Hespeler wastewater treatment plant (WWTP), as well as any changes incurred from 2009-2018, the period during which benthic macroinvertebrate samples occurred (RoW, 2018).

Treatment	Description	Changes from 2009-2018
Process	Extended aeration	N/A
Primary	Aerated grit removal	N/A
Secondary	Aeration; secondary clarification; alum added prior to clarifiers for phosphorous removal	N/A
Tertiary	N/A	N/A
Disinfection	Sodium hypochlorite; dechloronation	N/A
Biosolids	Sent to aerated sludge holding tank and gravity thickened and sent to Kitchener/Galt WWTPs for treatment	N/A

Table A5. Description of the treatment process at the New Hamburg wastewater treatment plant (WWTP), as well as any changes incurred from 2009-2018, the period during which benthic macroinvertebrate samples occurred (RoW, 2018).

Treatment	**Description**	**Changes from 2009-2018**
Process	Sequencing batch reactor (SBR)	N/A
Primary	Screening; vortex grit removal	N/A
Secondary	Aeration; secondary clarification; flow equalization	N/A
Tertiary	Tertiary sand filtration	N/A
Disinfection	UV	N/A
Biosolids	Stored in sludge holding tank and gravity thickened and sent to Kitchener/Galt WWTPs for treatment	N/A

Table A6. Description of the treatment process at the Preston wastewater treatment plant (WWTP), as well as any changes incurred from 2009-2018, the period during which benthic macroinvertebrate samples occurred (RoW, 2018).

Treatment	Description	Changes from 2009-2018
Process	Conventional activated sludge	N/A
Primary	Screening, vortex grit removal, primary clarification	N/A
Secondary	Aeration; secondary clarification; alum added at outlet of aeration tanks for phosphorus removal	Digestion upgrades (2013)
Tertiary	N/A	N/A
Disinfection	UV	N/A
Biosolids	Anaerobically digested and dewatered	Biosolids hauled to Galt WWTP for dewatering (2012)
Industrial Road Service Area	Treated at the Preston WWTP	Diverted to the Galt WWTP (2011)

Table A7. Description of the treatment process at the Wellesley wastewater treatment plant (WWTP), as well as any changes incurred from 2009-2018, the period during which benthic macroinvertebrate samples occurred (RoW, 2018).

Treatment	Description	Changes from 2009-2018
Process	Extended aeration package	N/A
Primary	Manually-cleaned screens	N/A
Secondary	Aeration; secondary clarification; alum added to aeration tank effluent for phosphorous removal	N/A
Tertiary	Tertiary filtration	N/A
Disinfection	Ozone	N/A
Biosolids	Stored in sludge holding tank and gravity thickened and sent to Kitchener/Galt WWTPs for treatment	N/A

Table A8. Description of the treatment process at the Kitchener wastewater treatment plant (WWTP), as well as any changes incurred from 2009-2018, the period during which benthic macroinvertebrate samples occurred (RoW, 2018).

Treatment	Description	Changes from 2009-2018
Process	Conventional activated sludge	N/A
Primary	Screening, vortex grit removal, primary clarification (Add ferric sulfate prior to primary clarification to remove P)	N/A
Secondary	Secondary clarification	New tanks/aeration – Plant 2 – Pass 1 (2012) New aeration – Plant 1 (2012) New tanks/aeration – Plant 2 – Passes 2 and 3 (2013)
Tertiary	N/A	Tertiary filters commissioned (Plant 1 not being filtered; 2017) Optimized coagulant dose for TP control (2017)
Disinfection	UV disinfection	N/A
Biosolids	Anaerobically digested and dewatered via centrifuge	N/A

Table A9. Description of the treatment process at the Waterloo wastewater treatment plant (WWTP), as well as any changes incurred from 2009-2018, the period during which benthic macroinvertebrate samples occurred (RoW, 2018).

