

***Identifying the microbial community and uncovering seasonal trends of Tellico
Reservoir, a reservoir in East Tennessee, USA***

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

David Junes Niknejad
August 2024

ACKNOWLEDGEMENTS

First, I would like to thank my principle investigator Dr. Steven Wilhelm for the support throughout my academic journey. I came to him as an undergraduate student and he was happy to help me kickstart my scientific career. He then allowed me to stay working in his laboratory and complete this master's thesis. I want to thank him for his support and advocacy for his students.

I would also like to thank my committee Dr. Forbes Walker, Dr. Erik Zinser, and Dr. Jennifer DeBruyn for the feedback throughout my degree. I am also grateful for the successful collaborations without this project would not be possible. Especially the citizens group WATeR, who chauffeured us around with their boat and spend their money to fuel it. I also want to thank Dr. Justin Chaffin, for running our nutrient samples.

In addition, I would like to thank everyone in the Wilhelm lab, as well as everyone on the 6th floor of SERF for all the laughs and the good support system. I thank especially Dr. Brittany Zepernick for training me as an undergrad and Dr. Robbie Martin for accompanying me on my Tellico trips. Finally, I would like to thank Rachelle Allen for administrative support and making sure that we all get paid.

ABSTRACT

Tellico reservoir is one of many manmade reservoirs in the Tennessee Valley Region. Rainfall upwards of 130 cm per year resulted in the flooding of cities in the Tennessee valley. To prevent flooding, the Tennessee Valley Authority (TVA) created several reservoirs that can hold the precipitation. Although the TVA does monitor the reservoirs for some ecological factors (including chlorophyll), the microbial status of the reservoir is widely unknown. In collaboration with a citizen group, Tellico Reservoir was sampled every three weeks from April until November 2023 for a total of 10 times. Six stations were selected and grouped into land use categories, which presumably influenced them. The primary land use inside the watershed of the land use categories correlated with the assigned label. The three land use categories were agriculture, developed and control. Primary land use inside the control category was forest. Physicochemical parameters showed a clear distinction between the three land use categories. It is unclear, however, if stations are influenced by land use, or distance to the main channel, due to the low retention time of the reservoir. Overall, chlorophyll was higher in stations influenced by agriculture. No differences were detected between control sites and stations influenced by urban development. This is also true for several nutrients that were measured. 16S/18S rRNA gene analyses showed *Actinobacteria*, *Cyanobacteria* and *Proteobacteria* were the most dominant phyla throughout the year, as well as at each land use category. Even with clear separation of physicochemical parameters, nMDS analyses of the microbial community show no trend between land use categories. Strong seasonal trends were observed. Spring is distinctly different than summer and fall, which are not different from each other. The most important parameter correlating with this separation was water temperature. With climate change increasing surface

water temperatures, it is vital to study how the microbial community responds so we can plan for the future.

TABLE OF CONTENTS

Chapter I: Introduction.....	1
Understanding the algal community of the Tennessee River	1
History of Tellico Reservoir	4
Microbial Diversity in the reservoir.....	6
TVA surveys	8
Changes in the reservoir.....	9
Synechococcus.....	10
Chapter II: Identifying the microbial community of Tellico Reservoir in 2023 during	
Summer Pool.....	12
Introduction.....	14
Methods.....	15
Sample sites and collection.....	15
Sample collection.....	15
Environmental metadata measurements	16
Measuring extracted chlorophyll <i>a</i>	16
DNA extraction and PCR amplification of gene sequences	19
16S/18S rRNA gene analysis.....	20
FlowCAM analysis	21
Data analysis	21
Results.....	22
Land use patterns	22
Variation within controls	22

Measurements of extracted chlorophyll <i>a</i> throughout the sampling season	23
Measurements of nutrients throughout the sampling season	31
Patterns in physical and chemical water properties	38
Results of amplicon sequencing	39
Analysis of 16S/18S rRNA gene sequences	39
Relative abundance of microbial organisms	39
Clustering of samples based of ASV presence and abundance highlighted by land use category and season	48
Physicochemical drivers of population diversity	48
Discussion	50
Site selection and implications for the study	50
Variation within controls	57
Seasonal chlorophyll <i>a</i> pattern	57
Seasonal nutrient dynamics	58
Physicochemical parameters driving differences between land use categories	59
Total community composition	60
Distance between land use categories in ordinate space	64
ASVs most contributing to similarity within a land use category and season	64
Physicochemical parameters correlating with species composition	67
Conclusion	67
Chapter III: Two genome sequences of <i>Synechococcus</i> sp. isolated from Tellico Reservoir, Tennessee, USA	70
Abstract	72

Announcement	72
Chapter IV: Conclusion	75
References.....	78
Appendix.....	82
Vita.....	101

LIST OF TABLES

Table 1 List of stations, coordinates, and land use	17
Table 2 Drainage area and percent land use of each station	24
Table 3 SIMPER results inside agriculture stations	52
Table 4 SIMPER results inside developed stations	52
Table 5 SIMPER results inside control stations	52
Table 6 SIMPER results between developed and control stations	53
Table 7 SIMPER results between developed and agriculture stations	53
Table 8 SIMPER results between control and agriculture stations.....	53
Table 9 SIMPER results inside the spring season	54
Table 10 SIMPER results inside the summer season.....	54
Table 11 SIMPER results inside the fall season	55
Table 12 SIMPER results between spring and summer seasons	55
Table 13 SIMPER results between spring and fall seasons.....	56
Table 14 SIMPER results between summer and fall seasons	56
Table 15 Tracking of 18S rRNA reads	83
Table 16 Tracking of 16S rRNA reads	86

LIST OF FIGURES

Figure 1 Map of the Mississippi River Watersheds.....	3
Figure 2 Sampling locations with presumed land use	18
Figure 3 Drainage area and land use of station Ag1	25
Figure 4 Drainage area and land use of station Ag2	26
Figure 5 Drainage area and land use of station D1	27
Figure 6 Drainage area and land use of station D2	28
Figure 7 Drainage area and land use of station C	29
Figure 8 Drainage area and land use of station C2	30
Figure 9 Seasonal chlorophyll <i>a</i> concentrations	32
Figure 10 Seasonal dissolved ammonium by station	33
Figure 11 Seasonal dissolved reactive phosphorus by station	34
Figure 12 Seasonal dissolved silicate by station	35
Figure 13 Seasonal dissolved nitrite by station.....	36
Figure 14 Seasonal dissolved nitrate by station	37
Figure 15 PCA of physicochemical data collected	40
Figure 16 Relative abundance of all microbes at agriculture sites	42
Figure 17 Relative abundance of microbes at developed sites	43
Figure 18 Relative abundance of microbes at control sites	44
Figure 19 Relative abundance of eukaryotes at agriculture sites.....	45
Figure 20 Relative abundance of eukaryotes at developed sites.....	46
Figure 21 Relative abundance of eukaryotes on control sites.....	47
Figure 22 nMDS of all samples highlighted by season	49

Figure 23 nMDS of all samples highlighted by land use.....	51
Figure 24 Phylogenetic tree of MAGs recovered	74
Figure 25 Recorded rainfall inside Tellico's watershed	82
Figure 26 Relative abundance of cyanobacterial genera at agriculture stations.....	89
Figure 27 Relative abundance of cyanobacterial genera at developed stations.....	90
Figure 28 Relative abundance of cyanobacterial genera at control stations	91
Figure 29 Principle component analysis (PCA) of environmental data collected from stations C1 and C2	92
Figure 30 nMDS of Bray-Curtis distance between microbial communities from stations C1 and C2.....	93
Figure 31 Correlation between the microbial communities of C1 and C2	94
Figure 32 Abundance of microbes captured by the FlowCAM at station Ag1	95
Figure 33 Abundance of microbes captured by the FlowCAM at station Ag2	96
Figure 34 Abundance of microbes captured by the FlowCAM at station D1	97
Figure 35 Abundance of microbes captured by the FlowCAM at station D2	98
Figure 36 Abundance of microbes captured by the FlowCAM at station C1.....	99
Figure 37 Abundance of microbes captured by the FlowCAM at station C2.....	100

Chapter I

Introduction

Understanding the algal community of the Tennessee River basin

The Tennessee River Basin is an ecological hotspot. It is considered one of the most important regions of biodiversity in the world. It holds more than half of the species of freshwater fishes, crawfish, mussels and gastropod in North America (Jacquemin et al., 2016). The Animalia living inside the watershed are well defined, however, little to nothing is known about the microbial life.

The Tennessee Valley region experiences major rainfall throughout the year. Hurricanes and tropical storms travel from the Gulf of Mexico across the Southeast, releasing bound moisture. The region averages 51 inches of rain each year, 23 inches, or 44%, of the rainfall ends up as runoff in the Tennessee River (TVA, <https://www.tva.com/environment/managing-the-river/flood-management/valley-rainfall>). In addition, during the later summer months, Tennessee is prone to droughts, resulting in unpredictable swings between the abundance and the scarcity of rain. This causes the region to cycle between flooding and drought (TVA, <https://www.tva.com/environment/managing-the-river>). As cities are built closer to the water, they are prone to flooding and erosion. In an effort to prevent flooding, as well as enable transportation across the river and produce electricity, the Tennessee Valley Authority (TVA) was founded in 1933. Several dams were constructed across the region. The construction of these dams resulted in the formation of new reservoirs. Although the TVA has conducted thorough post-impoundment surveys of the resulting new reservoirs, the microbial community has not been a focus. Today, the TVA scores the health of each of its reservoirs biyearly. One of these scores is a chlorophyll rating ranging from poor to good (TVA,

<https://www.tva.com/environment/environmental-stewardship/water-quality/reservoir-health-ratings/tellico-reservoir>). The method of chlorophyll collection and interpretation of the abundance to the resulting score is not publicly available.

The Tennessee River Basin is the largest tributary of the Ohio River. Shortly after flowing into the Ohio river, the water of the Tennessee River flows into the Mississippi and eventually into the Gulf of Mexico (Figure 1). The Tennessee River Basin has a drainage area of $\sim 105,950 \text{ km}^2$ (<https://streamstats.usgs.gov/ss/>). The watershed includes the states of Alabama, Georgia, Kentucky, Mississippi, North Carolina, Tennessee and Virginia. The Tennessee River flows for $\sim 1049 \text{ km}$ and has its origins in West Virginia, before flowing through East Tennessee, North Alabama, then Middle and West Tennessee before meeting the Ohio river in West Kentucky. Following the damming of the Tennessee river, several reservoirs and lakes were created. Some of which are Kentucky Reservoir, Tellico Reservoir, Norris Lake, Guntersville Lake, Wheeler Lake, Pickwick Lake, Douglas Lake, Cherokee Lake, Nottely Lake, and Blue Ridge Lake.

Post-impoundment, the reservoirs attracted many people due to recreational and economic opportunities. The increase in anthropogenic activity around several reservoirs of the Tennessee River Basin leads to the potential for eutrophication. Moreover, climate change is projected to enhance freshwater scarcity and may make conditions more favorable for bloom forming species (Nazari-Sharabian et al., 2018; O'Neil et al., 2012). With an imminent change in environmental conditions, it is important to record a baseline of the reservoir community prior to potential changes.

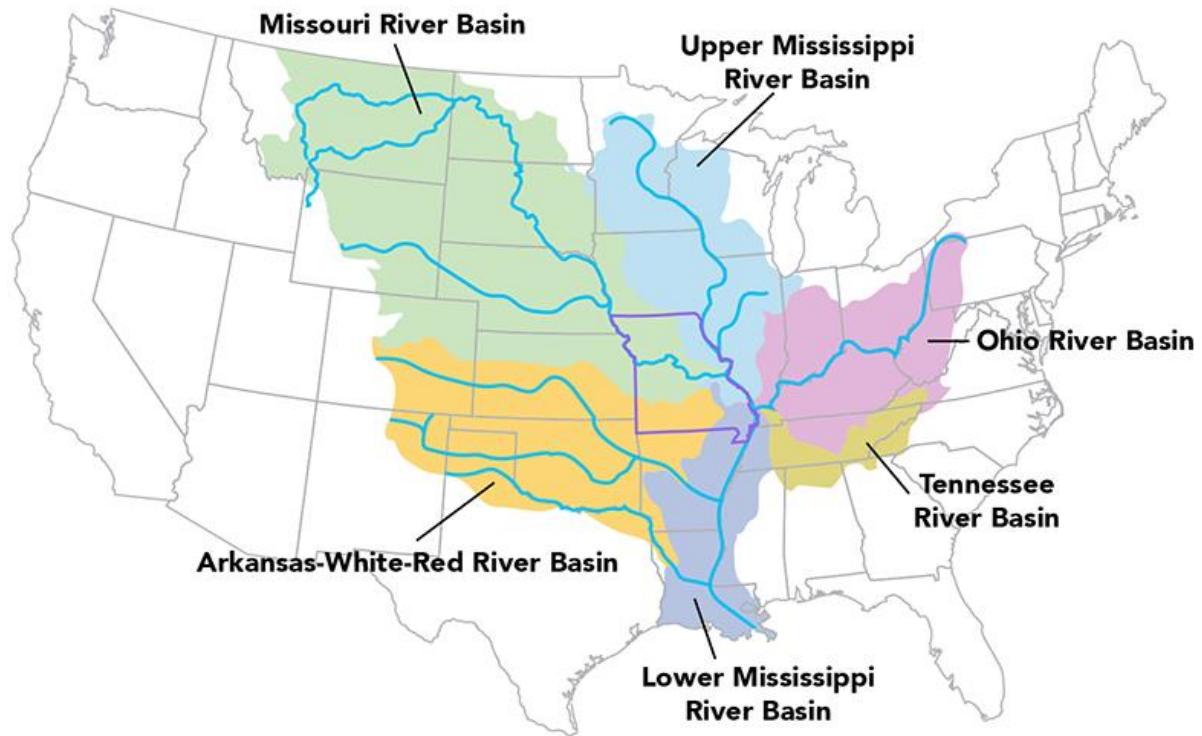


Figure 1: Map of the Mississippi river watersheds. Highlighted is the Tennessee River Basin.
(<https://dnr.mo.gov/water/how-water/state/interstate>)

History of Tellico reservoir

The Tellico Dam was completed in 1979 by the Tennessee Valley Authority (TVA). The resulting Tellico Reservoir provides ~575 km of shoreline and ~63 km² of water surface (TVA, <https://www.tva.com/about-tva/our-history/built-for-the-people/telling-the-story-of-tellico-it-s-complicated>). Together with the Fort Loudon Dam, the reservoirs help regulate flooding, enable transportation, and generate electricity. Tellico Reservoir was part of a construction project by TVA to use the shorelines around the reservoir for industrial, residential and recreational development. It is one of the few reservoirs which originally was planned to stimulate economic and recreational development, as opposed to flood control and hydroelectricity generation. It first drew national attention as construction of the dam resulted in a massive reshaping of local landscape / topography (Plater, 2013; Wheeler and McDonald, 1986). The once shallow and free flowing Little Tennessee River became a slow flowing, deep reservoir. As a result, native fish and mussels suffered, as over 20 aquatic species were lost from the reservoir (TVA, <https://www.tva.com/about-tva/our-history/built-for-the-people/telling-the-story-of-tellico-it-s-complicated>).

Another reason for controversy was the acquisition of ~89 km² of private property *via* eminent domain law (TVA, <https://www.tva.com/about-tva/our-history/built-for-the-people/telling-the-story-of-tellico-it-s-complicated>). Hundreds of people, mainly farmers, were displaced for recreational and industrial development (Rawls Jr., <https://www.nytimes.com/1979/11/11/archives/forgotten-people-of-the-tellico-dam-fight-right-of-eminant-domain.html>, 1979). In addition, the creation of the dam flooded several Cherokee archeological sites, including the city of Chota and Tanasi (the city for which the state of Tennessee is said to be named)(Chapman, 1985; Indian Affairs, <https://www.bia.gov/as->

ia/opa/online-press-release/origin-names-us-states#,1974). The proposed Tellico project was only partially achieved, as the dam never brought the economic development to the area that was promised. Despite protests and appeals, the construction of the dam moved forward until 1973, when an endangered fish species (*Percina tanasi*, the Snail Darter) was discovered to swim near the construction site (Hickman and Fitz, 1978; Starnes, 1977). The fish population was distributed to nearby bodies of water and construction resumed until 1977, when the court ruled in favor of the fish, even though the dam was 90% complete (Davis, 1979; TVA v. Hill, 1978). The Supreme Court upheld the lower court's decision in 1978, spearheading Tennessee Sen. Howard Baker push to make an amendment to the Endangered Species Act. His goal was to create a committee that can decide whether the projects were more valuable than the endangered species in question. Due to its power over other life, the committee was highly controversial, nicknamed "God Committee" (Gilmer, 2013).

In January 1979, the committee ruled against the snail darter and in favor of the Tellico Project, allowing construction of the Tellico Dam to move forward. Today, the industry and local businesses around the reservoir provide 1,800 jobs and \$137 million in capital investments around the reservoir. Notably the Snail Darter has been sighted again and is now thriving inside the reservoir (TVA, <https://www.tva.com/about-tva/our-history/built-for-the-people/telling-the-story-of-tellico-it-s-complicated>). It was taken off the endangered species list in Oct. 4, 2022 (Shollenberger et al., 2023).

Today, some 5 million people get their water from the Tennessee River and its tributaries every day (TVA, <https://www.tva.com/environment/managing-the-river>). Over 8 billion gallons of water are used daily, drawn from more than 700 municipal, industrial, agricultural, and thermoelectric intakes. However, over 95 percent of that water is recycled and returned to the river

to be used again (TVA, <https://www.tva.com/environment/managing-the-river>). In addition, development around the reservoir dramatically increased, as it became a popular retirement community with the main town known as Tellico Village. It is home to 7,311 residents, all along the western shore of the reservoir, while the eastern shoreline largely remains undeveloped (U.S. Census Bureau, <https://www.census.gov/quickfacts/fact/table/tellicovillagecdptennessee,n.d.>). Alongside countless homes, three golf courses are in the heart of the area, directly bordering the water.

Microbial diversity in the reservoir

Prior to impoundment in 1979, the Little Tennessee Rivers used to be a prominent fishing spot for both brown and rainbow trout. However, no information was recorded about the microbial community. Similarly, other rivers were also not characterized before damming. In 1939, however, streams in the watershed of two Tennessee rivers were investigated. The streams are from the Cumberland and Duck River. The Duck River is in the Tennessee River watershed, while the Cumberland River is in the Ohio River watershed. This assessment was conducted on streams, thus not indicating the microbial composition of the rivers. In addition, samples were centrifuged and identified *via* microscopy, which can lead to issues with identification, specificity and quantification. In this study they found, however, that the most abundant and diverse algal species were Chlorophyceae (74 species), Chlorococcales (50 species) and Volvocales (24 species). The most abundant taxon was *Pandorina* in the Cumberland River and *Pteromonas* in the Duck River. The most commonly found diatom taxon was *Navicula*. The most abundant diatom taxon was *Nitzschia* in the Cumberland and *Navicula* in the Duck River. *Cyclotella* sp. were also abundant, indicating an oligotrophic environment (Ossyssek et al.,

2020). Several *Synedra* sp. were commonly found. *Euglenia pisciformis* was not commonly found, but when found it was abundant in high numbers. Cyanobacteria were present, though in very low numbers. *Merismopedia tenuissima* was the most abundant cyanobacterium in both rivers. The bloom forming species *Oscillatoria* sp. was also present, while *Anabaena* spp. and *Microcystis aeruginosa* were also present in the Cumberland River, but not in the Duck River (Lackey, 1942).

After the construction of the Tellico dam concluded, a post-impoundment survey was conducted. There was a difference between community structure in the lower section (close to Tellico and Fort Loudon Dam), middle section and upper sections (inflow from Tellico river and Chilhowee Dam) of the reservoir. Generally, the lower section was dominated by blue-green algae (now known as Cyanobacteria), especially *Anacystis* (now known as *Synechococcus*), *Oscillatoria*, and in the summer months *Anabaena* (Sagona, <https://www.osti.gov/biblio/6124151>, 1983). The middle section was first dominated by the diatoms *Asterionella* or *Tabellaria*, but over the summer months the Cyanobacterium *Oscillatoria*, *Anabaena* and *Anacystis* spread upstream. The green algae *Pyramimonas* and *Scenedesmus* could also be found during the months of April and June respectively. The upper inflow section of the reservoir was generally dominated by the diatom *Tabellaria*. Other diatom species found were *Asterionella*, *Synedra* and *Meliorisa* in February, June and August. *Dinobyron* was found in February and August, as well as *Chlorococcum* and *Glenodinium* in July (Sagona, <https://www.osti.gov/biblio/6124151>, 1983). The distinct difference in community composition is hypothesized to be accredited to the difference in flows inside the reservoir. The lower section of the reservoir can receive a flow reversal from the Fort Loudon Dam due to elevation differences between the two reservoirs (Waldrop et al., 1980). Due to the flow reversal,

the lower section of the Tellico Reservoir receives inflow from the Tennessee River. The upper section of the reservoir receives inflow from the Chilhowee Dam. The water released from the dam is colder than the water inside the Tellico Reservoir, equating to density differences, which result in the released water to sink below warmer surface waters, further increasing stratification. During the colder months, where the water temperature is comparable between the two reservoirs, the inflow from the Chilhowee dam does not sink and the water is vertically mixed (Waldrop et al., 1980). Another thing to note is the presence of vegetation during impoundment. Although the reservoir was cleared from vegetation during the 1970s, the delay from the legal battle resulted in re-establishment of vegetative cover. When the dam was closed in 1979, vegetation was present. Decomposition of organic matter in the lower depths of the reservoir can affect algal composition close to impoundment (Sagona, <https://www.osti.gov/biblio/6124151>, 1983).

Not only algal species were impacted by impoundment. The TVA reported over 20 aquatic species no longer existing in the reservoir. This number may be underestimated, as mussel species alone were devastated. Prior to impoundment, 50 unionid species were reported to inhabit the Little Tennessee River. 12 years post impoundment, only six species were found present in the lake, while eight others invaded the watershed (Parmalee and Hughes, 1993).

TVA surveys

The TVA scores the ecological health of its reservoirs on a scale of 0 to 100. The scores are a cumulation of dissolved oxygen, chlorophyll, fish, bottom life and sediment ratings. The health score ratings are divided into “poor”, “fair” and “good” which results from a score of <59, 59-72, and >72, respectively. Tellico Reservoir gets scored every two years. Tellico Reservoir

scored consistently between 50 and 60, resulting in “poor” to “fair” scores. The highest score (69) was reported in 1994, while the lowest score (51) was reported in each of 1995, 2001, and 2011. In 2021, Tellico Reservoir got a health score of 64, its second highest ever. This is a six point increase from 2019, where it scored 56. In this reservoir, the forebay and mid-reservoir are scored separately. Both forebay and the mid-reservoir had chlorophyll ratings of “fair”.

According to the TVA, higher chlorophyll concentrations are found at the forebay, due to the input of nutrient rich water from the Fort Loudon reservoir. The upstream soils are reported to be nutrient poor, yet higher than expected chlorophyll concentrations are found at mid-reservoir (TVA, <https://www.tva.com/environment/environmental-stewardship/water-quality/reservoir-health-ratings/tellico-reservoir>). Chlorophyll ratings at the forebay are usually “poor” and “poor” to “fair” at the mid-reservoir. The TVA does note that these ratings are relative and compared to other reservoirs in the same ecoregion. Tellico Reservoir is in the Blue-Ridge ecoregion and has higher chlorophyll ratings than other reservoirs in this region (Lacher et al., 2019). If you compare Tellico Reservoir to other geographically close reservoirs outside of the Blue-ridge ecoregion, like Fort Loudon Reservoir, it would score a lot better (TVA, <https://www.tva.com/environment/environmental-stewardship/water-quality/reservoir-health-ratings/tellico-reservoir>).

Changes in the reservoir

The Tennessee River Region is projected to add 650,000 residents to the already existing 5.2 million by 2045 (TVA, [https://www.tva.com/environment/managing-the-river/managing-water-supply#:~:text=The%20Tennessee%20River%20Watershed%20will,the%20TVA%20region's%](https://www.tva.com/environment/managing-the-river/managing-water-supply#:~:text=The%20Tennessee%20River%20Watershed%20will,the%20TVA%20region's%20water%20supply)

20water%20resources.). Additionally, growth in urban areas around the region—some of which are already facing water-supply challenges—will put increased pressure on the TVA region’s water resources (TVA, <https://www.tva.com/environment/managing-the-river/managing-water-supply#:~:text=The%20Tennessee%20River%20Watershed%20will,the%20TVA%20region's%20water%20resources.>). In addition, development around the reservoir is steadily increasing. The three counties in which Tellico Reservoir is located (Blount county, Monroe county and Loudon county) all experienced an increase of development over the past 20 years. In Loudon County, high intensity, medium intensity and low intensity development in Loudon county have increased by 50%, 68% and 16%, respectively. Overall, the developed area has increased by 18% (<https://www.mrlc.gov/eva/>). This trend is also found in the other counties. Blount county had a 15% net increase in development, while Monroe county had a 12% increase. Both counties had significant increases in high and especially medium intensity development. Increase in anthropogenic activity results in an increase in nutrient pollution.

In addition to direct impacts on the reservoir due to an increase of development, a changing climate also will affect the reservoir. Due to anthropogenic input of carbon into the atmosphere, global temperatures are increasing. This affects freshwaters, as surface temperatures are increasing. This, in turn, impacts the microbial community and water quality. Increased surface temperatures proliferate growth of algal biomass. Especially bloom forming species have been found to thrive in warmer temperatures (Paerl and Huisman, <https://www.science.org/doi/abs/10.1126/science.1155398>,2008). The combined effects of anthropogenic pollution and climate change point to a deterioration of water quality. This is especially prevalent in Tellico Reservoir, as some urban areas already face water scarcity (TVA, <https://www.tva.com/environment/managing-the-river/managing-water->

supply#:~:text=The%20Tennessee%20River%20Watershed%20will,the%20TVA%20region's%20water%20resources.).

Synechococcus

The genus *Synechococcus* is a ubiquitous picocyanobacterium. It holds a critical role in both marine and freshwater, as it is among the most abundant autotrophic prokaryotes (Fucich and Chen, 2020). Its size is estimated to be 0.2 μm – 2 μm (Carrick and Schelske, 1997). This makes it challenging to get accurate estimates of primary productivity. However, nearly 20% of oceanic primary productivity is attributed to *Synechococcus*, making its role not negligible (Guidi et al., 2016; Jardillier et al., 2010). Picocyanobacteria play a vital role in freshwater too, as up to 30% of primary productivity can be attributed to them (Carrick and Schelske, 1997). In addition, *Synechococcus* also gets preyed on by grazers, transferring material and energy to higher trophic levels and shaping structure and diversity of the ecosystem (Wang et al., 2022). On average, the genome size of *Synechococcus* is larger (2.9 Mb) than other picocyanobacteria (Partensky et al., 1999). This allows it to have a wider variety of environmental response genes and exist in wider ranges of habitats (Pittera et al., 2018). *Synechococcus* has been isolated from freshwater, estuaries, coastal and oceanic water (Fucich and Chen, 2020). A larger genome is also a predictor of adaptability, as open ocean *Synechococcus* have smaller genomes than coastal ones (Palenik et al., 2006). Due to their large adaptability within their relatively large genome, in addition to their ubiquitous nature, *Synechococcus* contain highly diverse phylotypes and ecotypes (Fucich and Chen, 2020). This is especially true in freshwater, where the genus is phylogenetically poorly defined compared to the marine counterpart (Sanchez-Baracaldo et al., 2005). This, however, does not undermine the importance for the freshwater ecosystem.

Fahnenstiel et al. have shown that single *Synechococcus* cells accounted for an average of 10%, and up to 26%, of surface mixed-layer primary productivity in Lake Huron and Lake Michigan (Fahnenstiel et al., 1991).

Chapter II

Identifying the microbial community of Tellico Reservoir in 2023 during summer pool.

Author contributions:

David Niknejad – writing – original draft/editing/reviewing, sample collection, formal analysis

Robbie Martin – sample collection

Steven Wilhelm – writing – review/editing, conceptualization

Introduction

Tellico Reservoir is one of several reservoirs created by the TVA. The TVA samples its reservoirs every two years and assigns ecological health ratings based on five parameters. Notably, no nutrient data or microbial composition data is measured or included. The only record available of the microbial community of Tellico Reservoir comes from a post-impoundment survey conducted to describe the newly formed reservoir. In those surveys, the microbial community is described via light-microscopy. No microbial data via molecular sequencing is available describing the reservoirs. Afterwards, ecological health scores were given to the reservoirs every two years. Besides chlorophyll *a*, none of the parameters measured take the microbial community into account. With the advancements and reduction of costs of molecular microbial analyses, it is important to monitor the microbial community with molecular tools.

With climate change increasing the scarcity of freshwater sources, it is important to establish a molecular baseline to track changes in the microbial community (Gleick and Cooley, 2021). Using molecular tools to identify the current microbial community of Tellico Reservoir will act as a baseline for future studies in the reservoir, as well as other reservoirs in the southeastern United States.

Methods

Sample sites and collection

Six stations were chosen in the Tellico Reservoir based on presumed influence of land use on the watershed. Two stations with presumed influence from agricultural land use, two stations with presumed influence from urbanization and two stations with no presumed influence, one being at the upstream of development around the watershed (Table 1, Figure 2). Stations were sampled every three weeks starting on April 20, 2023 and continuing until October 24, 2023. Stations were sampled in the following order, due to ease of sampling: Ag1, D2, C1, D1, Ag2, C2. For downstream analysis, stations were combined by land use, resulting in agriculture (Ag1+Ag2), developed (D1+D2), and control (C1+C2). The control was chosen so the adjacent and watershed land use was of forest or undeveloped. To determine seasonal effects, samples were grouped into three periods: spring (sampling dates 4/20/23, 5/11/23, 6/1/23, 6/22/23), summer (7/13/23, 8/3/23, 8/24/23, 9/14/23) and fall (10/3/23, 10/24/23).

Sample collection

A bucket was rinsed with water from each station. It was then filled with one or two loads of water collected by a 4 L Van Dorn bottle. Lake water for FlowCam analyses (below) was collected in 50 mL Falcon tubes and preserved in pre-added ~1% glutaraldehyde (EM grade). For nucleic acids, lake water was pushed through a 0.2- μ m polycarbonate Sterivex™. The filtrate was collected in 50 mL Falcon tubes for nutrient analysis and enumerated for filtrate collected. Lake water was also collected in 300 mL Nalgene bottles for chlorophyll *a* quantification. All samples were immediately stored in an ice chest on the vessel. Sterivex™

were plugged with a female Luer lock as well as clay and immediately placed in a -80° C liquid nitrogen tank.

Environmental metadata measurements

A YSI EXO2 (*Xylem*) multiparameter sonde (SKU:599502-01) was used with temperature, conductivity, dissolved oxygen, ammonium and turbidity sensors installed. Sensors were calibrated per exo user manual (*Xylem*, <https://www.ysi.com/File%20Library/Documents/Manuals/EXO-User-Manual-Web.pdf>,2020). Data on water column parameters were collected at 1 m, 2 m, 3 m for each station. Secchi depth was taken by lowering of a Secchi disk into the water until the difference between black and white was not determinable. It was performed by the same individual for all runs but one (9/14/23).

Measuring extracted chlorophyll *a*

Lake water (100 mL) was filtered onto a 0.2- μ m, 47-mm diameter pore-size polycarbonate filter. The filter was placed into a 15 mL falcon tube and stored at -20° C until processing.

Chlorophyll *a* concentrations were estimated on a Turner Designs AU-10 fluorometer calibrated per manufacturers specifications (<http://docs.turnerdesigns.com/t2/doc/instructions/S-0148.pdf>). Chlorophyll *a* standards for AU-10 calibration were extracted from spinach leaves and quantified on ThermoSpectronic Biomate 5 spectrometer *via* an adapted protocol from Wetzel and Likens, 2000 (Wetzel and Likens, 2000; Wilhelm, <https://www.protocols.io/view/total-chlorophyll-a-measurements-by-spectrophotome-q26g73yqv wz1/v1>,2017).

Table 1: List of stations names, assigned land use category, land use per NLCD, and location.

Station name	Land use category	Land use per NLCD	Longitude (°)	Latitude (°)
Ag1	Agriculture	Agriculture	35.678	-84.280
D2	Developed	Developed	35.670	-84.247
C1	Control	Forest	35.662	-84.240
D1	Developed	Developed	35.658	-84.266
Ag2	Agriculture	Agriculture	35.641	-84.217
C2	Control	Forest	35.578	-84.184

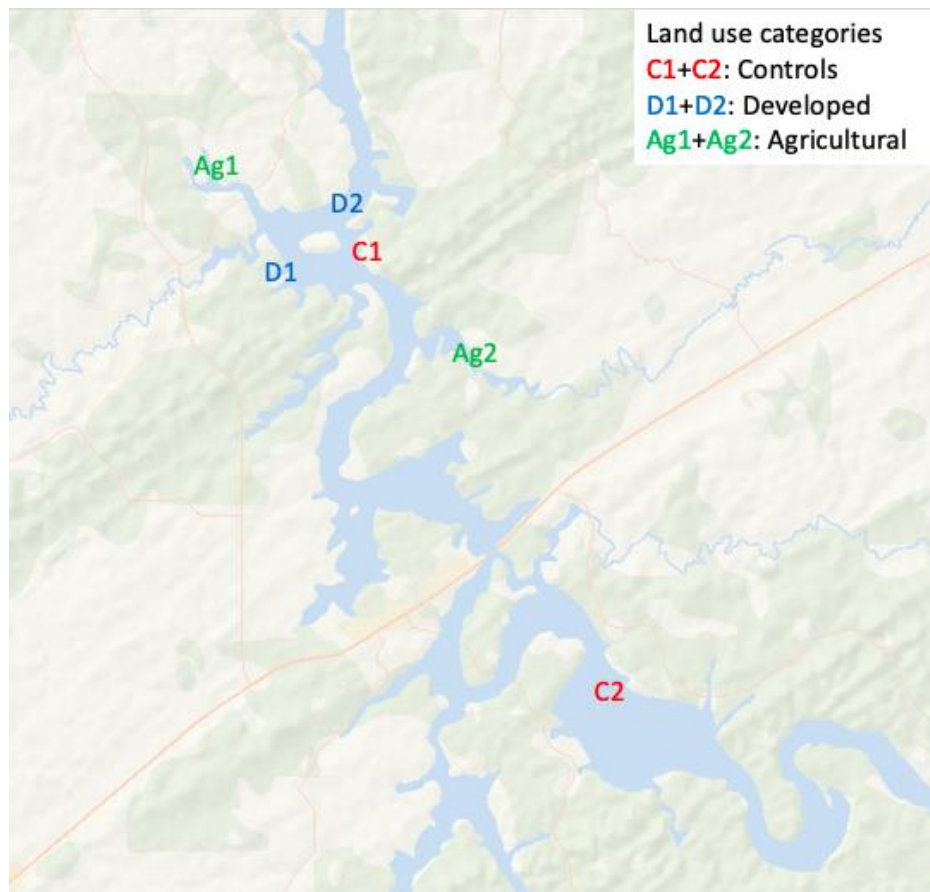


Figure 2: Sampling locations paired with presumed land use.

Polycarbonate filters containing biomass were placed inside a 15 mL Falcon tube with 5 mL of 90% acetone and stored at -20° C overnight. The following day, the acetone was decanted into a 13 x 100 mm glass tube. Chlorophyll *a* concentration was estimated in duplicates on the pre-calibrated AU-10 fluorometer.

Final chlorophyll *a* concentrations were calculated in the following fashion:

$$\text{Chlorophyll in Sample } \left(\frac{\mu\text{g}}{\text{L}} \right) \times \text{Acetone Volume} = \text{Total Chlorophyll extracted } (\mu\text{g})$$

$$\frac{\text{Total Chlorophyll extracted } (\mu\text{g})}{\text{Volume Filtered (L)}}$$

$$= \text{Concentration of Chlorophyll in original Sample } \left(\frac{\mu\text{g}}{\text{L}} \right)$$

DNA extraction and PCR amplification of gene sequences

Prior to extraction Sterivex™ were removed from the freezer and thawed at room temperature. The polycarbonate filter was cut loose from the Sterivex™ with a sterile scalpel and tweezers, then cut up into six pieces on a petri dish. The pieces of the filter were placed in a 2 mL bead beating tube prefilled with extraction buffer. DNA was extracted with solution A, phenol, and chloroform. DNA was precipitated using 100% and 70% ethanol and resuspended in 50µL of water.

For PCR amplification, a master mix containing 2.5µL of EconoTaq 2x plus green polymerase (*Biosearch Technologies*), 19 µL of PCR grade water, 2.5 µL of 10µM forward primer and 2.5 µL of 10 µM reverse primer. To each reaction, 1 µL of template was added. The forward primer (515F-Y) and the reserve primer (926R) were concatenated with an adapter

sequence necessary for amplicon sequencing resulting in the following primers: Forward 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTAA and reverse 5'-

GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCGYCAATTYMTTTRAGTTT

(Parada et al., 2015). Thermal cycling was conducted on a pre-heated Bio-Rad C1000 Touch™ thermal cycler and was adapted for environmental samples. The initial denature temperature was set to 95° C for 10 min. Thirty cycles were performed as follows: denature temperature 95° C for 30 s, 45 s annealing at 50° C and a 1.30 min of extension at 72°C. The final extension was for 8 min at 72° C.

16S/18S rRNA gene analysis

This workflow was adapted from Dr. Michael D. Lee (https://astrobiomike.github.io/amplicon/16S_and_18S_mixed). Primers were trimmed using cutadapt, then blasted against the pr2 database (v5.0) using magic-blast. The output was split into 16S and 18S sequences for downstream analysis. Fastq files were loaded into RStudio2023.09.1 (*Posit Software, PBC*) and analyzed with the package DADA2 v1.26.0 following and modifying the tutorial (<https://benjjneb.github.io/dada2/tutorial.html>). The forward strand was trimmed at base 250 and the reverse strand was trimmed at base 225, based on the quality control profile. For reads annotated as 16S rRNA gene, the forward and reverse strand were merged with a 75 base pair overlap. Chimeras were removed. For reads annotated as 18S, the same workflow was used, though read pairs were merged with the argument `justConcatenate=TRUE`. Taxonomy was assigned to the 16S reads using `silva_nr_99_v138.1_train_set` and species was assigned using

silva_species_assignment_v138.1, while silva v132 was used to assign taxonomy to 18S reads. The extension _unclassified was added to all reads that were annotated as NA or uncultured.

FlowCAM analysis

Lake water was collected in a 50 mL Falcon™ tube and preserved with ~1% glutaraldehyde. Samples were filtered through a 100 µm mesh and placed in the fridge until further processing. Before running the samples, a context file was created. The size minimum was 15 µm. Images were collected until ~2000 events were recorded. To classify the images into different groups, VisualSpreadSheet5 was used. Filters for each group were created by grouping similar images together and comparing parameters to each other. To identify parameters that best highlight the differences between the groups, a python script was used. Once the filters were set up, auto-classification was run. Afterwards, the groups were manually curated to detect outliers.

Data analysis

Land cover was retrieved from the 2021 National Land Cover Database (<https://www.mrlc.gov/data/nlcd-2021-land-cover-conus>). Drainage areas were calculated with StreamStats (<https://streamstats.usgs.gov/ss/>). Shapefiles of drainage areas were mapped in ArcGIS Pro 3.2.0. Prior to analysis, missing values in the Secchi depth dataset were estimated using the Expectation-maximization algorithm to allow for further processing of data.

Nitrate+Nitrite, YSI chlorophyll *a*, phycocyanin RFU, and Conductivity (C) were removed from the physicochemical dataset, as they illustrated high correlations (>0.9) with other included parameters. Data were normalized and analyzed using Principle component analysis (PCA) based on Euclidean distance and plotted using Primer version 7.0.23. Physicochemical graphs were

constructed using Prism version 10.0.0. In analyses regarding molecular data, station Ag1 on April 20th was removed due to low sequence abundances. Taxonomic data were standardized by total abundance and plotted using RStudio version 2023.09.1. A non-metric Multi-Dimensional Scaling plot (nMDS) was created using the Species composition data. A SIMPER using Bray-Curtis similarity was conducted to highlight what species caused the separation between samples in ordinate space. To investigate whether the factors season and/or land use were a good predictor of the separation highlighted by the nMDS, a two way crossed PERMANOVA was performed. The interaction term was removed due to it being not significant ($p=0.205$). In addition, a BIOENV using Spearman rank correlation was used to identify parameters correlating with the separation highlighted by the nMDS.

Results

Land use patterns

The characteristics of the different watersheds of each station were analyzed (Table 2). The drainage area of stations Ag1, Ag2, D1, D2, C1, and C2 is as follows: 5.9 km², 106.87 km², 0.21 km², 6717.25 km², 0.17 km², 5423.62 km². The primary land use of agricultural stations was pasture/hay. The primary land use of developed stations was developed, yet station D2's primary land use was forest. The primary land use of control station was forest (Figure 3-8).

Variation within controls

Due to the large distance between the control stations (C1+C2), they were compared to each other. To achieve this, the physicochemical properties and the molecular community of each station was compared. To compare physicochemical indicators, a PCA was conducted

(Appendix). The microbial community was compared in several ways. First, a nMDS of Bray-Curtis distance of the microbial community of the two stations was plotted (Appendix). Then, the relative abundances of the microbial community of each station was correlated against each other.

Measurements of extracted chlorophyll *a* throughout the sampling season

Chlorophyll *a* concentrations were measured in duplicate (Fig.9). All three station types started out ~5 µg/L on April 20th, 2023 not being statistically significant from each other. Every subsequent sampling date, the agricultural station had significantly higher chlorophyll *a* concentrations than both the developed and control stations. Developed and control stations were never significantly different from each other.

In control and developed stations, chlorophyll ranged from ~2.5 to ~7.5 µg/L. Agricultural stations ranged from ~5 to ~13 µg/L. A linear increase of chlorophyll *a* over time can be observed for all three stations. Chlorophyll *a* measured at agricultural stations had a steeper increase over time than the control and developed stations. The linear increase until September 14th, where it leveled off at ~13 µg/L. On June 22nd, 2023 chlorophyll measured was an outlier in the linear increase, as it was higher than expected. The same, though less steep, linear increase in extracted chlorophyll *a* can be observed for both control and developed stations. Interestingly, values drop from April 20th to May 11th from ~5 to ~2 µg/L, before steadily increasing to ~7.5 µg/L. Parallel to the agriculture station, on June 22nd, the chlorophyll values were higher than expected by the linear increase. In both of these stations, chlorophyll does not level out towards the end of the year, rather reaching a clear maximum at the last date sampled (October 24th).

Table 2: Drainage area and percent land use of every station. Different types of forest, agriculture, and developed land use were combined.

Station	Drainage Area (km ²)	Combined Forest land use	Combined Agriculture land use	Combined Developed land use
Ag1	5.9	33%	52%	13%
Ag2	106.87	18%	58%	18%
D1	0.21	2%	0%	90%
D2	6717.25	82%	7%	8%
C1	0.17	91%	0%	1%
C2	5423.62	88%	2%	7%

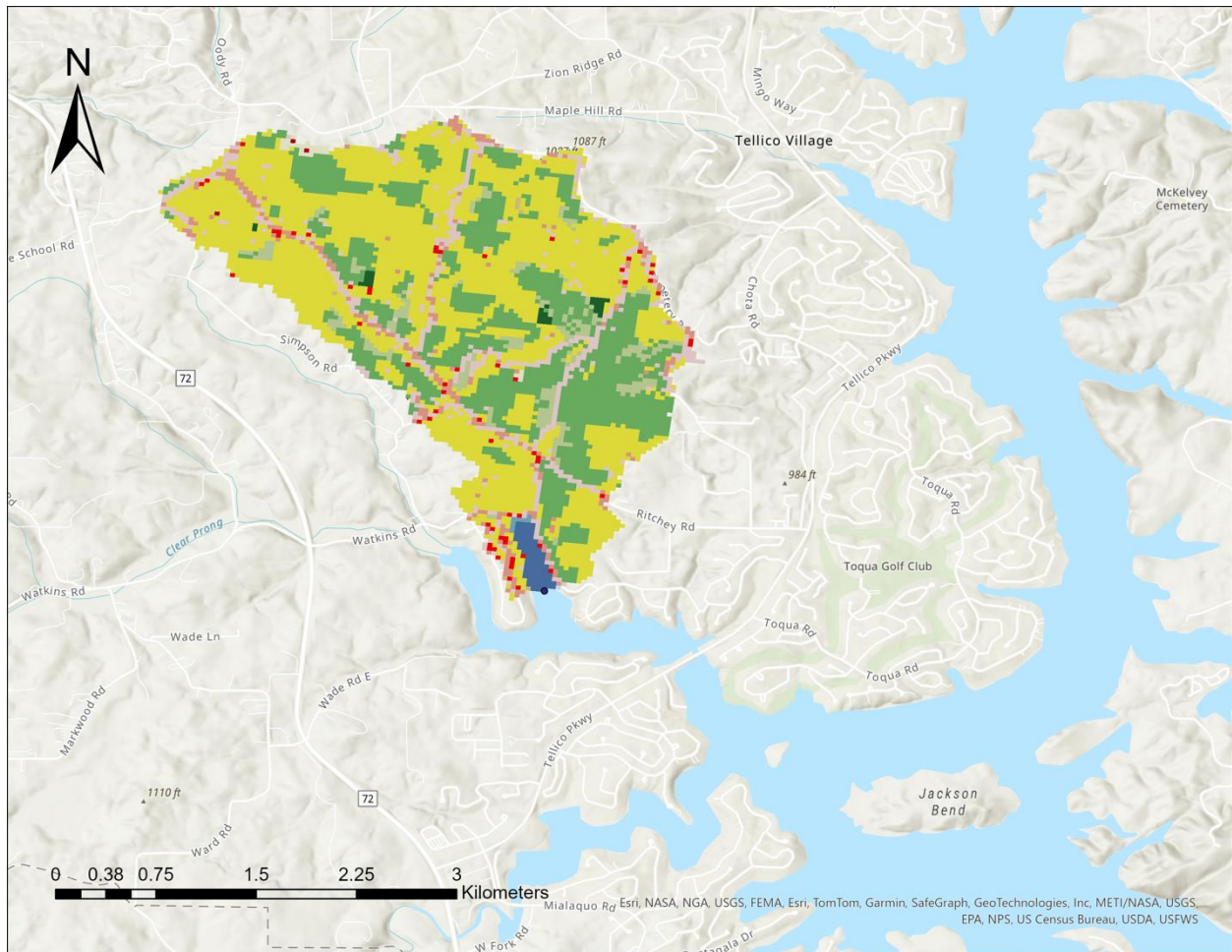


Figure 3: Drainage area and land use of station Ag1 mapped with ArcGIS Pro. Sampling point is highlighted by a dot. Drainage area was calculated with USGS's StreamStats (<https://streamstats.usgs.gov/ss/>) and land cover was determined via the 2021 National Land Cover Database. Land use is determined by the color. (Green: forest, yellow: pasture/hay, red: developed).

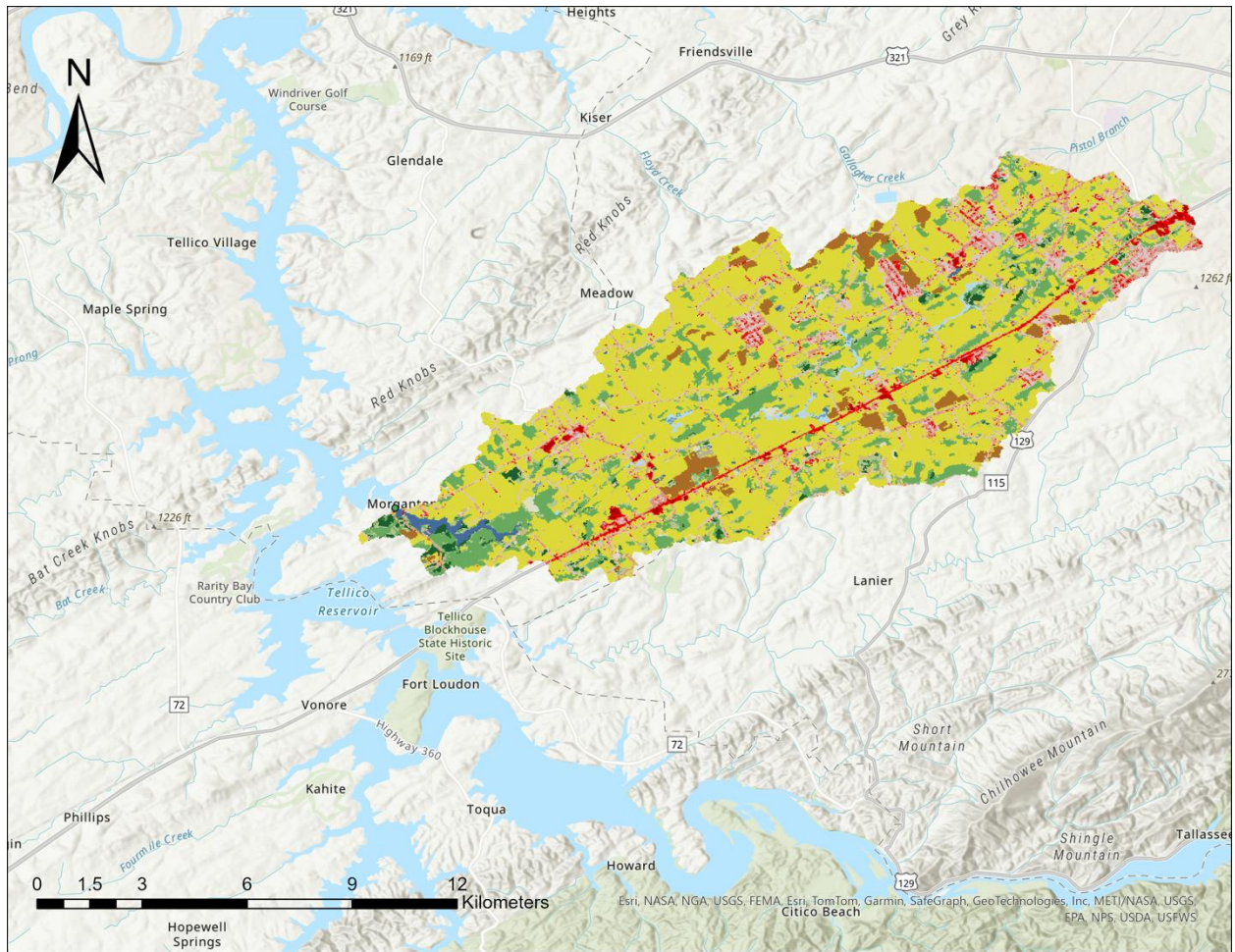


Figure 4: Drainage area and land use of station Ag2 mapped with ArcGIS Pro. Sampling point is highlighted by a dot. Drainage area was calculated with USGSs StreamStats (<https://streamstats.usgs.gov/ss/>) and land cover was determined via the 2021 National Land Cover Database. Land use is determined by the color. (Green: forest, yellow: pasture/hay, red: developed).