Treatment	Description	Changes from 2009-2018
Process	Conventional activated sludge	N/A
Primary	Screening, vortex grit removal, primary clarification (Add ferric chloride prior to primary clarification, ferrous chloride after)	N/A
Secondary	Secondary clarification	RAS re-aeration online (2015) Aeration tank 1 upgrades (2017) Aeration tank 2 online (2018)
Tertiary	N/A	N/A
Disinfection	UV disinfection	N/A
Biosolids	Anaerobically digested and dewatered via centrifuge	N/A

Appendix B

Statistical Analyses

Table A10. Summary of PERMANOVA results on the US (upstream), DSN (downstream-near), and DSF (downstream-far) benthic macroinvertebrate assemblages adjacent to the Waterloo wastewater treatment plant, as well as post-hoc, pairwise PERMANOVAs comparing sites within years. Additionally presented is the average Bray-Curtis dissimilarity between US & DSN, and US & DSF within years.

Waterloo - Analysis	Group	p-value
Two-way PERMANOVA	Site	0.0001
	Year	0.0001
	Site*Year	0.0002
Pairwise PERMANOVA (US vs DSN)	2009	0.0232
	2012	0.0069
	2015	0.0076
	2018	0.0324
Pairwise PERMANOVA (US vs DSF)	2009	0.0634
	2012	0.0074
	2015	0.0309
	2018	0.0076
Average Dissimilarity	**Group**	**Value**
Upstream vs. Downstream-near	2009	0.46 (± 0.11)
	2012	0.60 (± 0.14)
	2015	0.52 (± 0.12)
	2018	0.35 (± 0.08)
Upstream vs. Downstream-far	2009	0.44 (± 0.06)
	2012	0.48 (± 0.11)
	2015	0.42 (± 0.08)
	2018	0.33 (± 0.04)

Table A11. Summary of results from General Linear Model (GLM) analysis on average modified family biotic index (MFBI) values, % scrapers, % filterers, and family richness using benthic macroinvertebrate assemblage data collected from the US (upstream), DSN (downstream-near), and DSF (downstream-far) sites adjacent to the Waterloo wastewater treatment plant. Additionally, presented are the results of post-hoc tests using estimated marginal means between upstream and downstream sites.

Metric (GLM)	Group	p-value
MFBI (GLM)	Site	< 0.001
	Year	0.070
	Site*Year	0.211
% Scrapers (GLM)	Site	< 0.001
	Year	0.030
	Site*Year	0.044
% Scrapers (Post-hoc Test; US vs DSN)	2009	0.014
	2012	< 0.001
	2015	< 0.001
	2018	0.188
% Scrapers (Post-hoc Test; US vs DSF)	2009	0.020
	2012	0.002
	2015	0.321
	2018	0.449
% Filterers (GLM)	Site	< 0.001
	Year	0.001
	Site*Year	0.151
Family Richness (GLM)	Site	< 0.001
	Year	< 0.001
	Site*Year	0.013
Family Richness (Post-hoc Test; US vs DSN)	2009	< 0.001
	2012	< 0.001
	2015	0.003
	2018	0.025
Family Richness (Post-hoc Test; US vs DSF)	2009	< 0.001
	2012	0.008
	2015	0.446
	2018	0.328

Table A12. Summary of PERMANOVA results on the US (upstream), DSN (downstream-near), and DSF (downstream-far) benthic macroinvertebrate assemblages adjacent to the Kitchener wastewater treatment plant, as well as post-hoc, pairwise PERMANOVAs comparing sites within years. Additionally presented is the average Bray-Curtis dissimilarity between US & DSN, and US & DSF within years where applicable.

Kitchener - Analysis	**Group**	**p-value**
Two-way PERMANOVA	Site	< 0.001
	Year	< 0.001
	Site*Year	< 0.001
Pairwise PERMANOVA (US vs DSN)	2009	0.008
	2012	0.007
	2015	0.022
	2018	0.007
Pairwise PERMANOVA (US vs DSF)	2009	0.009
	2012	0.008
	2015	0.025
	2018	0.007
Average Dissimilarity	**Group**	**Value**
Upstream vs. Downstream-near	2009	0.49 (± 0.06)
	2012	0.40 (± 0.06)
	2015	0.41 (± 0.10)
	2018	0.34 (± 0.06)
Upstream vs. Downstream-far	2009	0.70 (± 0.08)
	2012	0.50 (± 0.11)
	2015	0.38 (± 0.05)
	2018	0.33 (± 0.05)

Table A13. Summary of results from General Linear Model (GLM) analysis on average modified family biotic index (MFBI) values, % scrapers, % filterers, and family richness using benthic macroinvertebrate assemblage data collected from the US the US (upstream), DSN (downstream-near), and DSF (downstream-far) sites adjacent to the Kitchener wastewater treatment plant. Additionally, presented are the results of post-hoc tests using estimated marginal means between upstream and downstream sites where applicable.