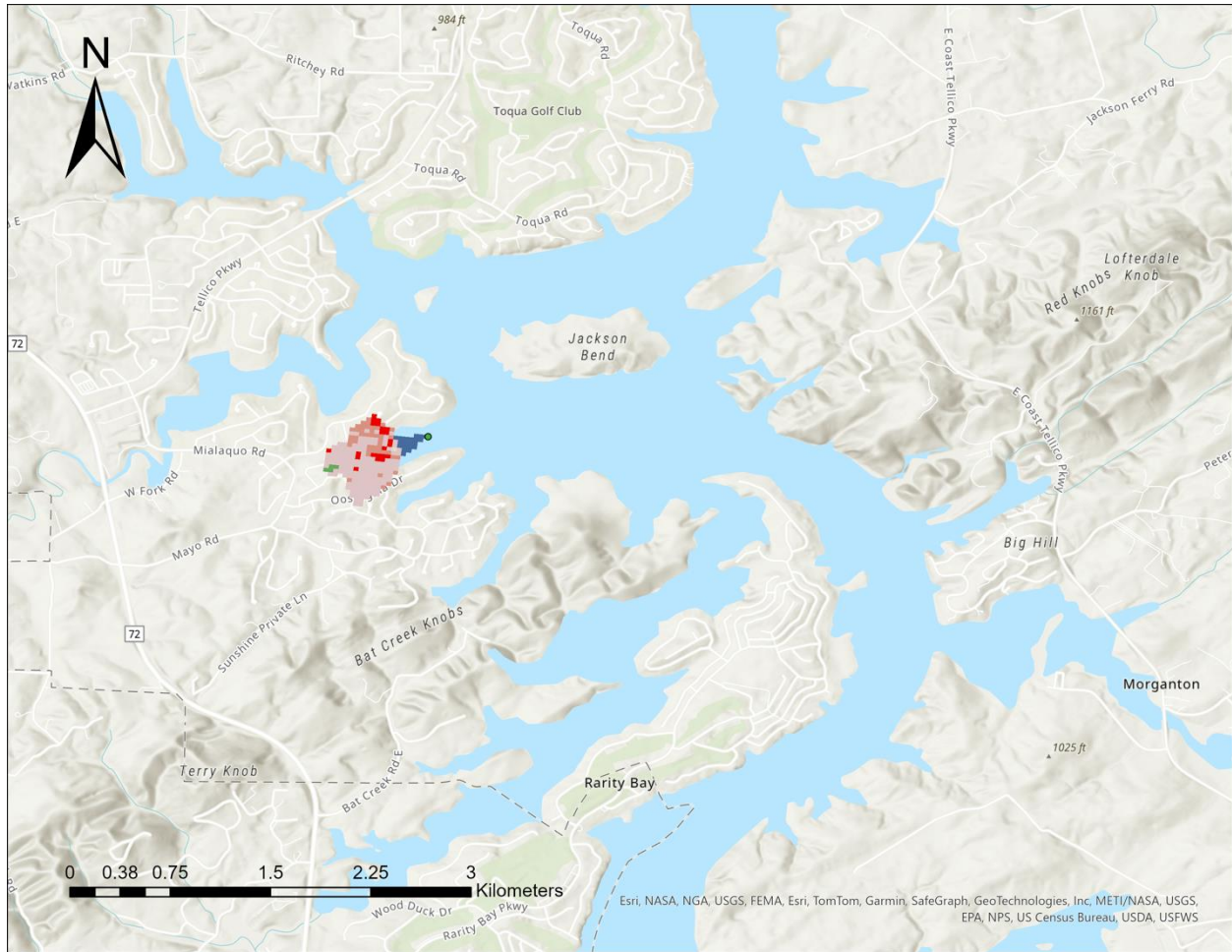


Figure 5: Drainage area and land use of station D1 mapped with ArcGIS Pro. Sampling point is highlighted by a dot. Drainage area was calculated with USGSs StreamStats (<https://streamstats.usgs.gov/ss/>) and land cover was determined via the 2021 National Land Cover Database. Land use is determined by the color. (Green: forest, yellow: pasture/hay, red: developed).

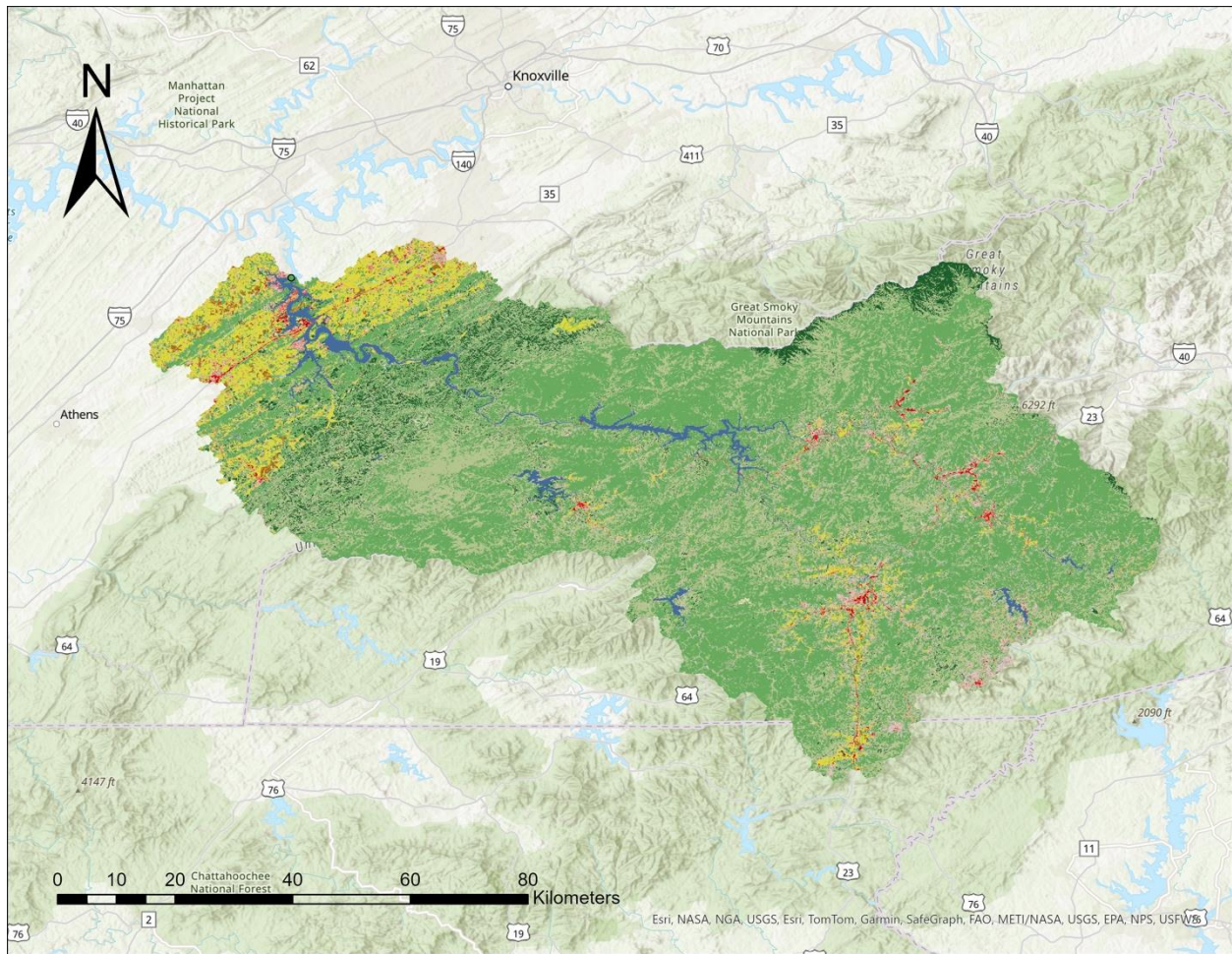


Figure 6: Drainage area and land use of station D2 mapped with ArcGIS Pro. Sampling point is highlighted by a dot. Drainage area was calculated with USGSs StreamStats (<https://streamstats.usgs.gov/ss/>) and land cover was determined via the 2021 National Land Cover Database. Land use is determined by the color. (Green: forest, yellow: pasture/hay, red: developed).

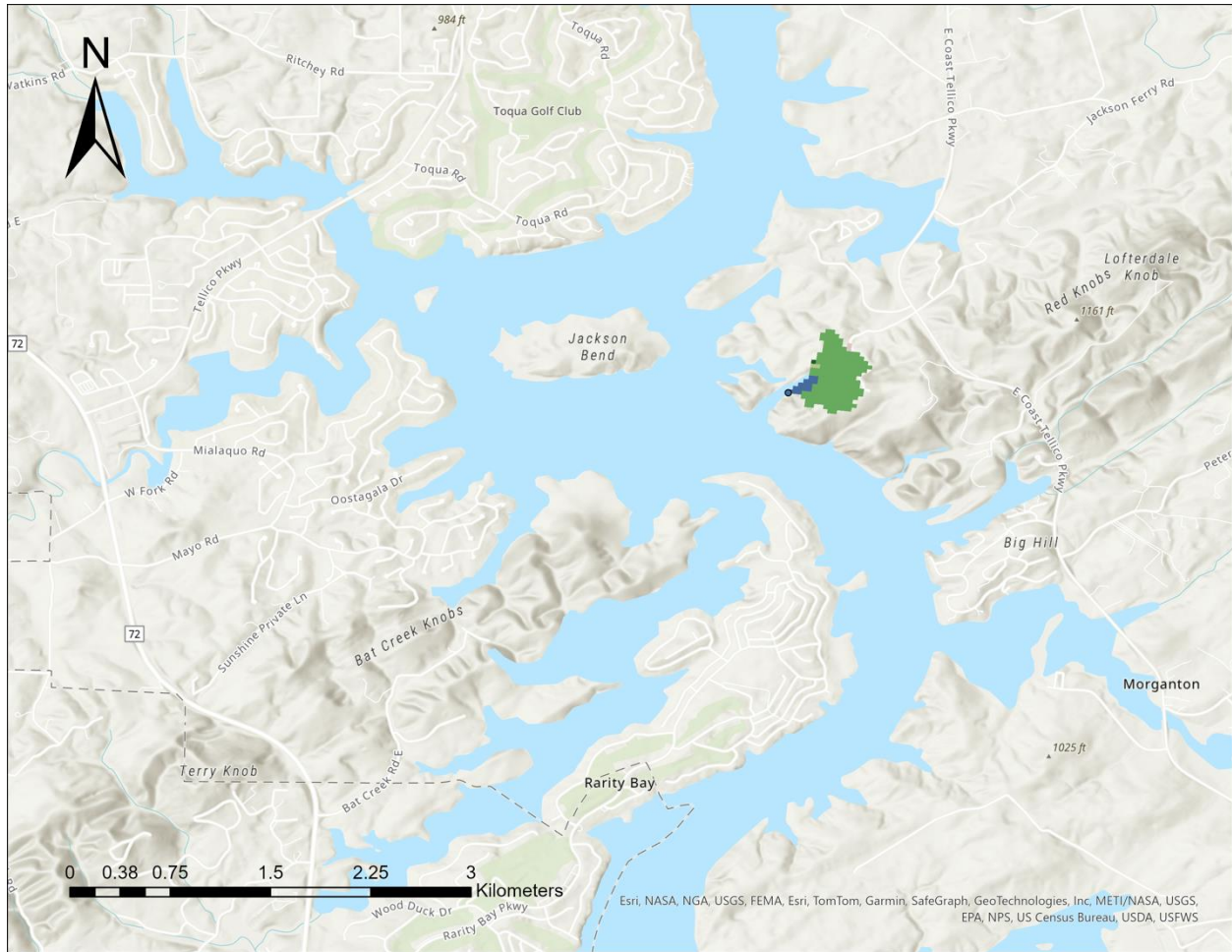


Figure 7: Drainage area and land use of station C1 mapped with ArcGIS Pro. Sampling point is highlighted by a dot. Drainage area was calculated with USGS's StreamStats (<https://streamstats.usgs.gov/ss/>) and land cover was determined via the 2021 National Land Cover Database. Land use is determined by the color. (Green: forest, yellow: pasture/hay, red: developed).

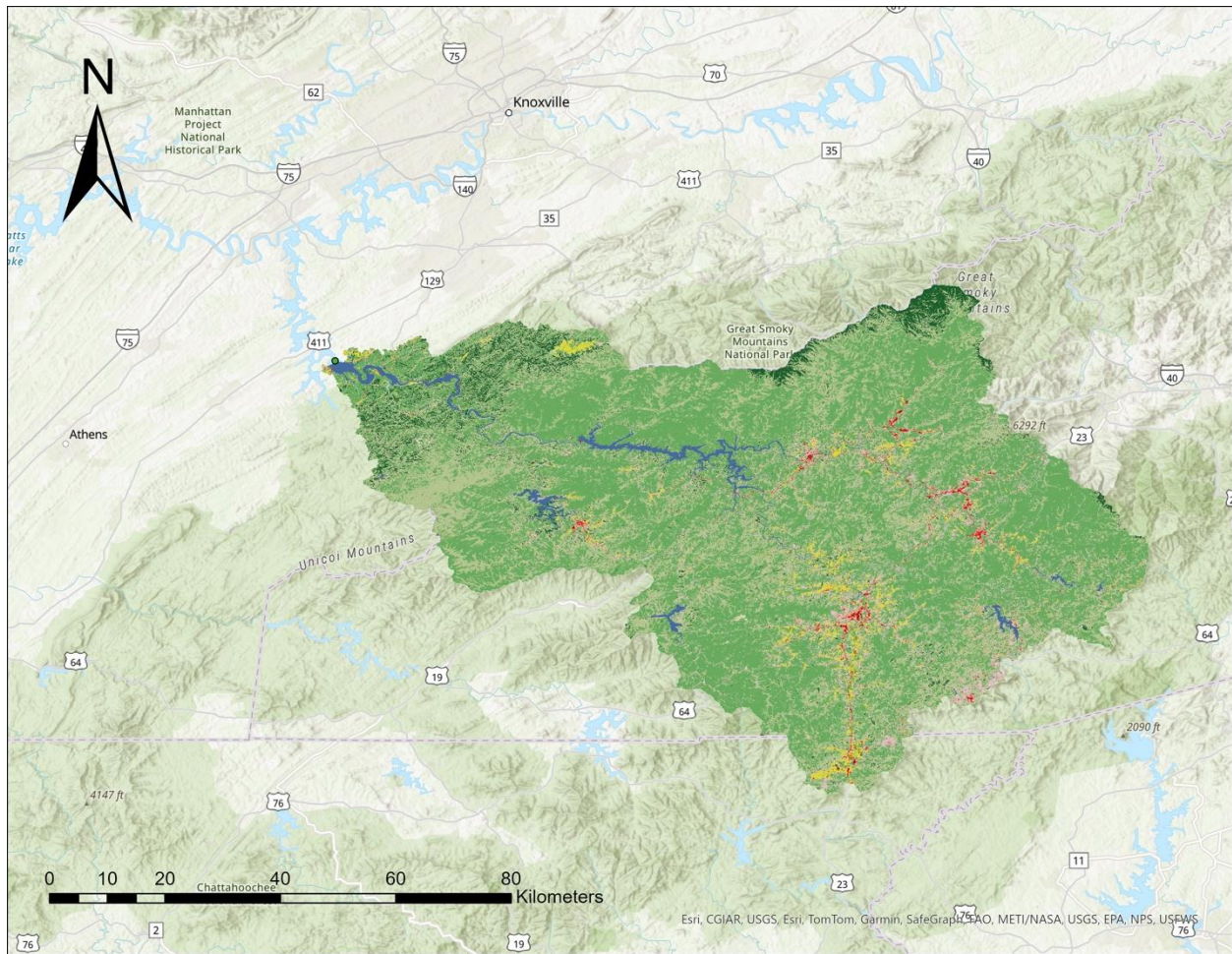


Figure 8: Drainage area and land use of station C2 mapped with ArcGIS Pro. Sampling point is highlighted by a dot. Drainage area was calculated with USGS's StreamStats (<https://streamstats.usgs.gov/ss/>) and land cover was determined via the 2021 National Land Cover Database. Land use is determined by the color. (Green: forest, yellow: pasture/hay, red: developed).

Measurement of nutrients throughout the sampling season

Nutrient samples were measured in two replicates. Dissolved nutrient concentrations were measured for ammonium, dissolved reactive phosphorus, silicate, nitrite, and nitrate (Figure 9-14). For ammonium all three stations follow a similar pattern and were never statistically significant from each other (Figure 10). Ammonium concentrations were around 0.25 to 0.6 $\mu\text{mol/L}$ in the late spring/early summer yet fell and was not detectable after August 3rd (Figure 10). Water collected at agricultural stations surged with ammonium on August 24, September 14, and October 24, yet was not significantly different from the other stations.

Similarly, stations associated with developed land use noted an increase in ammonium on September 14, but also not significant. Control stations stayed below detection limit from August 3rd onward. For dissolved reactive phosphorus (DRP) concentrations, all stations followed the same pattern and were not statistically significant from each other at any point (Figure 11). From April 20th until June 1st, as well as from September 14th until October 24th, DRP was below the detection limit. From June 22nd until August 24th, DRP was detectable and ranging from 0.1 to 0.2 $\mu\text{mol/L}$. An increase in concentration is detectable, staying constant $\sim 0.2 \mu\text{mol/L}$ on August 24th, before dropping below detection limit on September 14th. Silicate is between 50 and 100 $\mu\text{mol/L}$ throughout the whole year. Concentrations varied, yet converged to $\sim 75 \mu\text{mol/L}$. Generally Agricultural stations are higher in silicate, followed by developed and then control stations, though not supported by statistics (Figure 12). Nitrite concentrations were higher in the spring and fall ($\sim 0.15 \mu\text{mol/L}$), while dropping in the summer ($\sim 0.05 \mu\text{mol/L}$) (Figure 13). Agricultural sites measured higher nitrite concentrations on April 20th and May 11th, though not statistically significant. At all other sampling dates the three stations were measuring nitrite concentrations close to each other. Nitrate concentrations follow a similar pattern (Figure 14).

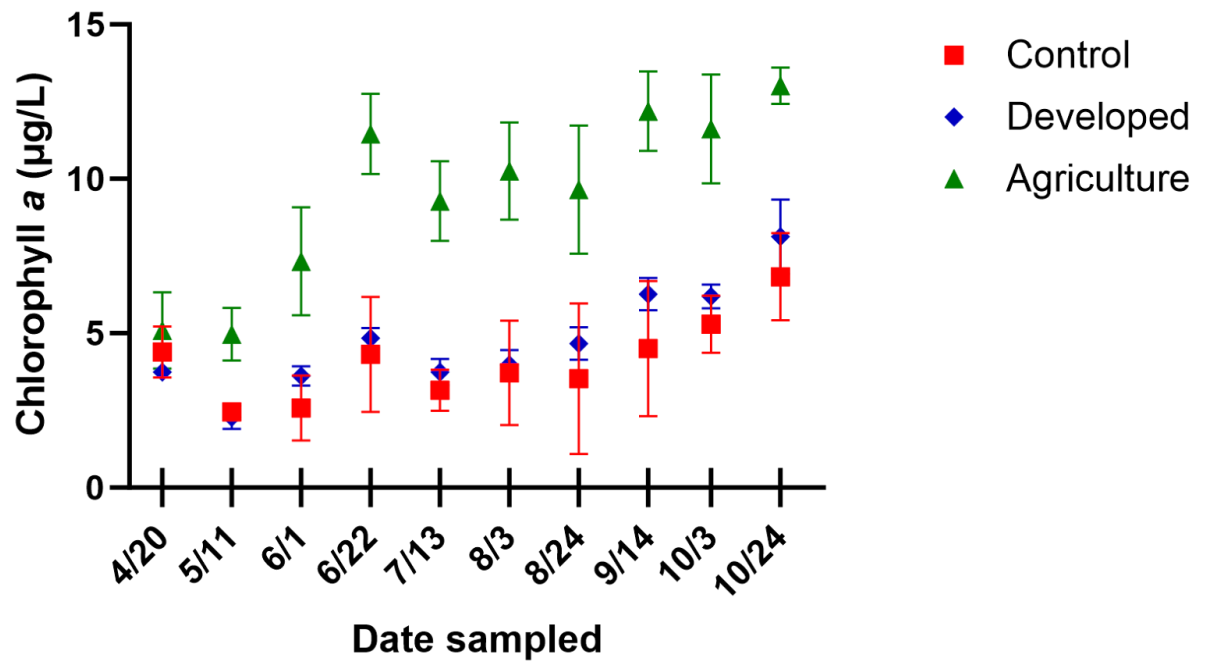


Figure 9: Chlorophyll *a* concentration averaged across the three different land use types from 4/20 until 10/24. Data points are month/day for 2023 and represent mean $\pm$ range of duplicate measures.

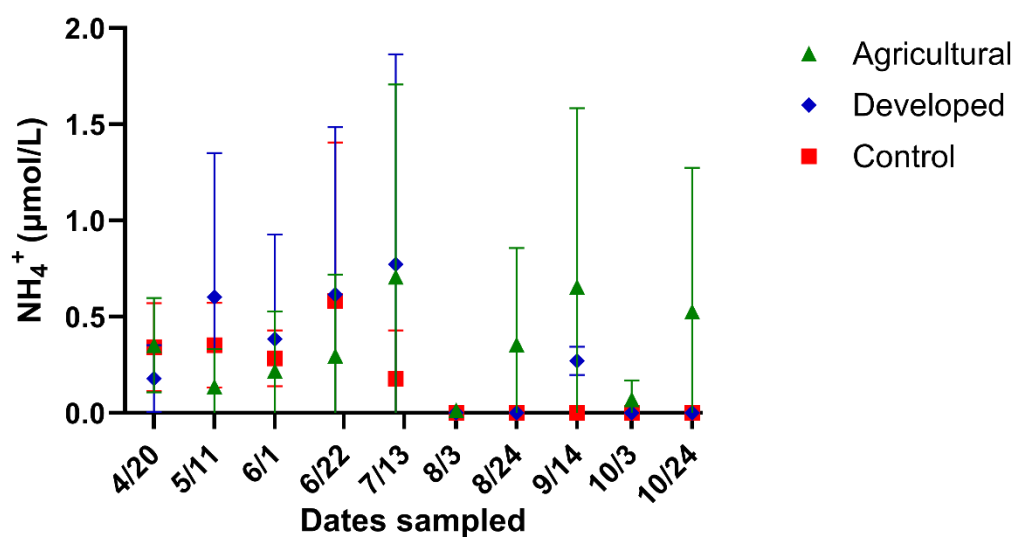


Figure 10: Dissolved ammonium concentrations in $\mu\text{mol/L}$ of land use categories through the year. Data points are means of two sampling sites for each land use type (control, developed, agricultural), error bars show standard deviation.

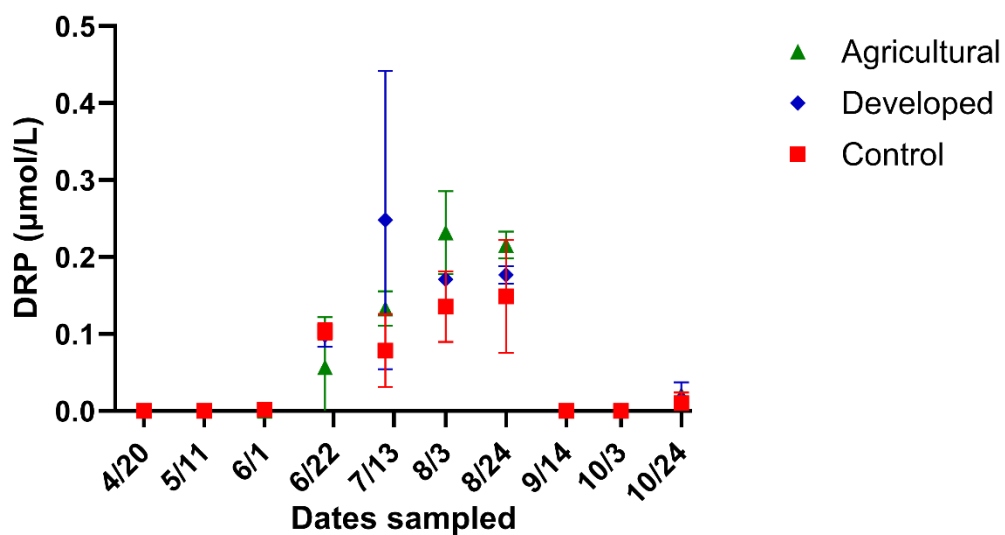


Figure 11: Dissolved reactive phosphorus (DRP) concentrations in $\mu\text{mol/L}$ of land use categories through the year. Data points are means of two sampling sites for each land use type (control, developed, agricultural), error bars show standard deviation.

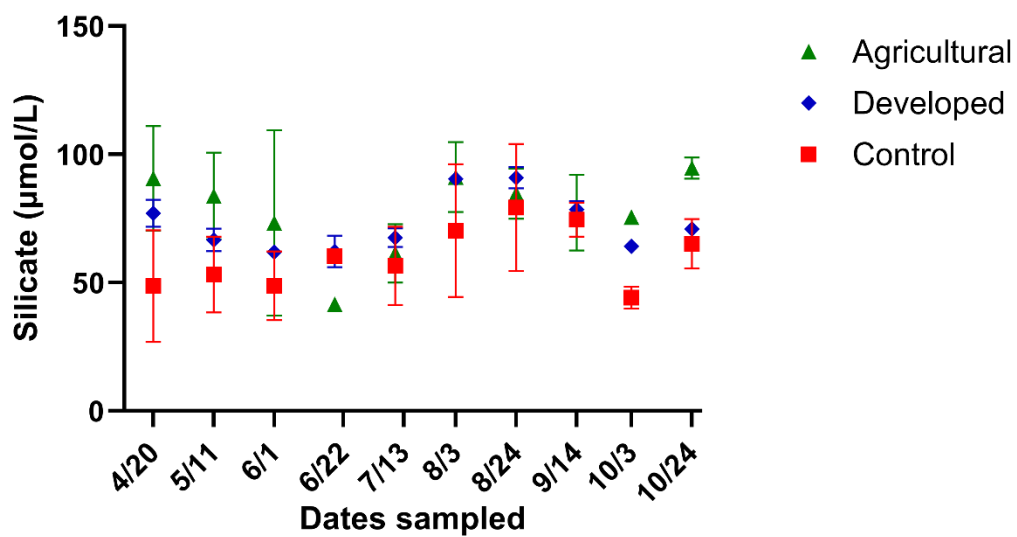


Figure 12: Dissolved silicate concentrations in µmol/L of land use categories through the year. Data points are means of two sampling sites for each land use type (control, developed, agricultural), error bars show standard deviation.

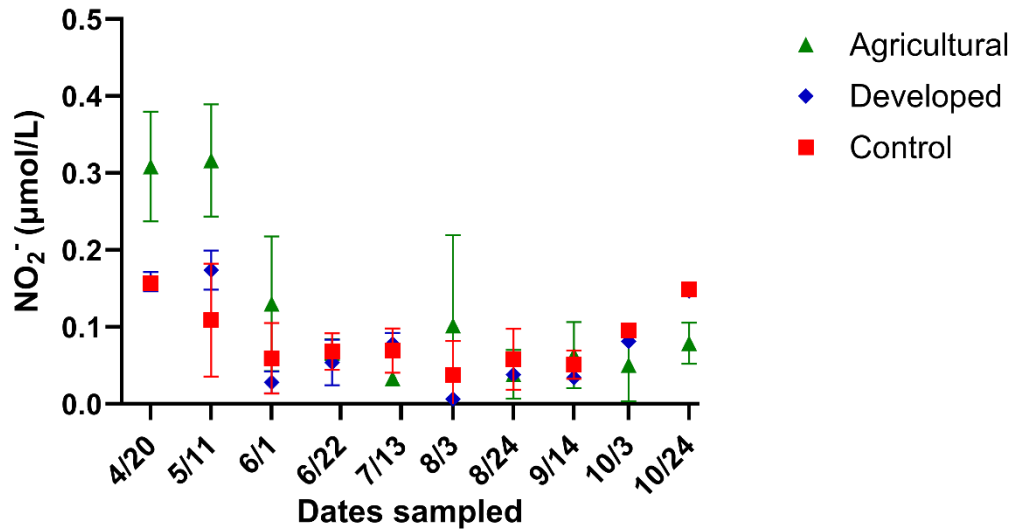


Figure 13: Dissolved nitrite concentrations in $\mu\text{mol/L}$ of land use categories through the year. Data points are means of two sampling sites for each land use type (control, developed, agricultural), error bars show standard deviation.

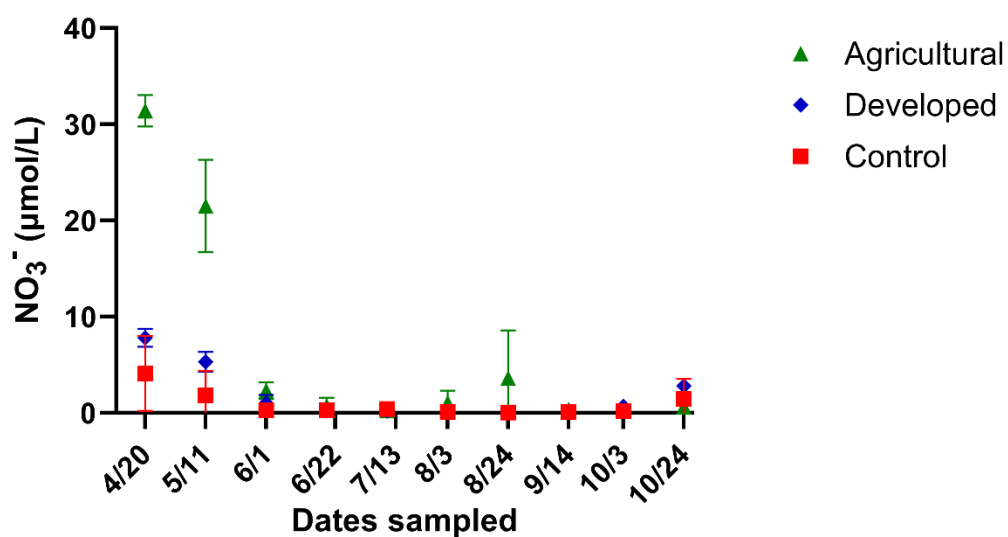


Figure 14: Dissolved nitrate concentrations in $\mu\text{mol/L}$ of land use categories through the year. Data points are means of two sampling sites for each land use type (control, developed, agricultural), error bars show standard deviation.

On April 20th and May 11th, the agricultural sites have higher nitrate concentrations than both control and developed sites. This is supported by statistical evidence. Control and developed stations, however, still had elevated nitrate levels on those dates. This increase in nitrate and nitrite during the first two sampling dates is of unknown origin. The source of these nutrients likely are not fertilizers, as urea is the fertilizer used by farmers. Due to global conflicts in 2023, urea fertilizers were scarce and not used a lot. The source could possibly be manure. Another possibility could be that the filling of the reservoir and the rewetting of coves lead to an increase in nitrogen. All three stations dropped off close to zero on June 1st and were not significantly different from each other throughout the rest of the year. A small increase in nitrate is recorded in the agricultural station on August 24th, yet not significant. Towards the end of the year, on October 24, nitrate is again increasing for all three stations.

Patterns in physical and chemical water properties

A principle component analysis was conducted. Parameters included in this plot were extracted chlorophyll, phycocyanin (PC), Secchi depth, conductivity (SPC), pH, water temperature, dissolved oxygen, nitrate, dissolved reactive phosphorus (DRP), nitrite, and silicate. Chlorophyll *a* (RFU), chlorophyll *a* ($\mu\text{g/L}$), phycocyanin (RFU), conductivity (C), and nitrate+nitrite ($\mu\text{mol/L}$) were not included in PCA analyses due to high correlation with other parameters (correlation above 0.9).

The resulting figure (Figure 15) depicts the clustering of sampling stations by land use. PC1 accounts for 38.6% of the total variation, while PC2 accounts for 23.3%. The control stations cluster towards the right. Perpendicular is the developed stations. Agricultural stations are more spread out and distance themselves from the control and developed stations. The

resulting parameters accountable for separation between Land use are Secchi depth, chlorophyll, phycocyanin and conductivity. The spread within land use groups can be explained by water temperature, nitrate, DRP, and nitrite.

Results of amplicon sequencing

The average mean quality score was 35.16% and the average percent of perfect barcodes was 91.43%. The average number of reads for 16S sequences were 102,109 and 2,130 for 18S sequences. On average, 78.0.2% of 16S and 72.8% of 18S reads survived (full table in Appendix).

Analysis of 16S/18S rRNA gene sequences

Prior to analyzing sequences were split into 16S and 18S rRNA gene sequences via a pipeline described in the methods. Then, sequences were taken and put through the same pipeline. Around 78.02% of 16S rRNA gene sequences survived the pipeline, while ~72.8% of 18S rRNA gene sequences passed quality control. 7776 different 16S rRNA gene ASVs and 2208 different 18S rRNA gene ASVs were annotated. rRNA gene 1.99% of total ASVs (97,485) were of eukaryotic descent, while the other 98.1% of ASVs (4,810,720) were prokaryotic.

Relative abundance of microbial organisms

Universal primers were used to sequence the 16S and 18S rRNA gene. Stacked bar charts were created with bacterial phyla contributing more than 1% to the total community. Eukaryotes were added as a separate color (Figure 16-19). There was little variation in the microbial community between land use groups. The top phyla were consistently *Actinobacteriota*, *Proteobacteria*, *Cyanobacteria*, *Bacteroidota*, *Planctomycetota*, *Verrucomicrobiota*, and *Chloroflexi*. Eukaryotic sequences made up ~2% of the total sequences. *Actinobacteriota*,

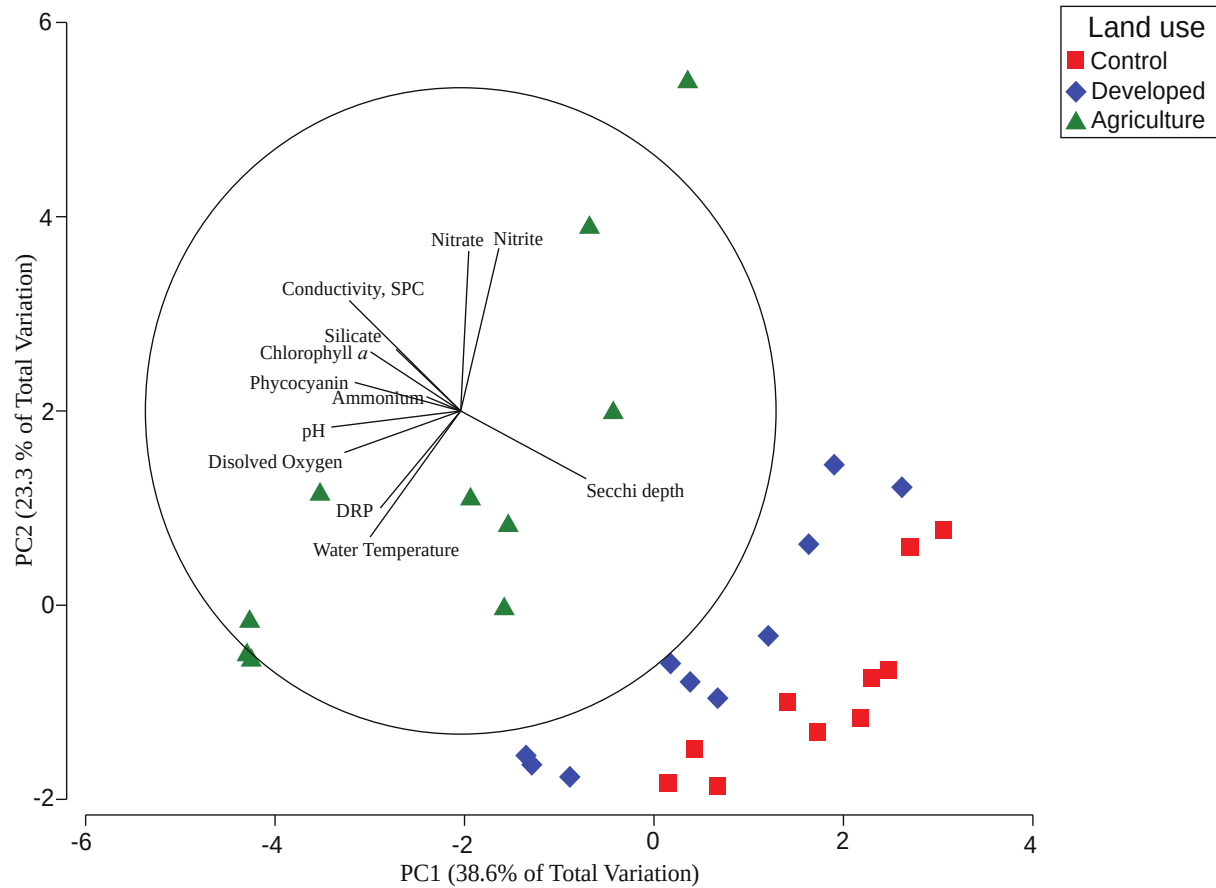


Figure 15: Principle component analysis (PCA) of environmental data collected. Data points represent a single sample, colored by land use type. Variables with longer vectors have a more significant influence on the corresponding principal component. Variables include Secchi depth (cm), water temperature ($^{\circ}\text{C}$), dissolved reactive phosphorus (DRP, $\mu\text{mol/L}$), dissolved oxygen, pH, Phycocyanin ($\mu\text{g/L}$), extracted Chlorophyll *a* ($\mu\text{mol/L}$), silicate ($\mu\text{mol/L}$), conductivity (SPC), nitrate ($\mu\text{mol/L}$), and nitrite ($\mu\text{mol/L}$).

Proteobacteria, and *Cyanobacteria* consistently made up 75% – 90% of the total community. *Actinobacteriota* were most abundant in the spring months, then remaining constant at around 25% of the community (Figure 16). *Proteobacteria* remain ~20% of the community throughout the year. *Cyanobacteria* increase throughout the spring, reaching their highest percent contribution in the summer months during July 13th and August 3rd. *Bacteroidota* concentrations were ~15% - 20% from April 20th until June 22nd, before dropping to ~10% for the rest of the year. *Planctomycetota* concentrations sharply increased to ~10% - 15% on August 3rd onward. Eukaryotic sequences were consistently less than 5% of the total community. The highest concentrations were on April 20th and September 14th. The most abundant eukaryotic phyla were *SAR* and *Cryptophyceae*. Other eukaryotic phyla that made up more than 1% of the total community were *Opisphokonta* and *Archaeplastida*.

The top eukaryotic sequences were from the phyla *SAR*, *Cryptophyceae*, *Opisthokonta*, and *Archaeplastida* (Figure 19-21). *SAR* consistently made up around or more than 50% of the total community. *Cryptophyceae* were a large part of the community, often making up over 25% of the community. During the dates of June 1st, June 22nd, September 13th, and August 3rd, the relative abundance declined steeply to ~10% of the community before increasing to ~25% for the rest of the year. *Opisthokonta* remained constant at ~5% of the community for the whole year, while *Archaeplastida* were barely detectable besides for the sampling dates of June 22nd, July 13th, August 3rd and August 24th, where it increased to up to 10% of the total community.

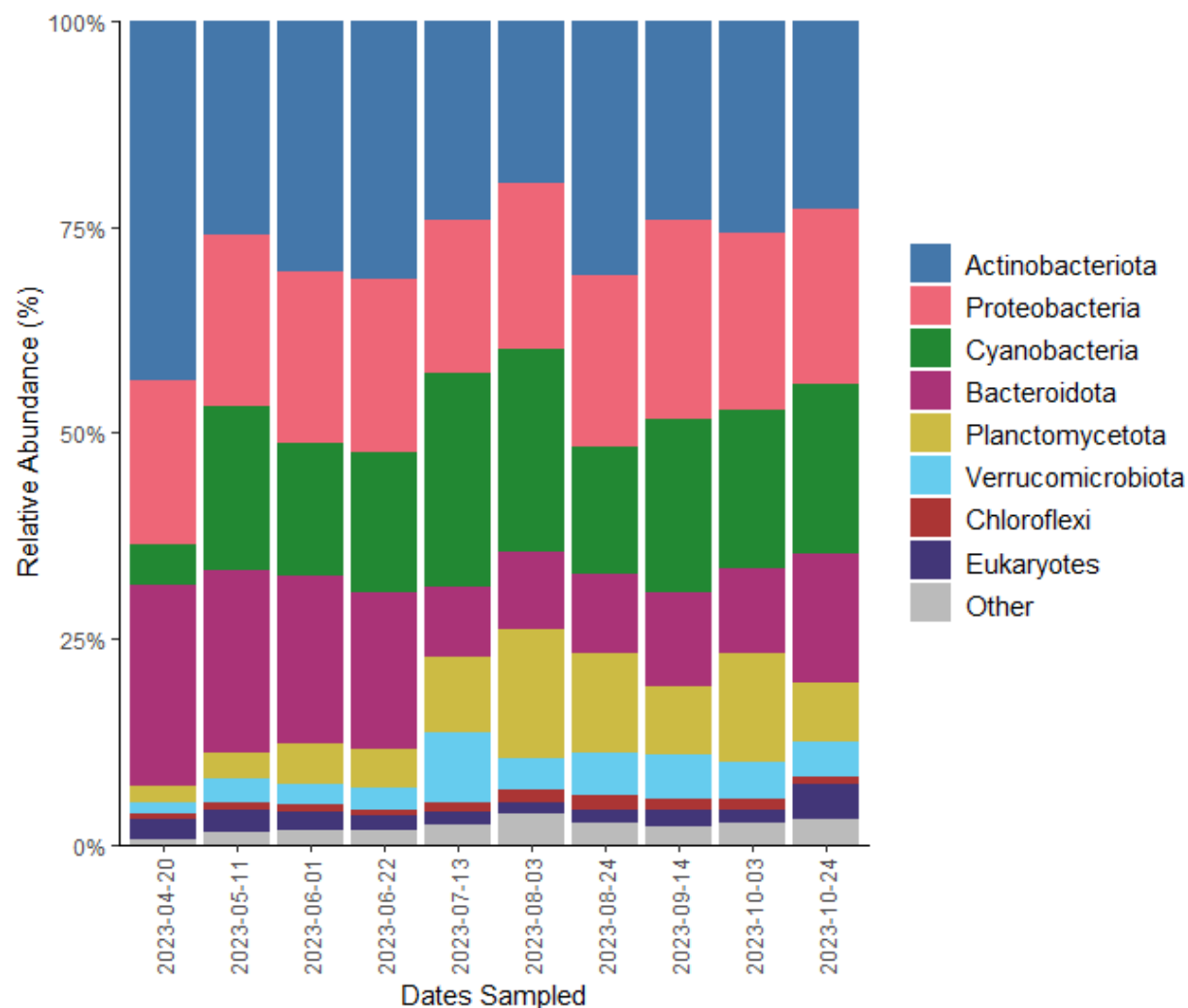


Figure 16: Relative abundance of microbes in sites influenced by agricultural land use. Bar plot showing the mean of relative abundance of microbes from stations Ag1 and Ag2 throughout the year at stations associated with agricultural land use. The label eukaryotes are all eukaryotic sequences and other is all phyla that made up less than 5% of the total community.

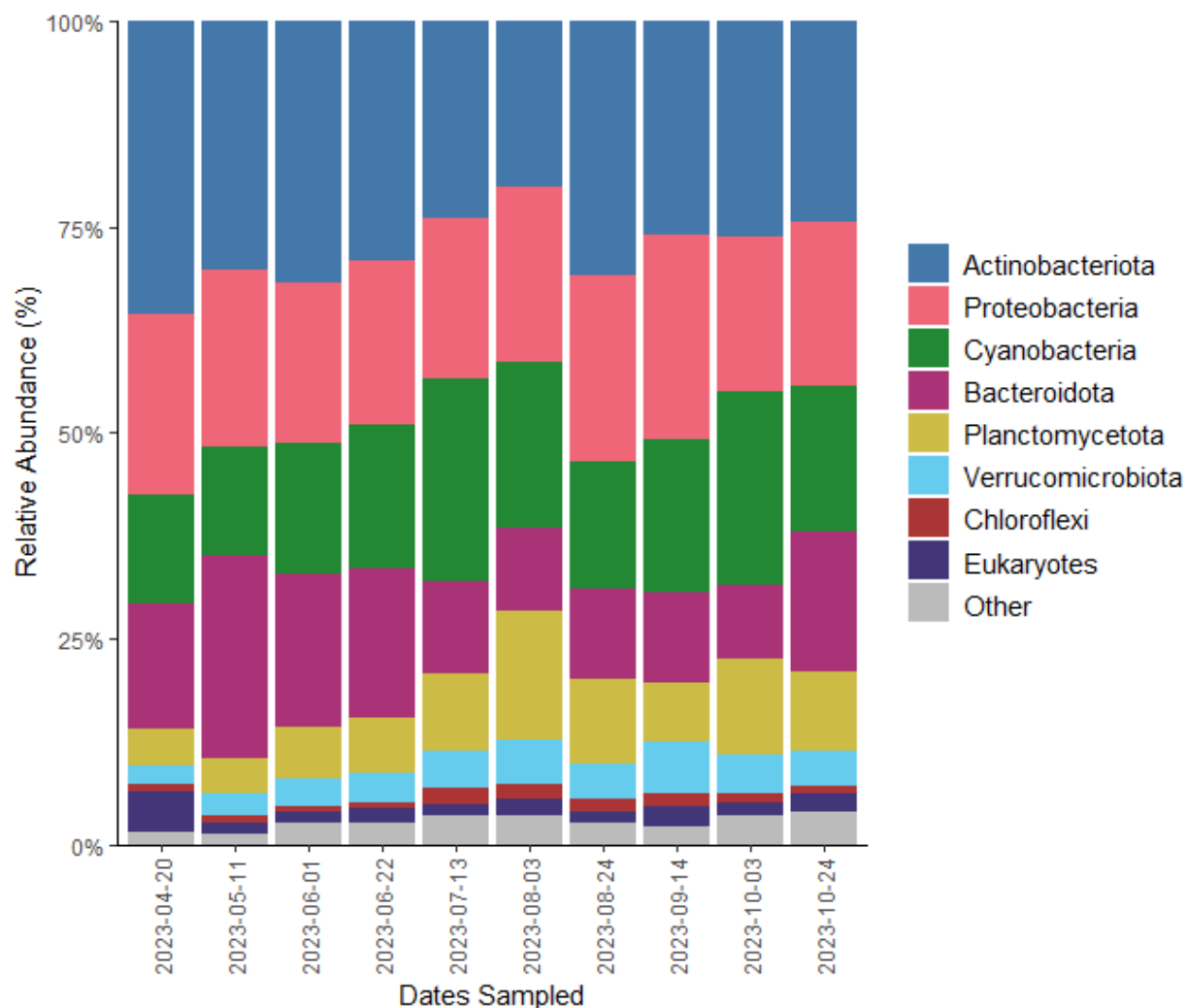


Figure 17: Relative abundance of microbes of stations associated with developed land use. Bar plot showing the mean of relative abundance of microbes from stations D1 and D2 throughout the year at stations associated with developed land use. The label eukaryotes are all eukaryotic sequences and other is all phyla that made up less than 5% of the total community.

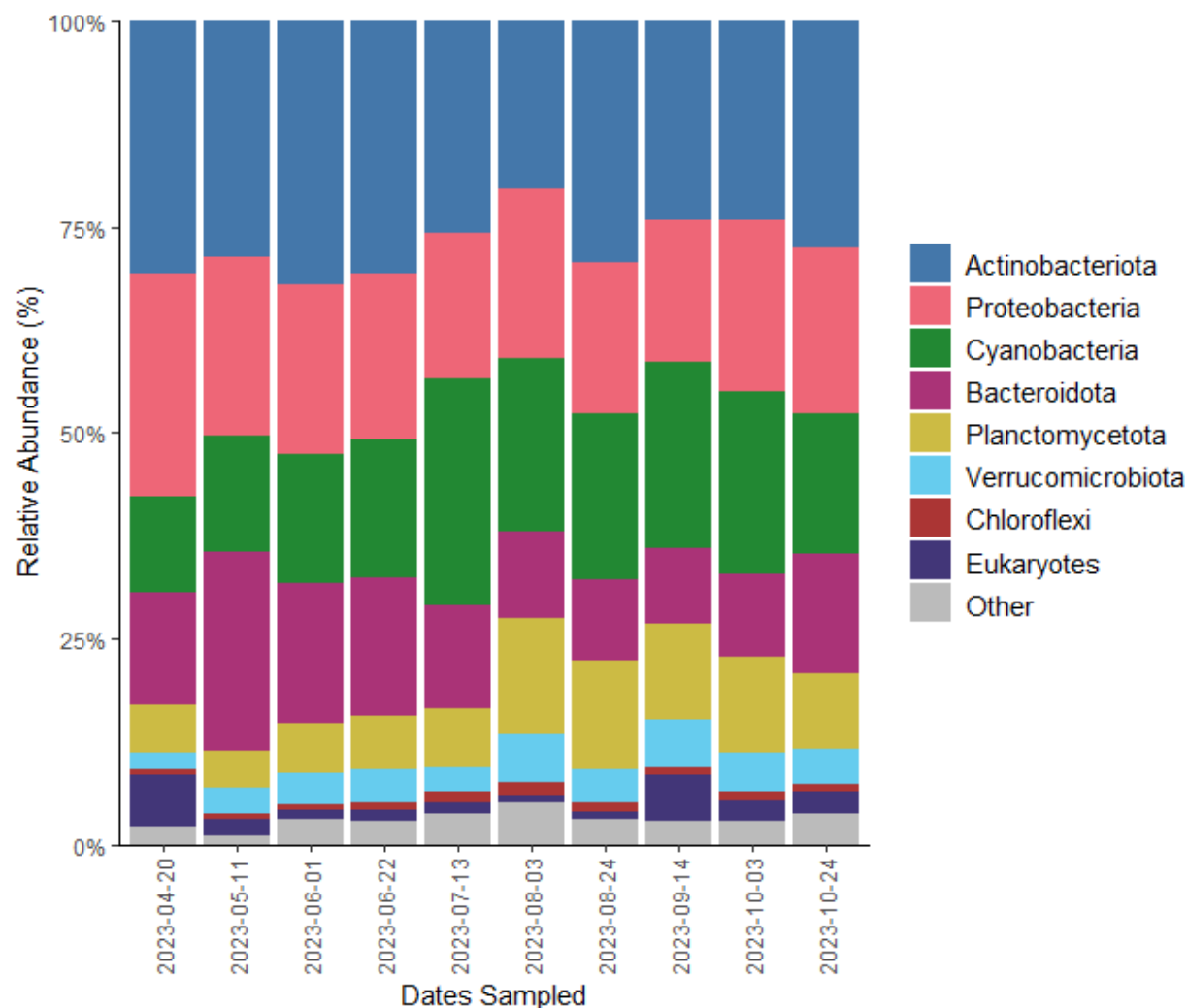


Figure 18: Relative abundance of microbes at control stations. Bar plot showing the mean of relative abundance of microbes from stations C1 and C2 throughout the year at control stations. The label eukaryotes are all eukaryotic sequences and other is all phyla that made up less than 5% of the total community.