Metric	Group	p-value
MFBI (GLM)	Site	< 0.001
	Year	0.080
	Site*Year	< 0.001
MFBI (Post-hoc Test; US vs DSN)	2009	0.572
	2012	0.770
	2015	0.003
	2018	0.220
MFBI (Post-hoc Test; US vs DSF)	2009	< 0.001
	2012	0.119
	2015	0.572
	2018	0.112
% Scrapers (GLM)	Site	0.024
	Year	< 0.001
	Site*Year	0.156
% Filterers (GLM)	Site	< 0.001
	Year	0.005
	Site*Year	0.110
Family Richness (GLM)	Site	< 0.001
	Year	< 0.001
	Site*Year	0.013
Family Richness (Post-hoc Test; US vs DSN)	2009	0.047
	2012	0.030
	2015	0.161
	2018	0.192
Family Richness (Post-hoc Test; US vs DSF)	2009	< 0.001
	2012	< 0.001
	2015	0.047
	2018	0.420

Table A14. Summary of PERMANOVA results on the US (upstream) and DS (downstream) benthic macroinvertebrate assemblages adjacent to the Preston wastewater treatment plant.

Preston - Analysis	Group	p-value
Two-way PERMANOVA	Site	0.281
	Year	< 0.001
	Site*Year	0.107

Table A15. Summary of results from General Linear Model (GLM) analysis on average modified family biotic index (MFBI) values, % scrapers, % filterers, and family richness using benthic macroinvertebrate assemblage data collected from the US (upstream) and DS (downstream) sites adjacent to the Preston wastewater treatment plant. Additionally, presented are the results of post-hoc tests using estimated marginal means between upstream and downstream sites where applicable.

Metric	Group	p-value
MFBI (GLM)	Site	0.929
	Year	< 0.001
	Site*Year	0.828
% Scrapers (GLM)	Site	0.508
	Year	< 0.001
	Site*Year	0.577
% Filterers (GLM)	Site	0.341
	Year	< 0.001
	Site*Year	0.188
Family Richness (GLM)	Site	0.958
	Year	< 0.001
	Site*Year	0.018

Table A16. Summary of PERMANOVA results on the US (upstream) and DS (downstream) benthic macroinvertebrate assemblages adjacent to the Wellesley wastewater treatment plant.

Wellesley - Analysis	Group	p-value
Two-way PERMANOVA	Site	0.158
	Year	< 0.001
	Site*Year	0.510

Table A17. Summary of results from General Linear Model (GLM) analysis on average modified family biotic index (MFBI) values, % scrapers, % filterers, and family richness using benthic macroinvertebrate assemblage data collected from the US (upstream) and DS (downstream) sites adjacent to the Wellesley wastewater treatment plant.

Metric	Group	p-value
MFBI (GLM)	Site	0.504
	Year	0.203
	Site*Year	0.341
% Scrapers (GLM)	Site	0.692
	Year	< 0.001
	Site*Year	0.472
% Filterers (GLM)	Site	0.235
	Year	0.077
	Site*Year	0.165
Family Richness (GLM)	Site	0.003
	Year	0.027
	Site*Year	0.933

Table A18. Summary of PERMANOVA results on the US (upstream) and DS (downstream) benthic macroinvertebrate assemblages adjacent to the New Hamburg wastewater treatment plant, as well as post-hoc, pairwise PERMANOVAs comparing sites within years. Additionally presented is the average Bray-Curtis dissimilarity between upstream and downstream sites.

New Hamburg - Analysis	Group	p-value
Two-way PERMANOVA	Site	< 0.001
	Year	< 0.001
	Site*Year	0.056
Pairwise PERMANOVA (US vs DS)	2009	0.007
	2012	0.008
	2015	0.009
	2018	0.009
Average Dissimilarity	**Group**	**Value**
Upstream vs. Downstream-near	2009	0.57 (± 0.07)
	2012	0.35 (± 0.04)
	2015	0.56 (± 0.05)
	2018	0.29 (± 0.05)

Table A19. Summary of results from General Linear Model (GLM) analysis on average modified family biotic index (MFBI) values, % scrapers, % filterers, and family richness using benthic macroinvertebrate assemblage data collected from the US (upstream) and DS (downstream) sites adjacent to the New Hamburg wastewater treatment plant. Additionally, presented are the results of post-hoc tests using estimated marginal means between upstream and downstream sites where applicable.