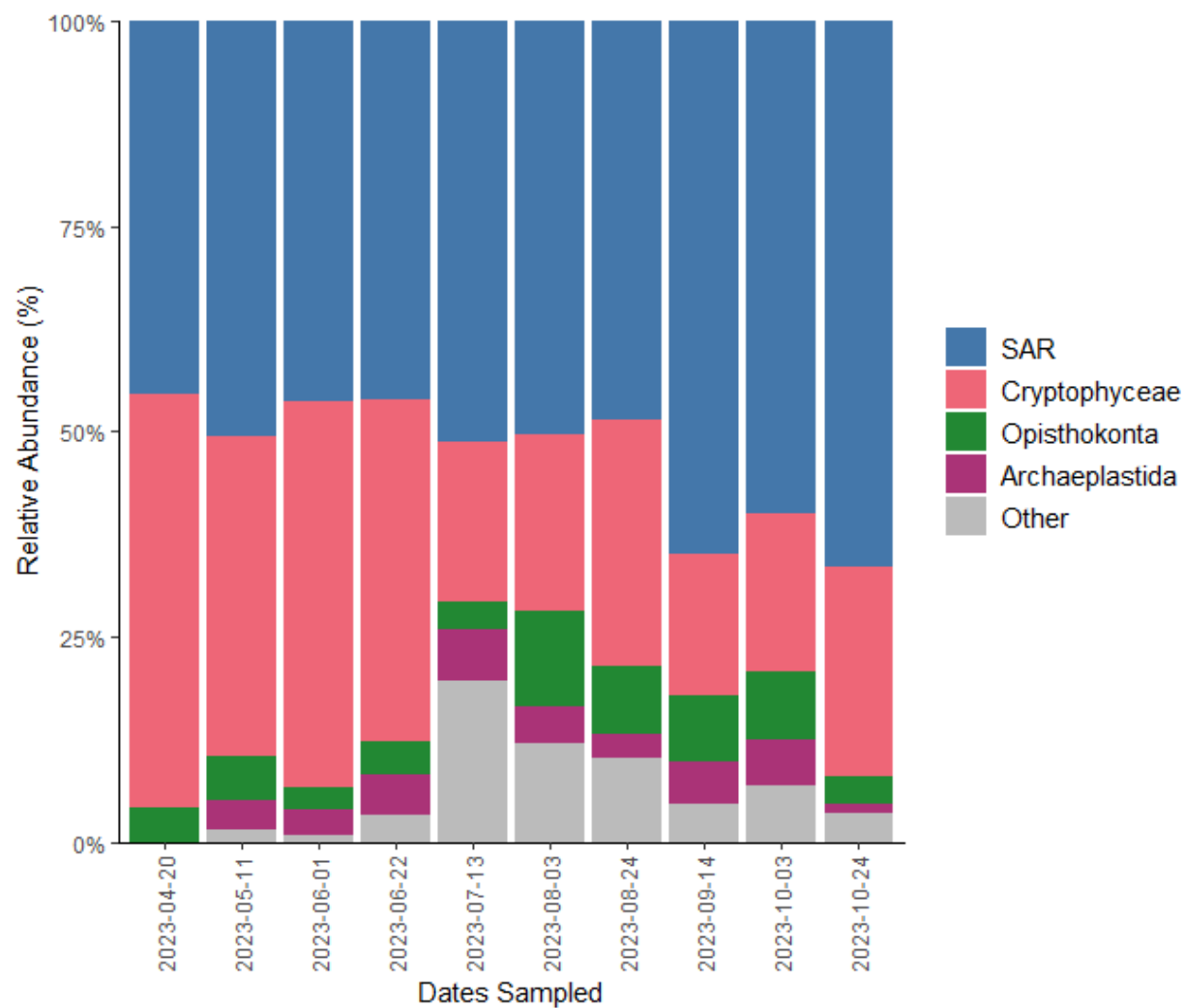


Figure 19: Relative abundance of eukaryotes at stations associated with agricultural land use. Bar plot showing the mean of relative abundance of eukaryotes from stations Ag1 and Ag2 throughout the year at stations associated with agricultural land use.

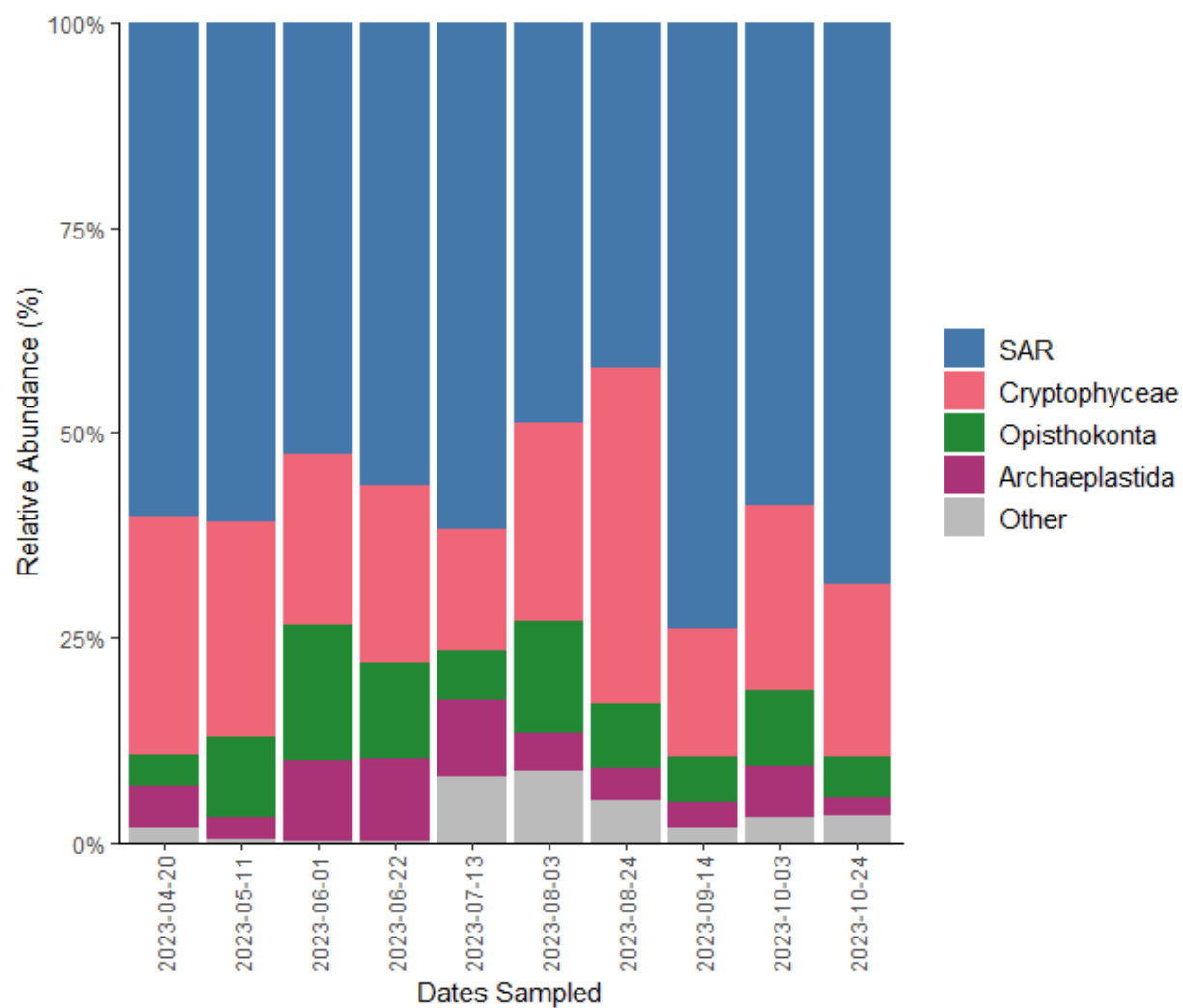


Figure 20: Relative abundance of eukaryotes at stations associated with developed land use. Bar plot showing the mean of relative abundance of eukaryotes from stations D1 and D2 throughout the year at stations associated with developed land use.

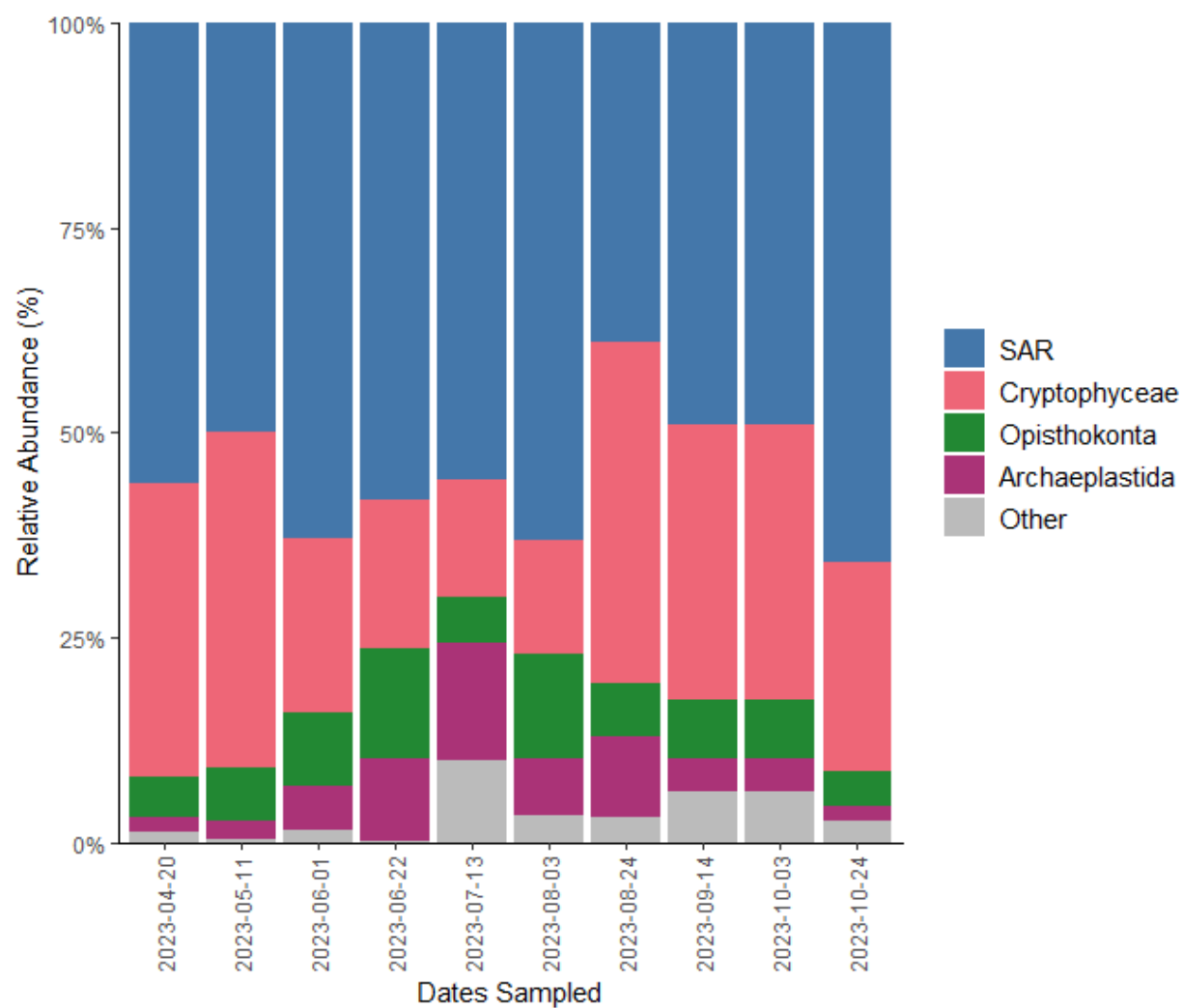


Figure 21: Relative abundance of eukaryotes at control stations. Bar plot showing the mean of relative abundance of eukaryotes from stations C1 and C2 throughout the year at control stations.

Clustering of samples based of ASV presence and abundance highlighted by land use category and season

A nMDS was performed on all ASVs. Station Ag1 was discarded from all further analysis due to very low read count compared to the other samples. Distinct clustering of samples collected in the spring can be observed towards the left side of the plot (Figure 22). Summer and Fall samples were overlapping and clustered towards the right side. Summer samples clustered closer together, while the Fall samples were more spread out. The clustering is supported by statistics. PERMANOVA analysis showed that the factor of season was a good predictor of clustering ($p=0.001$). When highlighting the samples by land use, no trend was clearly observable (Figure 23). Though PERMANOVA results suggested that the factor of land use was a good predictor ($p=0.004$). A two way crossed Simper was performed between land use and season to assess what ASVs were contributing the most to the dissimilarity between stations and seasons (Table 2-13).

Physicochemical drivers of population diversity

BIOENV for the factor land use

To determine what environmental variables might be contributing to the separation of stations, a BIOENV was performed (Clarke and Ainsworth, 1993). This method constructs a model to identify which physicochemical parameters best explain the patterns observed in the community composition data. To test for the environmental variables correlating with the spread resulted by highlighting stations by land use, the factor of land use was included in analysis. Correlations of 0.532, 0.532, 0.531, 0.530 were achieved by including six, seven, eight, or five variables,

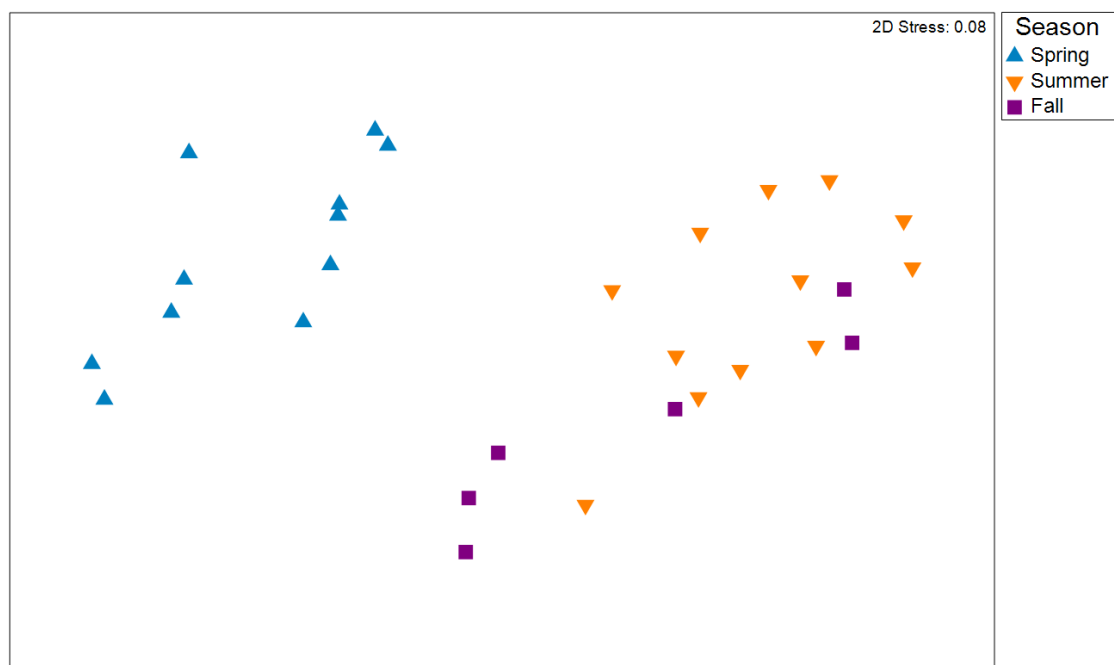


Figure 22: nMDS of Bray-Curtis distance between microbial communities from samples collected. Samples were color coded by season.

respectively. Variables always included were water temperature, phycocyanin, extracted chlorophyll *a*, silicate, and nitrate. Indeed, these variables alone result in a correlation of 0.53. Adding the sixth variable of Secchi depth, the correlation increases to 0.532. When adding nitrite as a seventh variable, the correlation stays at 0.532. Adding the variable ammonium as the eighth variable decreases the correlation to 0.531. Adding more variables decreases correlation (0.514 for nine, etc.)

BIOENV for the factor season

To test for the environmental variables correlating with spread of stations caused by highlighting samples by season, season was included as a factor in analysis. The highest correlation, 0.685, was achieved by including four variables. Those variables included water temperature, dissolved oxygen, phycocyanin, and nitrite. By adding pH and DRP, a correlation of 0.684 was recorded. Following closely, a correlation of 0.683 was achieved by water temperature, dissolved oxygen, and phycocyanin. The sole variable with the highest correlation was water temperature with 0.639.

Discussion

Site selection and implications for the study

Besides station D2, land use within the stations watershed mostly correlates with the labels assigned. D2 is in the middle of the reservoir and captures a large drainage area including the mountains, thus its main land use is forest. The imminent land use around the station, however, is developed. Although land use patterns correlate with the label assigned, other factors influence the reservoir. One of which is retention time. The retention time

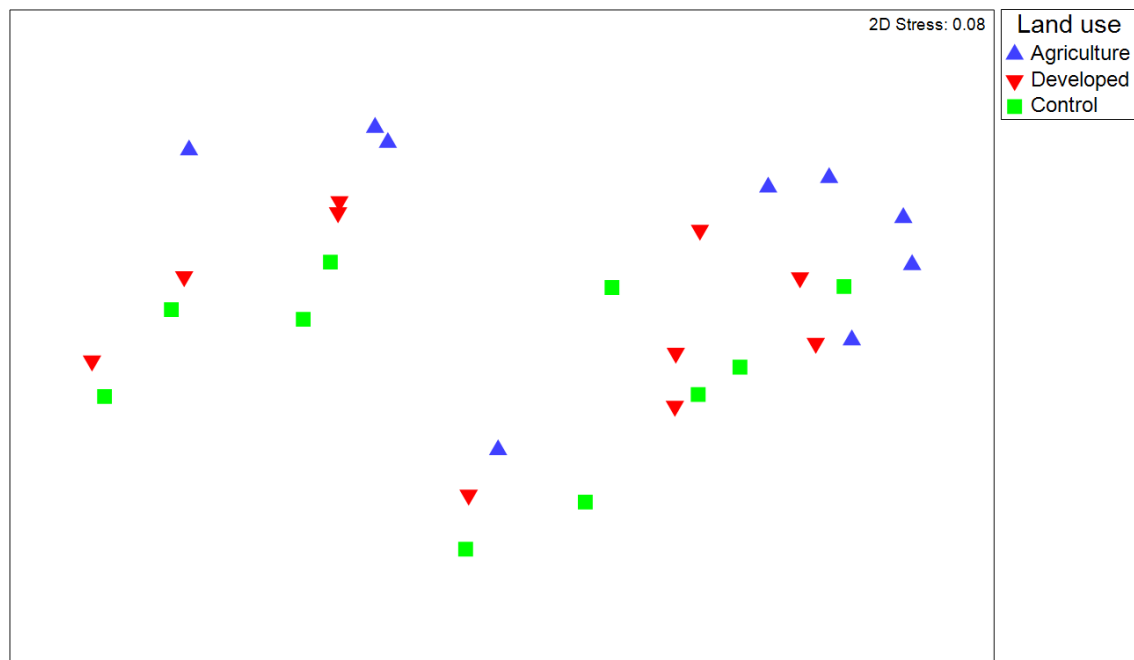


Figure 23: nMDS of Bray-Curtis distance between microbial communities from samples collected. Samples were color coded by land use category.

Table 3: ASVs contributing the most towards the similarity within stations associated with agriculture land use. The agriculture group had an average similarity of 55.12%.

Species	Average abundance	Contribution	Average Similarity	Consistency of Similarity	Inferred taxonomy
ASV_16S_1	109.38	2.09	109.38	5.67	Sporichthyaceae
ASV_16S_4	74.83	1.49	74.83	1.67	Sporichthyaceae
ASV_16S_3	69.34	1.26	69.34	1.68	Sporichthyaceae
ASV_16S_13	53.83	0.96	53.83	3.6	Cyanobium_PCC-6307

Table 4: ASVs contributing the most towards the similarity within stations associated with developed land use. The developed group had an average similarity of 53.88%.

Species	Average abundance	Contribution	Average Similarity	Consistency of Similarity	Inferred taxonomy
ASV_16S_1	116.25	2.04	1.1	4.55	Sporichthyaceae
ASV_16S_3	73.77	1.56	0.84	2.34	Sporichthyaceae
ASV_16S_5	58.87	1.17	0.63	2.47	Sporichthyaceae
ASV_16S_6	56.83	1.07	0.58	1.86	Sporichthyaceae
ASV_16S_4	67.92	1.07	0.58	1.47	Sporichthyaceae

Table 5: ASVs contributing the most towards the similarity within control stations. The average similarity between control stations was 54.07%.

Species	Average abundance	Contribution	Average Similarity	Consistency of Similarity	Inferred taxonomy
ASV_16S_1	102.71	2.02	1.09	4.89	Sporichthyaceae
ASV_16S_3	75.49	1.75	0.95	2.96	Sporichthyaceae
ASV_16S_5	62.59	1.45	0.78	3.85	Sporichthyaceae
ASV_16S_2	66.23	1.21	0.65	1.02	Cyanobium_PCC-6307
ASV_16S_6	55.12	1.16	0.63	2.35	Sporichthyaceae

Table 6: ASVs contributing the most towards the similarity between developed and control stations. The developed and control station had an average similarity of 54.28%.

Species	Group Developed Av. Abundance	Group Control Av. Abundance	Average dissimilarity	Consistency of dissimilarity	Contribution	Inferred taxonomy
ASV_16S_1	116.25	102.71	0.22	1.31	0.48	Sporichthyaceae
ASV_16S_4	67.92	52.47	0.15	0.96	0.33	Sporichthyaceae
ASV_16S_2	64.48	66.23	0.14	0.97	0.31	Cyanobium_PCC-6307
ASV_16S_7	47.63	46.84	0.14	1.33	0.30	Cyanobium_PCC-6307
ASV_16S_8	50.67	39.72	0.13	1.42	0.28	Proteobacteria

Table 7: ASVs contributing the most towards the similarity between developed and agriculture stations. The average similarity between developed and agriculture stations was 50.85%.

Species	Group Developed Av. Abundance	Group Agriculture Av. Abundance	Average dissimilarity	Contribution	Consistency of dissimilarity	Inferred taxonomy
ASV_16S_2	64.48	47.17	0.19	0.38	1.07	Cyanobium_PCC-6307
ASV_16S_1	116.25	109.38	0.18	0.38	1.16	Sporichthyaceae
ASV_16S_19	29.61	36.33	0.16	0.33	0.93	Bacteroidota
ASV_16S_7	47.63	29.00	0.13	0.26	1.45	Cyanobium_PCC-6307

Table 8: ASVs contributing the most towards the similarity between control and agriculture stations. The average similarity between control and agriculture was 43.55%.

Species	Group Control Av. Abundance	Group Agriculture Av. Abundance	Average dissimilarity	Contribution	Consistency of dissimilarity (Diss/SD)	Inferred taxonomy
---------	--------------------------------	------------------------------------	--------------------------	--------------	--	-------------------

ASV_16S_2	66.23	47.17	0.22	0.39	1.08	Cyanobium_PCC-6307
-----------	-------	-------	------	------	------	--------------------

Table 8 continued

ASV_16S_4	52.47	74.83	0.18	0.32	1.27	Sporichthyaceae
ASV_16S_1	102.71	109.38	0.18	0.32	1.10	Sporichthyaceae
ASV_16S_11	21.39	41.28	0.18	0.32	0.92	Sporichthyaceae
ASV_16S_13	31.57	53.83	0.17	0.31	1.73	Cyanobium_PCC-6307
ASV_16S_19	25.30	36.33	0.16	0.29	0.89	Bacteroidota

Table 9: ASVs contributing the most towards the similarity within the spring season. The spring had an average similarity of 55.31%.

Species	Average abundance	Contribution	Average Similarity	Consistency of Similarity	Inferred taxonomy
ASV_16S_3	97.33	2.24	1.24	17.04	Sporichthyaceae
ASV_16S_1	108.97	2.03	1.12	3.56	Sporichthyaceae
ASV_16S_9	68.96	1.56	0.86	9.60	Proteobacteria
ASV_16S_5	64.27	1.54	0.85	7.36	Sporichthyaceae
ASV_16S_6	63.77	1.50	0.83	5.61	Sporichthyaceae
ASV_16S_15	54.44	1.26	0.70	8.32	Sporichthyaceae
ASV_16S_7	58.54	1.10	0.61	3.27	Cyanobium_PCC-6307
ASV_16S_17	45.53	0.99	0.55	4.56	Proteobacteria

Table 10: ASVs contributing the most towards the similarity within the summer season. The summer had an average similarity of 55.91%.

Species	Average abundance	Contribution	Average Similarity	Consistency of Similarity	Inferred taxonomy
ASV_16S_1	110.33	2.03	1.13	10.82	Sporichthyaceae
ASV_16S_4	98.07	1.72	0.96	4.38	Sporichthyaceae
ASV_16S_2	98.27	1.63	0.91	2.67	Cyanobium_PCC-6307
ASV_16S_10	59.35	1.04	0.58	5.92	Proteobacteria

ASV_16S_3	57.95	1.02	0.57	3.73	Sporichthyaceae
-----------	-------	------	------	------	-----------------

Table 11: ASVs contributing the most towards the similarity within the fall season. The fall had an average similarity of 39.57%.

Species	Average abundance	Contribution	Average Similarity	Consistency of Similarity	Inferred taxonomy
ASV_16S_1	108.57	2.35	0.93	11.25	Sporichthyaceae
ASV_16S_4	83.02	1.51	0.60	7.18	Sporichthyaceae
ASV_16S_2	72.65	1.23	0.49	8.48	Cyanobium_PCC-6307
ASV_16S_3	58.42	1.13	0.45	3.72	Sporichthyaceae
ASV_16S_10	48.04	1.05	0.41	9.09	Proteobacteria
ASV_16S_5	56.14	0.99	0.39	3.26	Sporichthyaceae

Table 12: ASVs contributing the most towards the similarity between spring and summer. The average similarity between spring and summer groups were 29.98%.

Species	Group Spring Av. Abundance	Group Summer Av. Abundance	Average dissimilarity	Contribution	Consistency of dissimilarity (Diss/SD)	Inferred taxonomy
ASV_16S_2	10.58	98.27	0.57	0.82	1.98	Cyanobium_PCC-6307
ASV_16S_4	18.40	98.07	0.47	0.67	3.30	Sporichthyaceae
ASV_16S_14	0.00	53.26	0.33	0.47	2.59	Planctomycetota
ASV_16S_12	1.87	56.20	0.32	0.46	3.40	Cyanobium_PCC-6307
ASV_16S_9	68.96	17.60	0.32	0.45	4.18	Proteobacteria
ASV_16S_11	0.00	49.65	0.28	0.40	1.82	Sporichthyaceae

Table 13: ASVs contributing the most towards the similarity between spring and fall. The average similarity between spring and fall was 31.6%.

Species	Group Spring Av. Abundance	Group Fall Av. Abundance	Average dissimilarity	Contribution	Consistency of dissimilarity (Diss/SD)	Inferred taxonomy
ASV_16S_4	18.40	83.02	0.36	0.52	2.50	Sporichthyaceae
ASV_16S_2	10.58	72.65	0.35	0.51	1.55	Cyanobium_PCC-6307
ASV_16S_14	0.00	47.95	0.27	0.39	1.86	Planctomycetota
ASV_16S_11	0.00	44.03	0.24	0.36	1.26	Sporichthyaceae
ASV_16S_3	97.33	58.42	0.23	0.33	1.37	Sporichthyaceae
ASV_16S_9	68.96	29.53	0.22	0.33	1.74	Proteobacteria

Table 14: ASVs contributing the most towards the similarity between summer and fall. The average similarity between summer and fall was 45.07%.

Species	Group Summer Av. Abundance	Group Fall Av. Abundance	Average dissimilarity	Contribution	Consistency of dissimilarity (Diss/SD)	Inferred taxonomy
ASV_16S_2	98.27	72.65	0.21	0.39	1.49	Cyanobium_PCC-6307
ASV_16S_11	49.65	44.03	0.21	0.38	1.18	Sporichthyaceae
ASV_16S_4	98.07	83.02	0.19	0.34	1.29	Sporichthyaceae
ASV_16S_12	56.20	40.65	0.18	0.33	1.37	Cyanobium_PCC-6307
ASV_16S_6	33.02	64.11	0.18	0.32	1.25	Sporichthyaceae

for Tellico Reservoir is ~50 days. Even though this time is the highest of TVA's run-of-river reservoirs, it is less than that of other TVA tributaries. The short retention time may be more impactful on the reservoir than land use within a station's watershed. With a low retention time, proximity to the main channel has to be considered. Stations further from the main channel have longer retention times than stations closer.

Variation within controls

Although the land use of both control stations was mostly forest, the comparability of the stations is questioned, due to the distance from each other. One station was located in the developed area (C1), while the other was upstream (C2). Water, impacted from the land use around the reservoir, flows to C1. Thus, it is not clear whether the microbial community in C1 is impacted by the land use around the station, or by the impacted water flowing to the station. To understand the variability within the control group, physicochemical properties and community composition between C1 and C2 was compared. A PCA between C1 and C2 highlights physicochemical differences between the two stations (Appendix). This is mainly driven by higher chlorophyll and phycocyanin in station C1. However, this trend is not reflected in the community composition data. A nMDS of Bray-Curtis distance between the two stations reveals no difference between the two. This is supported when correlating the variance in the microbial communities against each other. The standardized abundances plotted against each other showed a R^2 value of 0.89 (Appendix). A strong correlation between the two stations shows the variation between the control stations is limited. In addition, a nMDS of Bray-Curtis distance did not reveal clustering of stations (Appendix).

Seasonal chlorophyll *a* pattern

As shown in the results, chlorophyll *a* tends to increase throughout the season. This pattern is not unexpected, as sampling started in the spring, where algal biomass is typically low. Outliers do occur on June 22nd, where values were higher than expected based on the trend. Higher chlorophyll *a* in transition from colder to warmer months can be attributed to the increase in water temperature and biological activity. The steady increase throughout the season can be explained by a shift in microbial community, where algae with higher chlorophyll contents dominate later in the year.

Seasonal nutrient dynamics

High ammonium concentrations early in the year suggest runoff from the surrounding area coming into the reservoir. This correlates with rainfall recorded around the reservoir, as ~3-5 in of rain was recorded. Increased ammonium concentrations, however, are not limited to rainfall runoff. A spike in ammonium was recorded in September and late October, yet no rain fell on those months. This suggests another variable influencing ammonium inside the reservoir. Though not significant, ammonium concentrations were higher in agricultural stations later in the year. This is especially interesting since farmers most likely did not fertilize their fields with anhydrous ammonia fertilizers. Due to global conflicts in 2022/2023, the price of ammonia fertilizers was high. The source of the spike is unclear, but it could lie in the usage of manure. Another reason for the increased nitrogen in the spring may be the rewetting of coves. Coves that run dry in the winter, when the water gets lowered, may accumulate nitrogen that gets released into the system when water levels increase in the spring. A spike in DRP could be accredited to large rainfall, as a spike was detected from June until August. Another reason for a spike in DRP

may be nitrogen limitation. Both nitrite and nitrate lowest concentrations correlate with the surge of DRP. Nitrogen limitation may result in a leftover of dissolved phosphorus, which can be detected by nutrient analysis. These results suggest a trend, where the reservoir starts phosphorus limited, then becomes nitrogen limited in the summer. After August, the DRP drops and the reservoir becomes phosphorus limited again. This trend is also observable in other freshwater bodies. Silicate concentrations were not limiting year round, suggesting that silica dependent organisms' abundance, like diatoms, were not limited by dissolved silicate availability.

Physicochemical parameters driving differences between land use categories

To further bolster the assumption that stations with different land use inside their watershed harbor different microbial communities, PCA was conducted. This will show whether environmental conditions inside the water is different with land use. Indeed, distinct clustering of sample taken by land use can be observed. The main drivers in the distinction are Secchi depth, chlorophyll, conductivity and phycocyanin. Secchi depth can be influenced by chromophoric dissolved organic matter (CDOM). CDOM can originate from both biological and geological sources. On one hand, chlorophyll and phycocyanin influence Secchi depth, as a higher algal biomass does reduce water penetration. In addition, dead cells float as debris, blocking sunlight. On the other hand, stirred-up sediments and runoff carrying large quantities of CDOM can lower Secchi depth. Agricultural stations had a lower Secchi depth, suggesting more CDOM inside the water (Appendix). This is supported by a large difference in conductivity between land use groups. Larger nitrate, nitrite, and ammonium concentration in agricultural groups could explain the larger chlorophyll and phycocyanin values. All this evidence suggests that agricultural stations receive higher CDOM carrying runoff throughout the year. Separation within a given

land use group is explained by water temperature, DRP, nitrate, and nitrite. The distinct spike in nitrate and nitrite observed in the spring and the large spike of DRP in the summer, causes separation in the PCA. Water temperature is another major driver of the separation within a land use group. As the season progresses, water temperature changes, which has an impact on the water column. Although not included in the PCA, one thing to consider is distance from the main channel as a predictor of separation. Both agricultural stations are located inside an arm of the reservoir. All of the control and developed stations are either collected from the main channel or from a cove directly neighboring the main channel. It is unclear whether Secchi depth is a direct result of the drivers discussed earlier, or by a larger residence time of water inside agricultural stations. However, control and developed stations do separate, even with their proximity to the main channel.

Total community composition

The most abundant phyla throughout the year were *Actinobacteriota*, *Proteobacteria*, and *Cyanobacteria*. Together they consistently made up close to 75% of the total community. All these phyla are common members of the freshwater community (Newton et al., 2011). *Actinobacteriota* were first thought to be dominant in soils, yet 16S rRNA studies revealed their presence and importance in freshwater habitats (Glöckner et al., 2000). It is not surprising that this phylum dominates the community, as it can make up to 50% of freshwater community (Glöckner et al., 2000). It is suggested that members of *Actinobacteria* produce secondary metabolites, due to their yellow/red color (Newton et al., 2011). Although they are primarily heterotrophic, their dominance throughout the year may be contributed to the presence of actinorhodopsins, which enables phototrophy (Sharma et al., 2009). The phylum was dominated

by the family *Sporichthyaceae*. Though originally isolated from soils, *Sporichthyaceae* is commonly a large member of the microbial community (He et al., 2022; Lefler et al., 2022; Trujillo and Normand, <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118960608.fbm00022.pub2>). It is especially common in either turbid waters or waters impacted by anthropogenic activity (Gołębiewski et al., 2017; Newton and McLellan, 2015; Zhang et al., 2021). The family also has been known to be able to consume complex carbon molecules, contributing to its dominance (Trujillo and Normand, <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118960608.fbm00022.pub2>). The dominant taxon was *Sporichthyaceae* *hgcl_clade* with the species being unclassified.

Proteobacteria are also common members of the freshwater epilimnions. The phylum is vastly studied in the apicultural and medical context (Chen et al., 2019). The majority of *Proteobacteria* fell under the class *Betaproteobacteria*. The class is often numerically dominant in freshwater lakes (Glöckner et al., 2000). Members of that class are often cocultured with *Cyanobacteria* and *Cryptomonas*, both found in the reservoir (Eiler et al., 2006; Pernthaler et al., 2001). Genera such as *Limnohabitants* and *Polynucleobacter* dominated the class. *Limnohabitants* is a chemoorganotrophic bacterium that responds quickly to nutrient pulses and can use algal derived substrates as growth source (Šimek et al., 2006; Šimek et al., 2010). The other genus found was *Polynucleobacter*. This genus is a chemoorganotroph that performs aerobic respiration, though is capable of facultative anaerobic growth. Both genera are living an opportunistic lifestyle, where they are especially prone to grazing with their high growth rates with nutrient influx (Hornák et al., 2006).

Above 95% of *Cyanobacteria* ASVs were annotated as *Cyanobium* PCC6307. This taxon is also known as *Cyanobium gracile* PCC6307. It falls under the order of *Synechococcales* and is often a large member of the picocyanobacterial community (Che et al., 2022). It dominates communities regardless of trophic state, though stays in higher light levels (Pantoja-Agreda and Pajares, 2024; Yanez-Montalvo et al., 2022). In addition, picocyanobacteria have been found to thrive in turbid lakes, conditions relevant to Tellico reservoir (Felföldi et al., 2009). Interestingly, a difference in cyanobacterial abundance between land use groups is not detectable, even though agricultural sites had significantly lower Secchi depth readings.

The phyla *Bacteroidota* and *Planctomycetota* together also made up ~20% of the community throughout the year. *Bacteroidata* have a large phenotypic and metabolic diversity, thus are commonly found in a variety of habitats. Most members of the phylum are chemoorganotrophs. They usually occupy a large proportion of particle associated bacteria, partly due to their ability to degrade complex biopolymers (Kirchman, 2002; Newton et al., 2011). Studies have shown an increase in *Bacteroidata* during periods of high input of DOC or algal-derived DOC (Eiler and Bertilsson, 2004; Zeder et al., 2009). Higher abundances of *Bacteroidota* agree with rainfall in the spring, however, declining abundance in the spring is influenced by an unknown factor. Furthermore, the phylum is also a copiotroph, strengthening the hypothesis of unregular spikes of nutrients into the reservoir.

In the context of freshwater, *Planctomycetota* is an understudied phylum, even though it is commonly found. It is studied extensively due to their anaerobic ammonium oxidation capabilities (annamox) and can degrade algal derived carbohydrates (Newton et al., 2011). The annamox capabilities, suggest the importance in the nitrogen cycle of Tellico Reservoir.

In addition to amplicon sequencing, the community was also assessed with the use of a fluid imaging device, the FlowCAM (Appendix). The FlowCAM can image most plankton in aquatic systems although it does have a minimum size requirement (15 μm), so picoplankton were not observed.

To process data, VisualSpreadsheet 5 was used to sort the images into different groups, based on parameters collected by the FlowCAM (Poulton, 2016). The microbial community observed by the FlowCAM was dominated by small colonial light-green microbes (Group A, Appendix). The exact taxonomy of this group remains unclear. Throughout the year, this group made up above 25% and up to 50% of the community regardless of land use category. Contrary to the microbial community, a spike in filamentous microbes was observed starting in early summer (6/22/23-8/04/23) This group (Group B, Appendix) rose to peak abundance during mid-summer, before fading in the fall. The characteristics of Group C are similar to Group A. Although they are also colonial microbes, they are darker in color (Appendix). Included are also smaller, non-round organisms. Due to the similarity between this group and Group A, it may be that the microbes in this group were images incorrectly and actually belong to Group A. Images with very round features were collected in Group D (Appendix). These microbes were constant throughout the year with an abundance of 5-10% of the total community. Group E consisted of large microbes (Appendix). Key members of this group were rotifers, large ciliates and diatoms like *Asterionella formosa*. This group was not abundant, often making up less than 1% of the community count wise, but, due to their size, likely much more if converted to biomass. It should be noted that about 25-35% of the images taken were not able to be classified (Group F). This was mainly due to blurry images although a large portion of this group looks to be from Group A.

Overall, these plankton communities were comparable between land use categories. Group A was the most abundant and the surge of Group B was recorded in every land use category. An exception to this trend was station C2 (Figure 37). Here, the group stayed at ~5% throughout the whole year. Generally, the microbial community in this station does not shift and stays constant throughout the year.

Differences in the community were highlighted between the two different techniques (amplicon sequences and FlowCAM). The FlowCAM could not capture the community smaller than 15 μm in size. Interestingly, the FlowCAM recorded large abundances of filamentous microbes, which were not recorded in the sequence data. This may be due to incomplete lysis of the filaments during DNA extraction or problems during sequencing.

Distance between land use categories in ordinate space

As shown earlier, a large difference between land use categories was observed with physicochemical parameters. However, the opposite was true based on species composition of the microbial community. Stations did not cluster out based on land use. Even though stations did not noticeably cluster based on land use, statistical analyses suggest an observable trend. In fact, it seems like agricultural stations cluster out to the top of the ordinate space, while control and developed stations are mixed in the middle and bottom. This trend was observed at several other times, where control and developed stations are similar to each other, and agricultural stations cluster out.

The opposite is true for season. A strong trend is noticeable with statistics when highlighting seasonal differences. The spring falls towards the left, while the summer and fall clusters towards the right. Although summer and fall overlap a little, the majority of the summer

samples cluster together. The fall samples are not clustering tightly. This may be explained by the overlapping seasons in the Southeast of the United States. Often, summer and fall blend together and have similar climates and temperatures, causing the samples collected to cluster together based on similar environmental conditions

ASVs most contributing to similarity in a land use category and season

Land use

All three land use groups had an average similarity within the land use category of ~54-55%. The family *Sporichthyaceae* was always a top member driving the similarity. This suggests that *Sporichthyaceae* is a member of the community throughout the year in all three land use categories. Its consistency and high abundance make it a vital microbe to study in Tellico Reservoir. The similarity within each site being around the same, also validates the selection of sample sites based on land use patterns.

When comparing land use categories against each other, differences in species composition can be detected. The highest similarity (54.28%) could be observed between the stations control and developed. They were more similar to each other than dissimilar. Here, *Sporichthyaceae* and *Cyanobium* PCC6307 were driving the differences. Although both of these genera are very abundant in both land use groups, the developed group had higher abundances.

The average similarity of developed and agriculture stations was 50.85%. Still more similar to each other than dissimilar, due to the similar dominance of *Sporichthyaceae*. Contributing to the difference is a higher abundance of *Cyanobium* PCC6307 in the developed group, as well as higher *Sporichthyaceae* abundance.

Control and agriculture groups are less similar. They are 43.55% similar with *Cyanobium* PCC6307 and *Sporichthyaceae* being the main drivers. *Sporichthyaceae* abundances are higher in agricultural stations, while one species of *Cyanobium* PCC6307 is more abundant in one station and another in the control station.

Comparing stations to another validates the hypothesis that control, and developed stations are very similar, while control and agriculture stations more are dissimilar. Developed stations find themselves between the two, sharing similarities with both stations. Molecular analysis has showed, however, that the difference between land use groups are not as dramatic as previously highlighted by environmental analyses.

Season

The similarity within each season was comparable to that in land use categories. Both spring and summer had a similarity of ~55%. The fall had a similarity of 39.57%, lower than the other two. This is reflected visually in the nMDS, as there are two clusters of fall groups. The cluster on the bottom of the graph were samples taken later in the year. Although fall and summer overlap very closely in the Southeast of the US, temperature decrease in the later fall months. It is predicted that further samples taken would cluster closer to the fall samples at the bottom, which are the last samples collected.

Comparing seasons to each other highlights a strong seasonal trend. Spring and summer had an average similarity of ~30%. High abundances of *Cyanobium* PCC6307, *Sporichthyaceae* and *Planctomycetota* were all detectable in the summer. The only genus that was more abundant in the spring was *Limnohabitans* (*Proteobacteria*). This genus is generally found in freshwater,

including colder waters, yet also showed adaptations to warmer weather, so it is unclear why the abundance diminished in the summer (Cruaud et al., 2020; Zufiaurre et al., 2022).

Similar to the spring and summer, the spring and fall also had a low similarity of 32%. Again, a steep increase in abundance of *Sporichthyaceae* and *Cyanobium* PCC6307 characterized the difference. In addition, a decrease of *Limnohabitans* could be detected, as well as an increase in *Planctomycetota*.

The similarity between summer and fall is the highest (~45%). No drastic shifts in species abundance could be detected. No species emerged or disappeared. Rather, different species from the same phylum seem to overtake one-another. Both *Cyanobium* PCC6307 and *Sporichthyaceae* species were the biggest drivers of dissimilarity.

Although there is not a strong observable trend in relative abundance throughout the season, SIMPER analyses suggest a strong seasonal trend. Some phyla seem to dominate throughout the season (*Actinobacteria*, *Cyanobacteria*), while others are more abundant in the spring (*Proteobacteria*). The results of this test fortify the trend observed in previous data. Spring and summer/fall are different from one another, while summer and fall are similar.

Physicochemical parameters correlating with species composition

To investigate what environmental variables drive the separation observed by the nMDS, a BIOENV was conducted. Trying to predict the separation of land use results in a model using six or seven variables. Several nutrient concentrations were included, as well as chlorophyll and phycocyanin. Interestingly, water temperature was also included. The exact water circulations and physical conditions in the reservoir are not explored yet, but water temperature could be included due to the different time in sampling the stations. Station Ag1 was generally sampled in

the morning, while station C2 was sampled midday. Secchi depth was also included, further leaning into the importance of turbidity and light penetration of each land use group.

When looking for variables driving the seasonal variation, a higher correlation of 0.685 was achieved between the community composition data and selected physicochemical parameters. Only four variables were included in this model. Differences in nitrite reflect seasonal runoff. Differences in water temperature favor phycocyanin producing microbes. Interestingly, chlorophyll *a* was not included in this model. This may be due to the fact that in colder months the community shifts towards higher chlorophyll producing organisms, increasing chlorophyll values. Water temperature was the most important variable. Indeed, it alone results in a correlation of 0.639. The shift in water temperature from spring to summer is reflected in the nMDS, as well as the prolonging of warm temperatures. High surface water temperatures explain the similarities between summer and fall communities.

Conclusion

The presumed land use of selected stations correlated with the assigned land use from the national land cover database. Due to its large drainage area, station D2 primary land use was forest, yet the immediate land use around the sampling location was developed. Although stations C2 was upstream, separated from all the other stations, the microbial community did not differ from the other control station (C1). A linear increase in chlorophyll *a* was observed in each land use category. This can be attributed to an increase in temperature in the spring and summer and a shift in the community in the fall. In addition, elevated nitrite and nitrate levels were observed in the spring, especially in agricultural stations. Interestingly, the source is likely not nutrient rich runoff from farms that use fertilizer, as fertilizer prices were high in 2023. A spike

in DRP was observed in the summer. Coupled with a drop in nitrite and nitrate suggests a change in the limiting nutrient. In the spring and fall, the reservoir was phosphorus limited, while in the summer, it was nitrogen limited.

Overall, land use categories were different from each other based on physicochemical parameters collected. The main drivers were Secchi depth, extracted chlorophyll *a* and conductivity. Difference within one land use category were also observed. The parameters that drove the difference within one land use category were water temperature, DRP, nitrate and nitrite.

The most abundant phyla throughout the sampling season in each land use category were *Actinobacteria*, *Proteobacteria*, and *Cyanobacteria*. Together, they made up above 75% of the community throughout the year. All phyla are shown to be abundant in other freshwater systems. Dominant genera within the *Actinobacteria* and *Cyanobacteria* phyla are found in turbid lakes, while *Proteobacteria* and *Bacteroidota* respond well to nutrient influxes. These conditions are relevant to Tellico Reservoir, as the retention time is very low, and spikes of nutrients were observed.

The separation between land use categories that was based on physicochemical data was not reflected in the microbial community. When plotting an nMDS based on Bray-Curtis distance between the microbial communities of each land use category, the same trend cannot be observed. However, PERMANOVA results show land use category to be a predictor of the spread observed. Although no strong separation between land use categories was observed in the nMDS, differences in microbial composition were observed. Higher similarities in microbial communities between control and developed stations than control and agricultural stations were observed. This is similar to the trend observed in the physicochemical data.

When highlighting the samples by season in which they were collected in, a strong trend can be observed. Samples cluster strongly with other samples from the same season. Rather a complete takeover from one taxon to another, a shift in species composition from the same phylum caused the separation.

A model was constructed to find the physicochemical parameters that best correlate with the seasonal patterns observed in the nMDS. The best model included four parameters and achieved a correlation of 0.685. The most important parameter was water temperature, as it alone achieved a correlation of ~0.64. As climate change is expected to increase surface water temperatures, the microbial community of Tellico Reservoir potentially can change. This further highlights the need of a microbial baseline that we establish in this study.

Chapter III

**Two genome sequences of *Synechococcus* sp. isolated from Tellico Reservoir,
Tennessee, USA.**

Author contributions:

David Niknejad – writing – original draft/editing/reviewing, sample collection, formal analysis

Robbie Martin – sample collection

Steven Wilhelm – writing – review/editing, conceptualization

Abstract

The Tennessee River Basin is one of the most ecological diverse regions for freshwater biodiversity. In contrast to Animalia, the molecular community is poorly defined. In addition, freshwater supplies are challenged by urban development. Here we present the genome of two freshwater *Synechococcus* sp. strains, a ubiquitous and ecologically important picocyanobacterium, isolated from Tellico Reservoir, Tennessee, USA. Results show two *Synechococcus* sp. MAGs with 73.4% and 98.5% completeness. In a phylogenetic tree, the MAGs closest relative was *Synechococcus* sp. MINOS11 and *Synechococcus* sp. CBW1107, respectively.

Announcement

The Tennessee River basin is one of the hotspots for biodiversity (Jacquemin et al., 2016). It is home to a large amount of different fish, mussels, crayfish, and gastropod species (Boschung and Mayden, 2004; Cavender, 1986; Etnier and Starnes, 1993; Page and Burr, 2011; Parmalee and Hughes, 1993). Established by the Tennessee Valley Authority (TVA), the region is home to 5.2 million residents. In addition, the Tennessee River Basin is currently undergoing urban development, as it will add 650,000 more residents by 2045. Additionally, growth in urban areas around the region—some of which are already facing water-supply challenges—will put increased pressure on the TVA region’s water resources. Some 5 million people get their water from the Tennessee River and its tributaries every day. Over 8 billion gallons of water are used daily, drawn from more than 700 municipal, industrial, agricultural, and thermoelectric intakes. However, over 95 percent of that water is recycled and returned to the river to be used again. The reservoir first gained national attention from the altercation between the state of Tennessee and

the Supreme Court about the removal of the habitat of the then endangered species *Percina tanasi* (Snail Darter). In 2022, long term monitoring of Tellico Reservoir was established.

Here we present two metagenome-assembled genomes (MAGs) of *Synechococcus* sp., isolated from Tellico Reservoir.

Lake water was pushed through a Sterivex™ using a peristaltic pump. It was then frozen in liquid nitrogen and stored at -80°C until further processing. Sterivex™ were opened with pliers and the filter was cut into six pieces and put into a 2 mL Lysing Matrix E tube (MP Biomedicals™). DNA was extracted using modified phenol-chloroform extraction methods (Martin and Wilhelm, 2020). Shotgun libraries were prepared and sequenced by the SeqCenter (Pittsburg, PA).

Default parameters were used unless specified. Forward and reserve reads were concatenated separately and assembled with metaWRAP (v1.3.2) which uses Megahit (v1.1.3). The assembled contigs were binned using the metaWRAP package with the packages metaBAT2 (v2.12.2), MaxBin2 (v2.2.6) and CONCOCT (v1.0.0). Bins were refined using the bin refinement module of the metaWRAP package. Bin quality was assessed using Checkm2 (v1.0.1) and Busco (v5.7.1), as well as Quast (v5.2.0). A phylogenetic tree of all *Synechococcus* genomes available in the NCBI RefSeq database was built using GToTree (v1.8.6) with the two MAGs included.

We recovered two MAGs of *Synechococcus*. Bin5 MAG clustered closely with *Synechococcus* sp. CBW1107, while the bin13 MAG clustered with *Synechococcus* sp. MINOS11. They consisted of 1,586,444 bp (N50=9070 bp) and 2,535,063 bp (N50=92,698 bp), respectively. They were 73.4% and 98.5% complete according to Busco and contained 0.71% and 0.05% contamination according to Checkm2.