Metric	Group	p-value
MFBI (GLM)	Site	< 0.001
	Year	< 0.001
	Site*Year	< 0.001
MFBI (Post-hoc Test; US vs DS)	2009	< 0.001
	2012	0.007
	2015	< 0.001
	2018	0.102
% Scrapers (GLM)	Site	0.876
	Year	0.235
	Site*Year	0.628
% Filterers (GLM)	Site	< 0.001
	Year	< 0.001
	Site*Year	< 0.001
% Filterers (Post-hoc Test; US vs DS)	2009	< 0.001
	2012	0.011
	2015	< 0.001
	2018	0.388
Family Richness (GLM)	Site	0.250
	Year	< 0.001
	Site*Year	0.876

Table A20. Summary of PERMANOVA results on the US (upstream) and DS (downstream) benthic macroinvertebrate assemblages adjacent to the Ayr wastewater treatment plant, as well as post-hoc, pairwise PERMANOVAs comparing sites within years. Additionally presented is the average Bray-Curtis dissimilarity between upstream and downstream sites.

Ayr - Analysis	**Group**	**p-value**
Two-way PERMANOVA	Site	0.011
	Year	< 0.001
	Site*Year	0.015
Pairwise PERMANOVA (US vs DS)	2012	0.014
	2015	0.009
	2018	0.049
Average Dissimilarity	**Group**	**Value**
Upstream vs. Downstream-near	2012	0.40 (± 0.05)
	2015	0.31 (± 0.04)
	2018	0.30 (± 0.06)

Table A21. Summary of results from General Linear Model (GLM) analysis on average modified family biotic index (MFBI) values, % scrapers, % filterers, and family richness using benthic macroinvertebrate assemblage data collected from the US (upstream) and DS (downstream) sites adjacent to the Ayr wastewater treatment plant. Additionally, presented are the results of post-hoc tests using estimated marginal means between upstream and downstream sites where applicable.

Metric	Group	p-value
MFBI (GLM)	Site	< 0.001
	Year	< 0.001
	Site*Year	0.271
% Scrapers (GLM)	Site	< 0.001
	Year	< 0.001
	Site*Year	0.009
% Scrapers (Post-hoc Test; US vs DS)	2012	< 0.001
	2015	0.334
	2018	0.070
% Filterers (GLM)	Site	0.039
	Year	< 0.001
	Site*Year	0.189
Family Richness (GLM)	Site	0.043
	Year	< 0.001
	Site*Year	0.860

Table A22. Summary of PERMANOVA results on the US (upstream) and DS (downstream) benthic macroinvertebrate assemblages adjacent to the Hespeler wastewater treatment plant.

Hespeler - Analysis	Group	p-value
Two-way PERMANOVA	Site	< 0.001
	Year	< 0.001
	Site*Year	0.306

Table A23. Summary of results from General Linear Model (GLM) analysis on average modified family biotic index (MFBI) values using benthic macroinvertebrate assemblage data collected from the US (upstream) and DS (downstream) sites adjacent to the Hespeler wastewater treatment plant. Additionally, presented are the results of post-hoc tests using estimated marginal means between upstream and downstream sites where applicable.

Hespeler - MFBI	Group	p-value
MFBI (GLM)	Site	0.013
	Year	0.671
	Site*Year	0.419
% Scrapers (GLM)	Site	0.210
	Year	< 0.001
	Site*Year	0.048
% Scrapers (Post-hoc Test; US vs DS)	2009	0.004
	2012	0.681
	2015	0.418
	2018	0.897
% Filterers (GLM)	Site	0.235
	Year	0.759
	Site*Year	0.002
% Filterers (Post-hoc Test; US vs DS)	2009	0.224
	2012	0.005
	2015	0.210
	2018	0.024
Family Richness (GLM)	Site	0.760
	Year	0.005
	Site*Year	0.023
Family Richness (Post-hoc Test; US vs DS)	2009	0.048
	2012	0.191
	2015	0.541
	2018	0.038