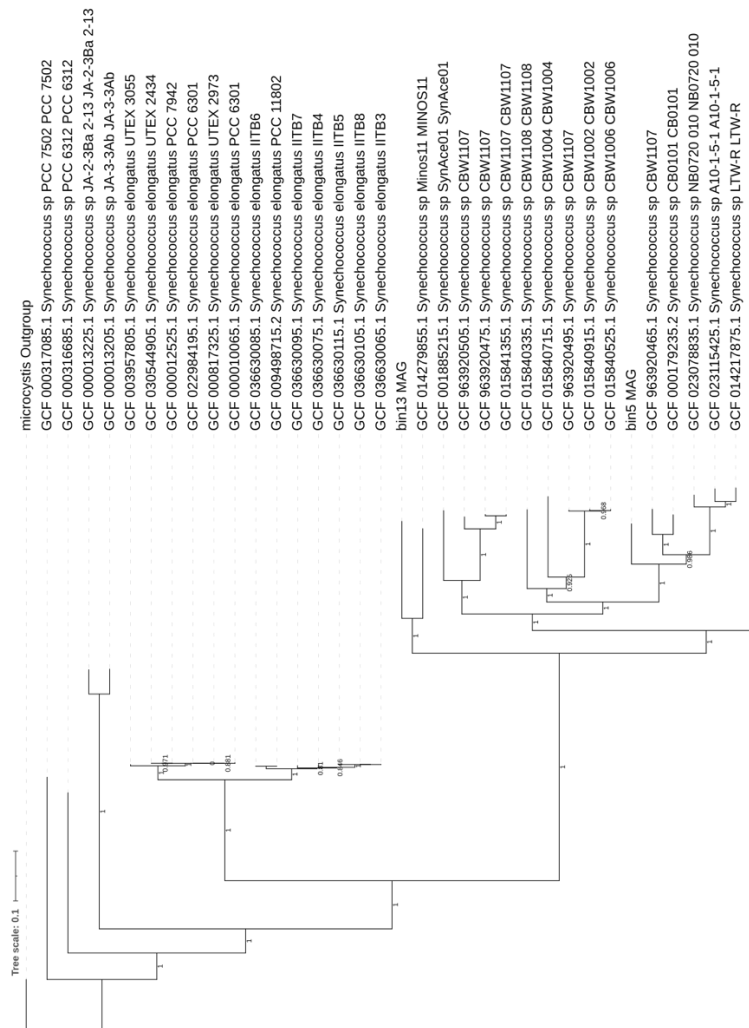


Figure 24: Phylogenetic tree of all publicly available *Synechococcus* genomes from the NCBI RefSeq databased constructed by GToTree (v1.8.6) including the two MAGs presented.

Chapter IV

Conclusion

Tellico Reservoir is one of several reservoirs throughout the Tennessee River Basin. Originally build to prevent flooding, it now sees prominent recreational use. As development increased around the reservoir, water quality concerns arose. Since the Tennessee Valley Authority provides limited testing of water quality, deterioration of water quality is a concern. In addition, analysis of the aquatic microbial community of the reservoir using modern molecular methods is therefore non-existent. With environmental changes associated with climate change set to impact freshwaters, it is important to establish a baseline to which any future changes can be compared.

The reservoir was sampled in 2023, where six stations that are influenced by adjacency to one of three land use categories (agriculture, developed, control). The stations are presumably heavily influenced by the chosen land use category to which they are adjacent. The land use categories were chosen based off the most prominent land use inside and around the watershed. Chlorophyll *a* analyses showed an increase in chlorophyll throughout the year. Agricultural stations were significantly higher in chlorophyll contents than both control and developed stations. The latter two were never different from another. Agriculturally-influenced stations reported increased nitrite and nitrate concentrations during the spring months and ammonium during the fall. A spike in DRP was recorded in all stations during the summer. In other months, nutrient concentrations were similar between the three stations.

Validation of site selection was conducted using a PCA. It showed distinct separation of clusters from the three sites. Control and agriculture stations were widely separated in the PCA, with stations influenced by development interspersed between the two, yet closer to the control

category. Drivers for this separation were mostly biological activity and turbidity of the water. The influx of fertilizer and manure, as well as the influx of Baker Creek into one of the agricultural stations, is one explanation to increased turbidity and nutrient levels at stations associated with agricultural land use. Separation within the stations was caused by nitrate, water temperature and DRP. Spikes in nutrient levels correlated with rainfall, suggesting runoff being the source. Water temperature changes throughout the year, thus being a good predictor of separation between stations. One thing left unexplored, however, is the distance to the main channel. Tellico's low retention time creates a current throughout the reservoir. Both presumably agriculture influenced stations may be influenced stronger influenced by flow rate.

To shed light on the microbial community, the 16S/18S rRNA gene was amplified. The results show little differences in community composition based off relative abundance. Although fluctuations in abundance occur, the main members of the community remain the same throughout the year. The community was largely dominated by *Actinobacteria*, *Cyanobacteria* and *Proteobacteria*. The phyla thriving in the reservoir are tolerant of turbidity and debris inside its habitat. As a constant stream of water flows through Tellico Reservoir, these phyla are well equipped to thrive under the conditions listed. Interestingly, stations with lower light penetration do not exhibit larger abundances of certain microbes, giving need to further study why these microbes are so dominant in the reservoir. Eukaryotes were significantly less abundant than prokaryotes. Similar to the prokaryotes, the community did not change drastically, as the phyla *SAR* and *Cryptophyceae* were the most abundant eukaryotes throughout the year.

Visualizing the species abundance of each stations with an nMDS, a strong seasonal trend can be observed. Physicochemical parameters show a land use category trend, the nMDS does

not. Additionally, analyses that shows low similarity between spring and summer, as well as spring and fall, yet higher similarity between summer and fall.

To understand what environmental variables are driving the difference, a BIOENV was conducted. A better model was created when tracking seasonal separation, than the separation by land use category. In the seasonal model, the most important driver was water temperature. This highlights the importance of water temperature for the reservoir community.

This study is suited as a baseline about the microbial community of Tellico Reservoir. Although imminent land use within a station's watershed influences physicochemical parameters, it influences the microbial community to a lesser extent. It is unknown, however, if the differences between stations truly are rooted in land use or if it is an artifact by distance to the main channel and lesser flow rate. What is apparent though, is the seasonality of the reservoir. Strong seasonal pattern can be observed. Water temperature has the biggest impact on seasonal patterns. With climate change increasing surface water temperature, it is vital to study how the microbial community responds and track long term changes.

References

- Boschung H. T., Mayden R. L. 2004. *Fishes of alabama*, vol. 564. Smithsonian Books Washington, DC.
- Carrick H. J., Schelske C. L. 1997. Have we overlooked the importance of small phytoplankton in productive waters? *Limnology and Oceanography*, 42:1613-1621.
- Cavender T. 1986. Review of the fossil history of North American freshwater fishes. *The zoogeography of North American freshwater fishes.*:699-724.
- Chapman J. 1985. Tellico archaeology: 12,000 years of Native American history. *(No Title)*.
- Che R., Bai M., Xiao W., Zhang S., Wang Y., Cui X. 2022. Nutrient levels and prokaryotes affect viral communities in plateau lakes. *Science of the Total Environment*, 839:156033.
- Chen H., Bian Z., Ravichandran V., Li R., Sun Y., Huo L., Fu J., Bian X., Xia L., Tu Q. 2019. Biosynthesis of polyketides by trans-AT polyketide synthases in Burkholderiales. *Critical reviews in microbiology*, 45:162-181.
- Clarke K. R., Ainsworth M. 1993. A method of linking multivariate community structure to environmental variables. *MARINE ECOLOGY PROGRESS SERIES*, Vol. 92: 205-219.
- Cruaud P., Vigneron A., Fradette M. S., Dorea C. C., Culley A. I., Rodriguez M. J., Charette S. J. 2020. Annual bacterial community cycle in a seasonally ice-covered river reflects environmental and climatic conditions. *Limnology and Oceanography*, 65:S21-S37.
- Davis R. K. 1979. Tellico Dam and Reservoir: Staff Report to the Endangered Species Committee.
- Eiler A., Bertilsson S. 2004. Composition of freshwater bacterial communities associated with cyanobacterial blooms in four Swedish lakes. *Environmental microbiology*, 6:1228-1243.
- Eiler A., Olsson J. A., Bertilsson S. 2006. Diurnal variations in the auto-and heterotrophic activity of cyanobacterial phycospheres (*Gloeotrichia echinulata*) and the identity of attached bacteria. *Freshwater Biology*, 51:298-311.
- Etnier D. A., Starnes W. C. 1993. The fishes of Tennessee.
- Fahnenstiel G. L., Carrick H. J., Iturriaga R. 1991. Physiological characteristics and food-web dynamics of *Synechococcus* in Lakes Huron and Michigan. *Limnology and Oceanography*, 36:219-234.
- Felföldi T., Somogyi B., Marialigeti K., Vörös L. 2009. Characterization of photoautotrophic picoplankton assemblages in turbid, alkaline lakes of the Carpathian Basin (Central Europe). *Journal of Limnology*, 68:385.
- Fuchich D., Chen F. 2020. Presence of toxin-antitoxin systems in picocyanobacteria and their ecological implications. *The ISME Journal*, 14:2843-2850.
- Gilmer R. A. 2013. Snail darters and sacred places: Creative application of the Endangered Species Act. *Environmental management*, 52:1046-1056.
- Gleick P. H., Cooley H. 2021. Freshwater scarcity. *Annual Review of Environment and Resources*, 46:319-348.
- Glöckner F. O., Zaichikov E., Belkova N., Denissova L., Pernthaler J., Pernthaler A., Amann R. 2000. Comparative 16S rRNA analysis of lake bacterioplankton reveals globally distributed phylogenetic clusters including an abundant group of actinobacteria. *Applied and Environmental Microbiology*, 66:5053-5065.

- Gołębiewski M., Całkiewicz J., Creer S., Piwosz K. 2017. Tideless estuaries in brackish seas as possible freshwater-marine transition zones for bacteria: the case study of the Vistula river estuary. ***Environmental Microbiology Reports***, 9:129-143.
- Guidi L., Chaffron S., Bittner L., Eveillard D., Larhlimi A., Roux S., Darzi Y., Audic S., Berline L., Brum J. R. 2016. Plankton networks driving carbon export in the oligotrophic ocean. ***Nature***, 532:465-470.
- He Q., Xiao Q., Fan J., Zhao H., Cao M., Zhang C., Jiang Y. 2022. The impact of heterotrophic bacteria on recalcitrant dissolved organic carbon formation in a typical karstic river. ***Science of the Total Environment***, 815:152576.
- Hickman G. D., Fitz R. B. 1978. ***A Report on the Ecology and Conservation of the Snail Darter (Percina Tanasi Etnier), 1975-1977***. Tennessee Valley Authority, Division of Forestry, Fisheries, and Wildlife
- Hornák K., Jezbera J., Nedoma J., Gasol J. M., Simek K. 2006. Effects of resource availability and bacterivory on leucine incorporation in different groups of freshwater bacterioplankton, assessed using microautoradiography. ***Aquatic microbial ecology***, 45:277-289.
- Indian Affairs. 1974. Origin of Names of US States. 6-6-24, 2024.
- Jacquemin S. J., Ebersole J. A., Dickinson W. C., Ciampaglio C. N. 2016. Late Pleistocene fishes of the Tennessee River Basin: an analysis of a late Pleistocene freshwater fish fauna from Bell Cave (site ACb-2) in Colbert County, Alabama, USA. ***PeerJ***, 4:e1648.
- Jardillier L., Zubkov M. V., Pearman J., Scanlan D. J. 2010. Significant CO₂ fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. ***The ISME Journal***, 4:1180-1192.
- Kirchman D. L. 2002. The ecology of Cytophaga–Flavobacteria in aquatic environments. ***FEMS Microbiology Ecology***, 39:91-100.
- Lacher I., Akre T., Mcshea W. J., Fergus C. 2019. Spatial and temporal patterns of public and private land protection within the Blue Ridge and Piedmont ecoregions of the eastern US. ***Landscape and urban planning***, 186:91-102.
- Lackey J. B. 1942. The plankton algae and protozoa of two Tennessee rivers. ***The American Midland Naturalist***, 27:191-202.
- Lefler F. W., Barbosa M., Berthold D. E., Zimba P. V., Laughinghouse IV D. 2022. Diversity of cyanobacteria and associated microbial community structure within Lake Okeechobee and the St. Lucie River, Florida, USA.
- Martin R. M., Wilhelm S. W. 2020. Phenol-based RNA extraction from polycarbonate filters. March 30, 2023. [dx.doi.org/10.17504/protocols.io.bivuke6w](https://doi.org/10.17504/protocols.io.bivuke6w).
- Nazari-Sharabian M., Ahmad S., Karakouzian M. 2018. Climate change and eutrophication: a short review. ***Engineering, Technology and Applied Science Research***, 8:3668.
- Newton R. J., Jones S. E., Eiler A., McMahon K. D., Bertilsson S. 2011. A guide to the natural history of freshwater lake bacteria. ***Microbiology and molecular biology reviews***, 75:14-49.
- Newton R. J., McLellan S. L. 2015. A unique assemblage of cosmopolitan freshwater bacteria and higher community diversity differentiate an urbanized estuary from oligotrophic Lake Michigan. ***Frontiers in Microbiology***, 6:158642.
- O’Neil J. M., Davis T. W., Burford M. A., Gobler C. J. 2012. The rise of harmful cyanobacteria blooms: the potential roles of eutrophication and climate change. ***Harmful algae***, 14:313-334.

- Ossyssek S., Geist J., Werner P., Raeder U. 2020. Identification of the ecological preferences of *Cyclotella comensis* in mountain lakes of the northern European Alps. **Arctic, Antarctic, and Alpine Research**, 52:512-523.
- Paerl H. W., Huisman J. 2008. Blooms Like It Hot. **Science**, 320:57-58.
- Page L. M., Burr B. M. 2011. **Peterson field guide to freshwater fishes of North America north of Mexico**. Houghton Mifflin Harcourt.
- Palenik B., Ren Q., Dupont C. L., Myers G. S., Heidelberg J. F., Badger J. H., Madupu R., Nelson W. C., Brinkac L. M., Dodson R. J. 2006. Genome sequence of *Synechococcus* CC9311: insights into adaptation to a coastal environment. **Proceedings of the National Academy of Sciences**, 103:13555-13559.
- Pantoja-Agreda F., Pajares S. 2024. Occurrence and diversity of bacterioplankton in drinking water tropical reservoirs of contrasting trophic state. **Aquatic Ecology**:1-16.
- Parmalee P. W., Hughes M. H. 1993. Freshwater mussels (Mollusca: Pelecypoda: Unionidae) of Tellico Lake: twelve years after impoundment of the Little Tennessee River. **Annals of the Carnegie Museum**, 62:81-93.
- Partensky F., Blanchot J., Vaulot D. 1999. Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. **Bulletin-Institut Oceanographique Monaco-Numero Special**:-457-476.
- Pernthaler J., Posch T., Šimek K., Vrba J., Pernthaler A., Glöckner F. O., Nübel U., Psenner R., Amann R. 2001. Predator-specific enrichment of actinobacteria from a cosmopolitan freshwater clade in mixed continuous culture. **Applied and Environmental Microbiology**, 67:2145-2155.
- Pittera J., Jouhet J., Breton S., Garczarek L., Partensky F., Maréchal É., Nguyen N. A., Doré H., Ratin M., Pitt F. D. 2018. Thermoacclimation and genome adaptation of the membrane lipidome in marine *Synechococcus*. **Environmental microbiology**, 20:612-631.
- Plater Z. J. 2013. **The snail darter and the dam: How pork-barrel politics endangered a little fish and killed a river**. Yale University Press.
- Poulton N. J. 2016. FlowCam: quantification and classification of phytoplankton by imaging flow cytometry. **Imaging flow cytometry: methods and protocols**:237-247.
- Rawls Jr. W. 1979. Forgotten People of the Tellico Dam Fight, New York Times. The New York Times Company.
- Sagona F. J. 1983. Tellico Reservoir postimpoundment water quality.
- Sanchez-Baracaldo P., Hayes P., BLANK C. E. 2005. Morphological and habitat evolution in the Cyanobacteria using a compartmentalization approach. **Geobiology**, 3:145-165.
- Sharma A. K., Sommerfeld K., Bullerjahn G. S., Matteson A. R., Wilhelm S. W., Jezbera J., Brandt U., Doolittle W. F., Hahn M. W. 2009. Actinorhodopsin genes discovered in diverse freshwater habitats and among cultivated freshwater Actinobacteria. **The ISME Journal**, 3:726-737.
- Shollenberger K. R., Janosik A. M., Johnston C. 2023. Detection of the threatened snail darter *Percina tanasi* in the Tennessee River system using environmental DNA. **Journal of fish biology**, 102:373-379.
- Šimek K., Horňák K., Jezbera J., Nedoma J., Vrba J., Straškrábová V., Macek M., Dolan J. R., Hahn M. W. 2006. Maximum growth rates and possible life strategies of different bacterioplankton groups in relation to phosphorus availability in a freshwater reservoir. **Environmental microbiology**, 8:1613-1624.

- Šimek K., Kasalický V. c., Jezbera J., Jezberová J., Hejzlar J., Hahn M. W. 2010. Broad habitat range of the phylogenetically narrow R-BT065 cluster, representing a core group of the betaproteobacterial genus *Limnohabitans*. ***Applied and Environmental Microbiology***, 76:631-639.
- Starnes W. C. 1977. The ecology and life history of the endangered snail darter, *Percina (Imostoma) tanasi* Etnier.
- Trujillo M. E., Normand P. Sporichthyaceae, p. 1-2, ***Bergey's Manual of Systematics of Archaea and Bacteria***.
- TVA. Managing the Tennessee River for Multiple Benefits. 5/24/24, 2024.
- TVA. Managing Water Supply. May 28, 2024.
- TVA. Tellico Reservoir: Reservoir Health Ratings. May 28, 2024.
- TVA. Telling the Story of Tellico: It's Complicated. March 2024, 2024.
- TVA. Valley Rainfall. May 28, 2024.
- TVA v. Hill. 1978. p. 153, US, vol. 437. Supreme Court.
- U.S. Census Bureau. n.d. QuickFacts: Tellico Village CDP, Tennessee. U.S. Department of Commerce. Accessed
- Waldrop W. R., Ungate C. D., Harper W. L. 1980. ***Computer simulation of hydrodynamics and temperatures of Tellico Reservoir***. Tennessee Valley Authority, Office of National Resources, Division of Water
- Wang T., Li J., Jing H., Qin S. 2022. Picocyanobacterial *Synechococcus* in marine ecosystem: insights from genetic diversity, global distribution, and potential function. ***Marine Environmental Research***, 177:105622.
- Wetzel R. G., Likens G. 2000. ***Limnological analyses***. Springer Science & Business Media.
- Wheeler W. B., McDonald M. J. 1986. TVA and the Tellico Dam, 1936-1979: A bureaucratic crisis in post-industrial America. ***(No Title)***.
- Wilhelm S. W. 2017. Total Chlorophyll a Measurements by Spectrophotometer. June 4, 2024. [dx.doi.org/10.17504/protocols.io.ijpccmn](https://doi.org/10.17504/protocols.io.ijpccmn).
- Xylem. 2020. Exo User Manual. July 19, 2024.
- Yanez-Montalvo A., Aguila B., Gómez-Acata E. S., Guerrero-Jacinto M., Oseguera L. A., Falcón L. I., Alcocer J. 2022. Shifts in water column microbial composition associated to lakes with different trophic conditions: “Lagunas de Montebello” National Park, Chiapas, México. ***PeerJ***, 10:e13999.
- Zeder M., Peter S., Shabarova T., Pernthaler J. 2009. A small population of planktonic Flavobacteria with disproportionally high growth during the spring phytoplankton bloom in a prealpine lake. ***Environmental microbiology***, 11:2676-2686.
- Zhang R., Liu W.-c., Liu Y., Zhang H.-l., Zhao Z.-h., Zou L.-y., Shen Y.-c., Lan W.-S. 2021. Impacts of anthropogenic disturbances on microbial community of coastal waters in Shenzhen, South China. ***Ecotoxicology***, 30:1652-1661.
- Zufiaurre A., Felip M., Camarero L., Sala-Faig M., Juhanson J., Bonilla-Rosso G., Hallin S., Catalan J. 2022. Bacterioplankton seasonality in deep high-mountain lakes. ***Frontiers in Microbiology***, 13:935378.

Appendix

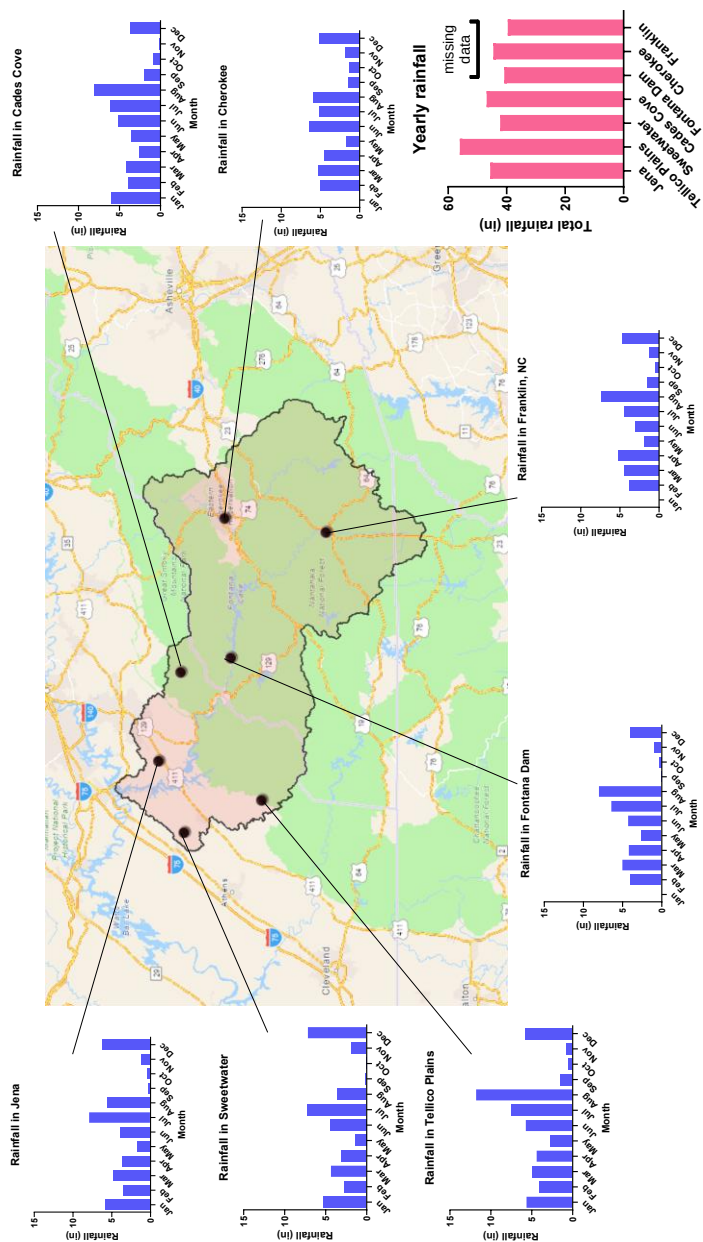


Figure 25: Rainfall at each TVA operated rainfall gauge inside the watershed of Tellico Reservoir.

Table 15: Tracking of 18S reads at certain steps of the DADA2 pipeline

Station						
name	dada2_input	filtered	denoised	merged	no_chimeras	perc_reads_survived
1Ag	282	235	198	181	167	59.2
1D	918	801	669	650	648	70.6
1C	5693	4907	4723	4675	4648	81.6
1D	4204	3644	3476	3441	3392	80.7
1Ag	412	351	274	272	271	65.8
1C	2684	2353	2208	2183	2170	80.8
2Ag	3513	3079	2975	2953	2869	81.7
2D	1021	870	736	714	699	68.5
2C	2251	1928	1817	1755	1724	76.6
2D	1287	1117	1007	967	953	74
2Ag	3118	2711	2575	2549	2523	80.9
2C	824	717	627	613	602	73.1
3Ag	4032	3562	3417	3373	3313	82.2
3D	1697	1492	1372	1321	1282	75.5
3C	764	659	556	529	474	62
3D	1501	1314	1201	1183	1155	76.9
3Ag	865	762	669	640	611	70.6
3C	1010	898	775	759	745	73.8
4Ag	2136	1855	1697	1673	1614	75.6
4D	1475	1275	1168	1122	1077	73

4C	2122	1850	1737	1656	1609	75.8
----	------	------	------	------	------	------

Table 15 continued

4D	3085	2705	2529	2470	2392	77.5
4Ag	2703	2329	2187	2123	2027	75
4C	408	362	301	288	278	68.1
5Ag	2896	2545	2351	2328	2279	78.7
5D	2643	2320	2121	2081	2026	76.7
5C	1047	899	791	760	751	71.7
5D	946	833	680	652	631	66.7
5Ag	470	378	246	231	230	48.9
5C	2233	1956	1834	1783	1736	77.7
6Ag	1541	1344	1201	1168	1138	73.8
6D	740	653	558	508	494	66.8
6C	1434	1212	1026	992	927	64.6
6D	1912	1690	1508	1446	1372	71.8
6Ag	2932	2560	2332	2272	2188	74.6
6C	865	755	650	639	605	69.9
7Ag	1037	912	783	729	689	66.4
7D	3228	2832	2594	2537	2496	77.3
7C	881	783	659	636	628	71.3
7D	1885	1615	1416	1356	1312	69.6
7Ag	2050	1797	1593	1561	1515	73.9
7C	591	513	407	372	349	59.1

8Ag	202	174	127	119	119	58.9
8D	3892	3351	3155	3117	3060	78.6

Table 15 continued

8C	412	362	264	255	240	58.3
8D	3997	3445	3227	3159	3050	76.3
8Ag	4290	3738	3542	3502	3412	79.5
8C	1914	1630	1470	1450	1309	68.4
9Ag	1539	1350	1213	1174	1152	74.9
9D	1645	1466	1290	1255	1205	73.3
9C	3242	2761	2588	2532	2519	77.7
9D	3068	2697	2514	2478	2413	78.7
9Ag	1381	1184	1044	1015	1003	72.6
9C	4216	3630	3384	3304	3210	76.1
10Ag	1841	1597	1410	1380	1363	74
10D	1215	1021	882	835	770	63.4
10C	5746	5015	4765	4660	4582	79.7
10D	3915	3459	3228	3139	3097	79.1
10Ag	3338	2958	2758	2699	2654	79.5
10C	4617	4023	3810	3755	3718	80.5

Table 16: Tracking of 16S reads at certain steps of the DADA2 pipeline

Station						
name	dada2_input	filtered	denoised	merged	no_chimeras	perc_reads_survived
1Ag	10381	9210	8804	7996	7980	76.9
1D	19940	17862	17276	15697	15647	78.5
1C	76482	67449	66024	60688	59961	78.4
1D	78485	69799	68410	63052	62360	79.5
1Ag	13219	11807	11349	10186	10138	76.7
1C	49816	43848	42909	39653	39426	79.1
2Ag	123488	109551	107917	101881	100994	81.8
2D	62506	54092	53019	49269	48820	78.1
2C	104750	93764	92409	87022	86149	82.2
2D	90146	80307	78985	73911	73094	81.1
2Ag	103096	91966	90425	85962	84939	82.4
2C	30909	27605	27027	25399	25269	81.8
3Ag	168523	150416	148002	138868	135852	80.6
3D	114731	102413	100474	91944	89125	77.7
3C	55221	49175	47821	42678	41665	75.5
3D	96396	86018	84144	76355	73609	76.4
3Ag	40341	36145	35004	31295	30545	75.7
3C	87290	77953	76866	71288	70035	80.2
4Ag	111372	98072	96168	88723	85879	77.1
4D	85842	76031	74384	67394	65702	76.5

4C	142700	127528	125099	113912	109971	77.1
----	--------	--------	--------	--------	--------	------

Table 16 continued

4D	174944	155509	152582	139966	134724	77
4Ag	129902	115713	113575	103925	100175	77.1
4C	35526	31638	30830	27697	27262	76.7
5Ag	171060	152244	149989	142003	139900	81.8
5D	167630	148291	145605	134726	131764	78.6
5C	76579	67208	65696	60658	59723	78
5D	83892	74538	72483	66190	64503	76.9
5Ag	25049	20553	19533	17234	16914	67.5
5C	161238	142587	140781	133244	131247	81.4
6Ag	100736	87955	86136	79850	78649	78.1
6D	39105	35086	33877	30631	30236	77.3
6C	91364	80197	78218	71995	71119	77.8
6D	130132	116179	113564	103840	101680	78.1
6Ag	195762	173033	170237	156728	151522	77.4
6C	102294	91410	89876	82922	81550	79.7
7Ag	57075	50879	49725	45476	44826	78.5
7D	179995	158628	156052	142807	140321	78
7C	38737	34530	33407	29701	29350	75.8
7D	162974	139151	136193	121933	117539	72.1
7Ag	126807	112222	109839	98767	94403	74.4
7C	133207	116812	114924	106203	102541	77

8Ag	52142	46237	45241	41796	41217	79
8D	135768	119823	117451	107452	106138	78.2

Table 16 continued

8C	16696	14891	14028	12501	12318	73.8
8D	167839	146990	143933	131472	129449	77.1
8Ag	166338	147985	145083	131764	129263	77.7
8C	105413	93039	91370	83831	81643	77.5
9Ag	119241	103743	101892	94754	93457	78.4
9D	62792	55668	54393	49702	49340	78.6
9C	131153	114419	112336	103493	102758	78.3
9D	208109	185117	183161	173242	171455	82.4
9Ag	59794	52899	51565	47139	46784	78.2
9C	158640	139557	136922	126944	125276	79
10Ag	28221	25168	24056	20892	20753	73.5
10D	44589	38383	37119	33585	33348	74.8
10C	216522	192926	189570	175434	174074	80.4
10D	162555	144970	142290	131601	130633	80.4
10Ag	86083	77094	75198	68801	68104	79.1
10C	154970	138441	136524	128716	127602	82.3

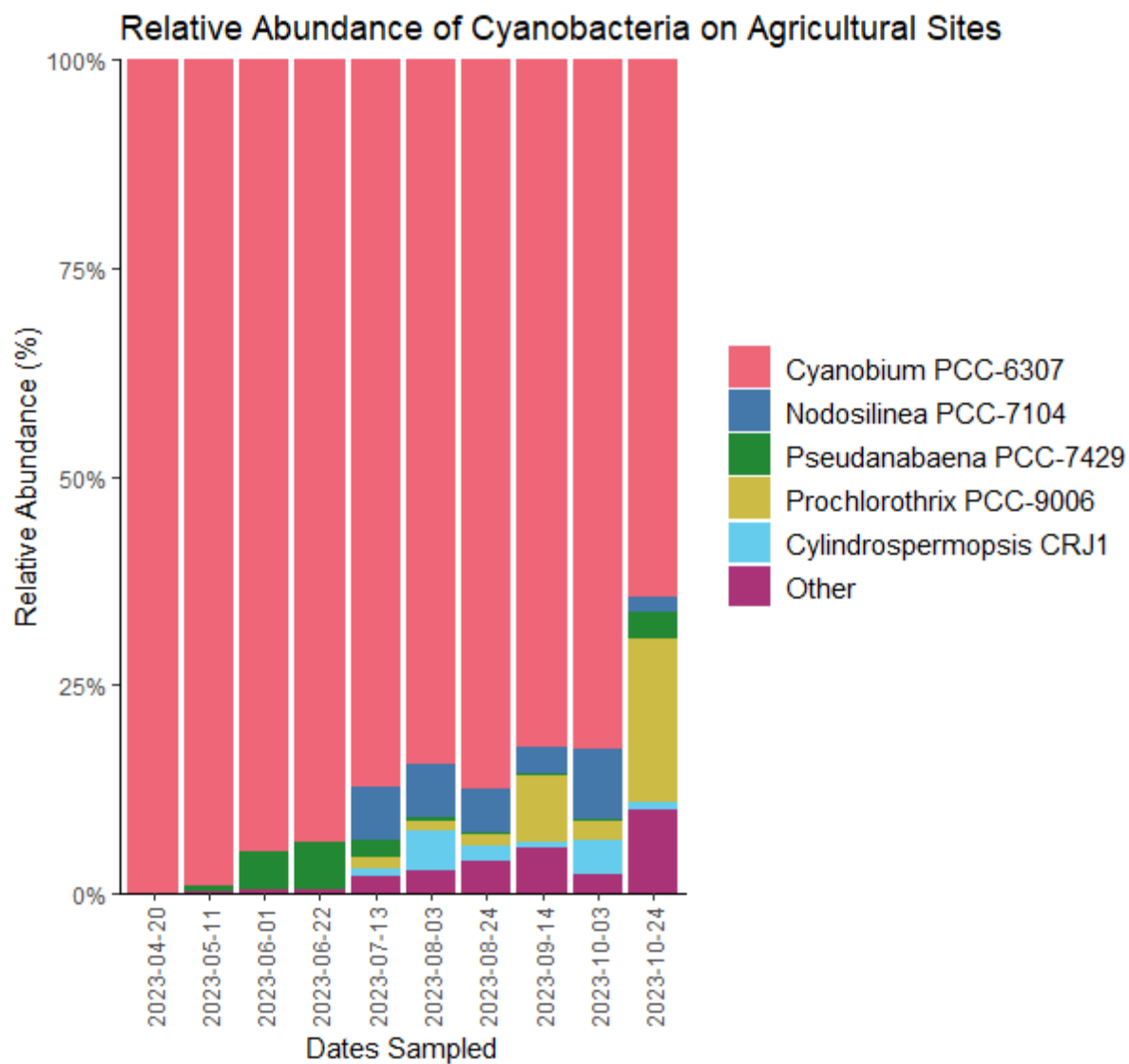


Figure 26: Relative abundance of genera inside the Cyanobacteria phylum of sites presumably influenced by agricultural land use. The label other is made out of each genera that made up less than 5% of the total community.

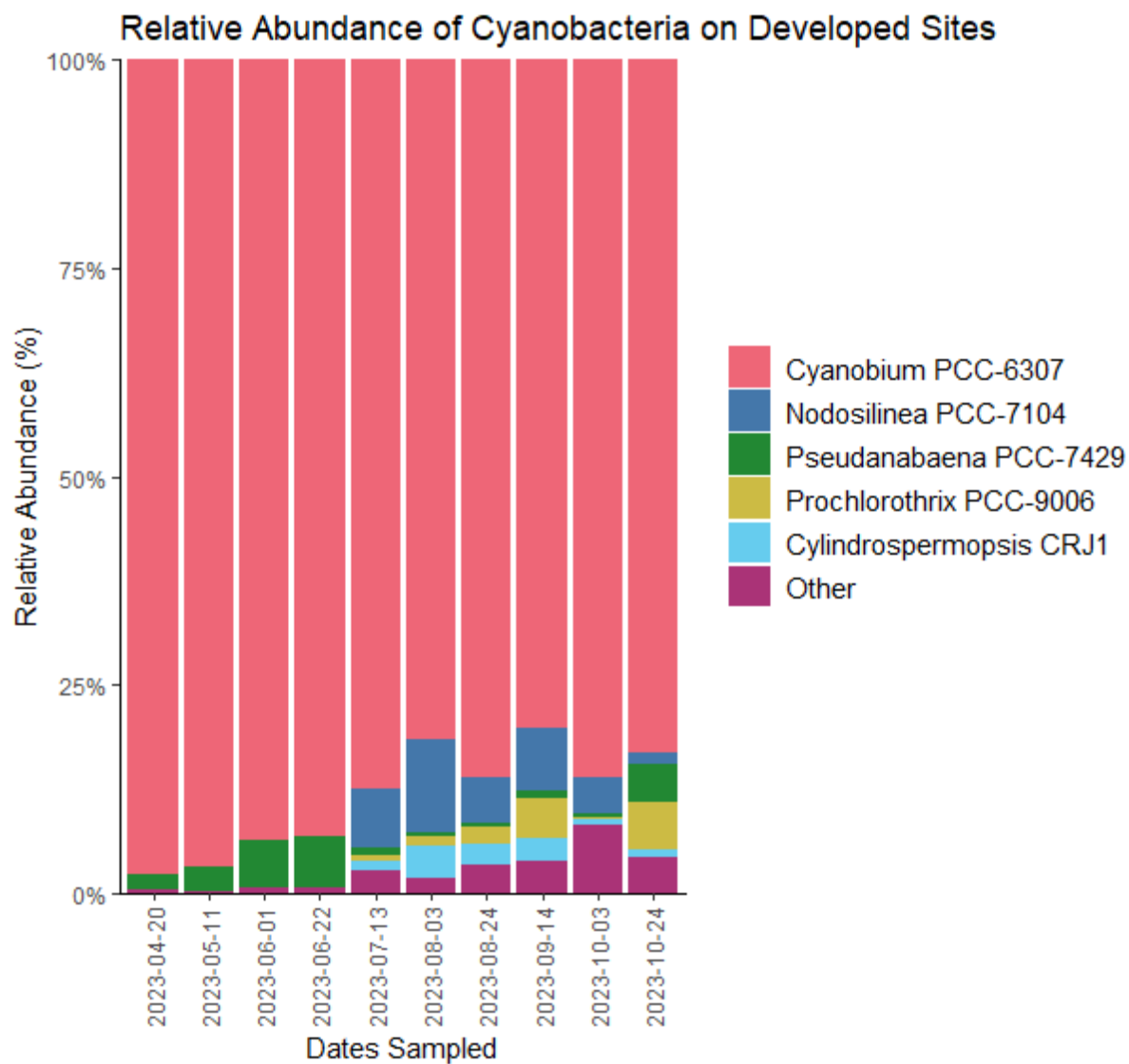


Figure 27: Relative abundance of genera inside the Cyanobacteria phylum of sites presumably influenced by developed land use. The label other is made out of each genera that made up less than 5% of the total community.

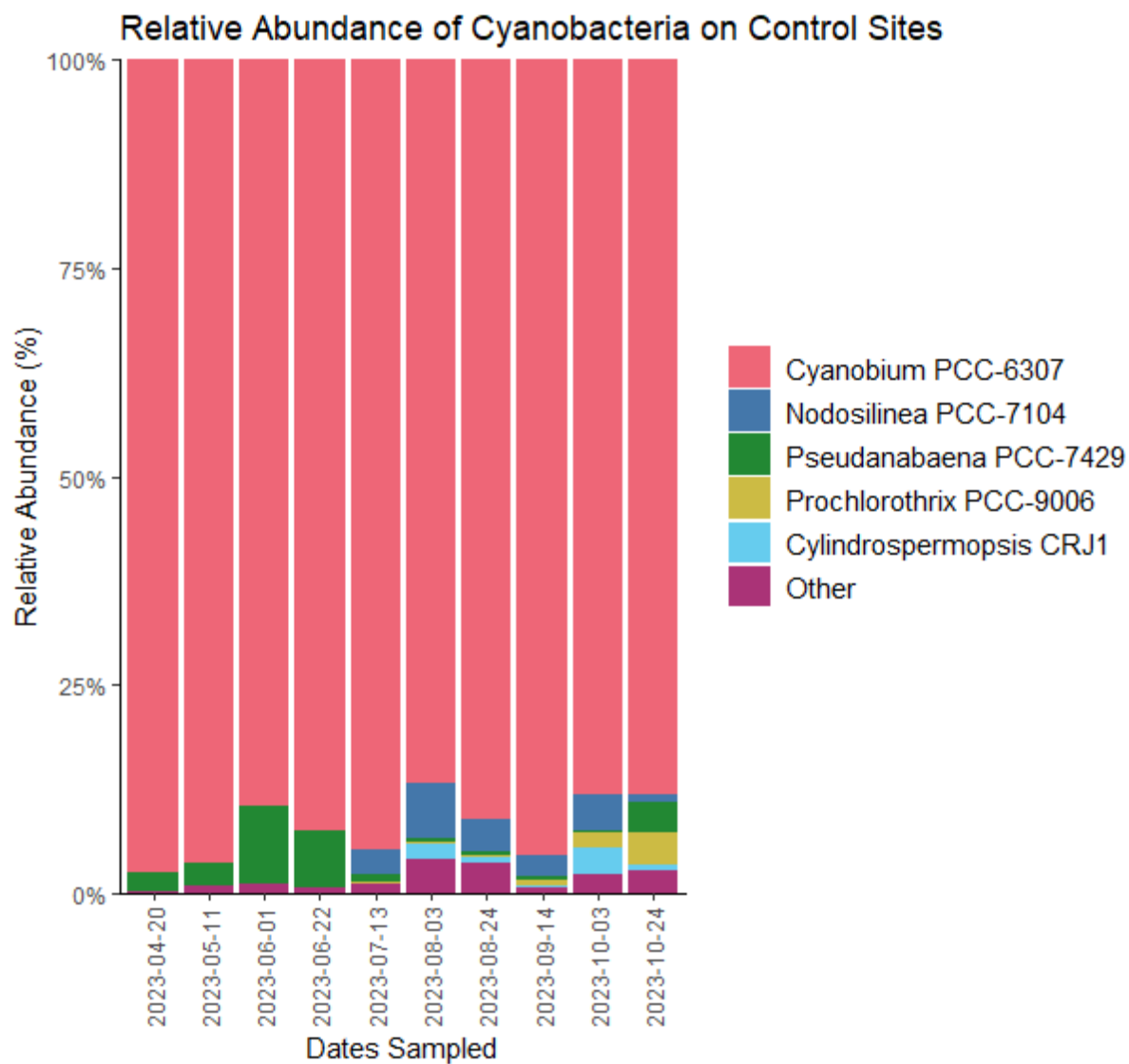


Figure 28: Relative abundance of genera inside the Cyanobacteria phylum of control sites. The label other is made out of each genera that made up less than 5% of the total community.

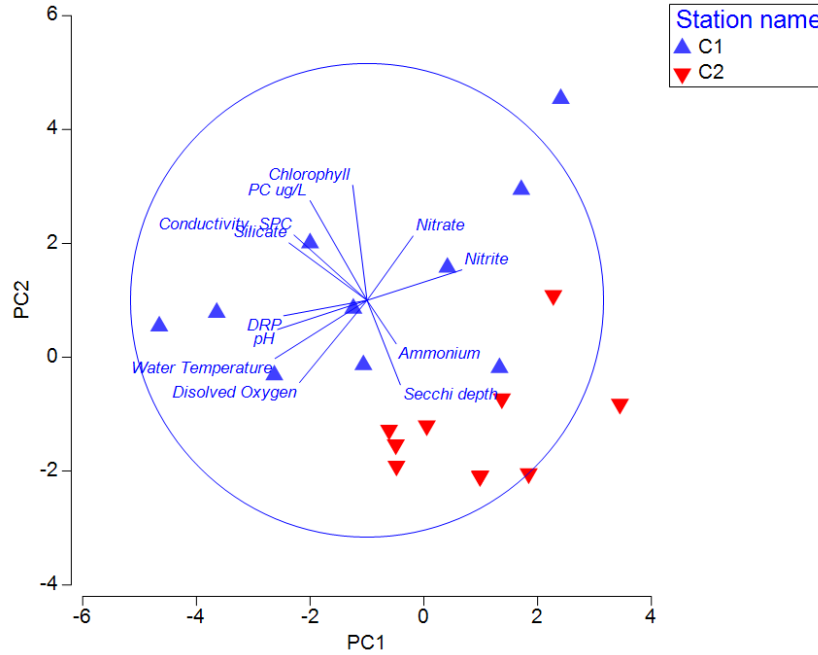


Figure 29: Principle component analysis (PCA) of environmental data collected from stations C1 and C2. Data points represent a single sample, colored by land use type. Variables with longer vectors have a more significant influence on the corresponding principal component. Variables include Secchi depth (cm), water temperature ($^{\circ}\text{C}$), dissolved reactive phosphorus (DRP, $\mu\text{mol/L}$), dissolved oxygen, pH, Phycocyanin ($\mu\text{g/L}$), extracted Chlorophyll *a* ($\mu\text{mol/L}$), silicate ($\mu\text{mol/L}$), conductivity (SPC), nitrate ($\mu\text{mol/L}$), and nitrite ($\mu\text{mol/L}$).

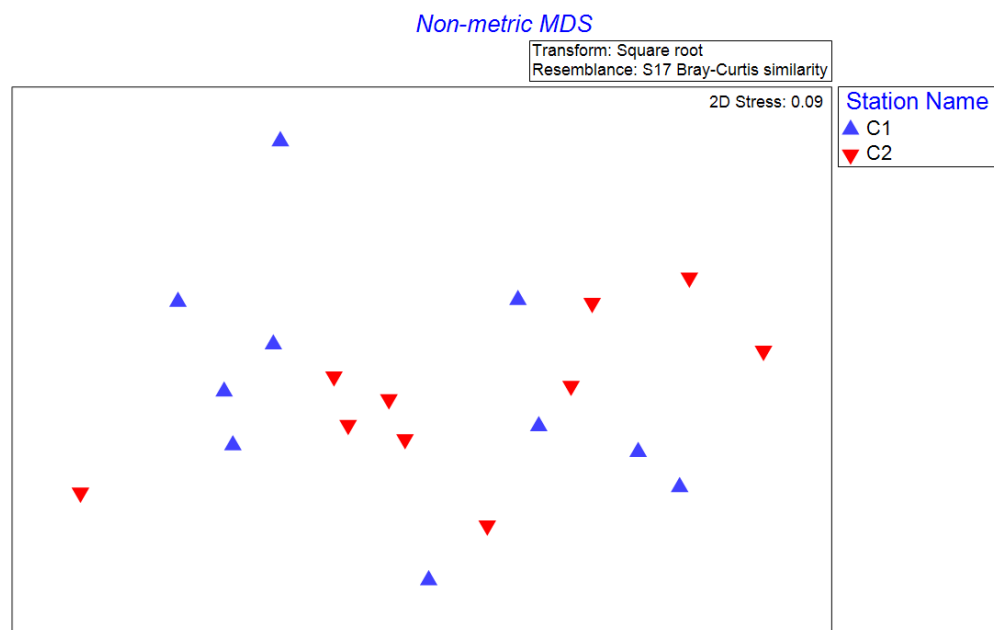


Figure 30: nMDS of Bray-Curtis distance between microbial communities from stations C1 and C2. Different stations were color coded.

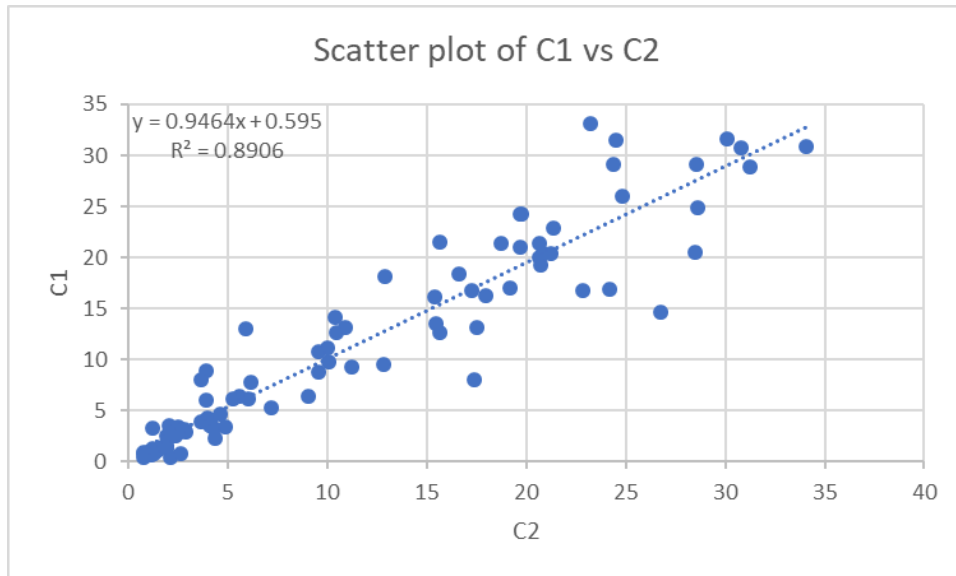


Figure 31: Correlation between the microbial communities of C1 and C2.

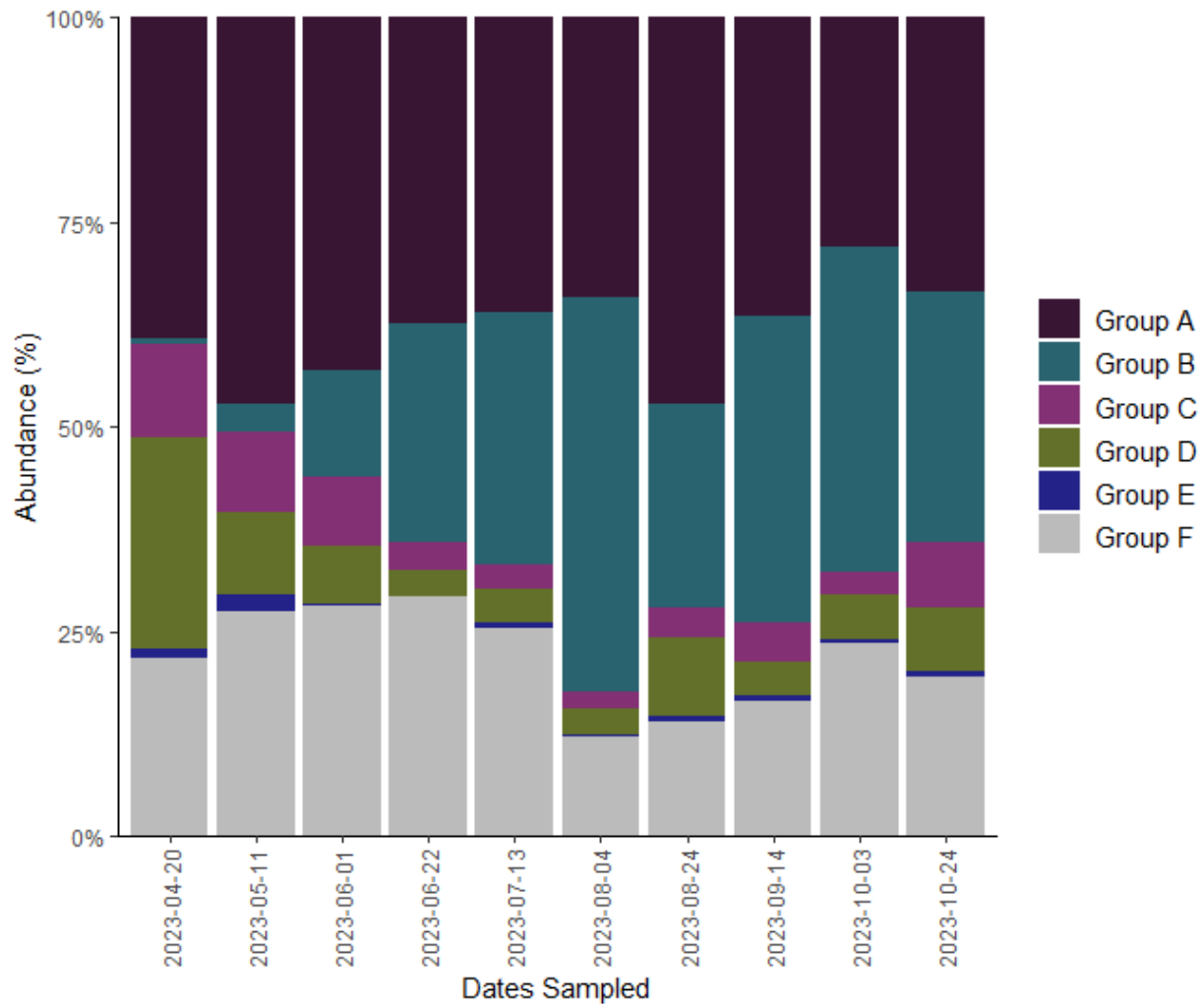


Figure 32: Abundance of microbes captured by the FlowCAM at station Ag1. Bar plot showing the abundance of microbes, 15 μm or larger, from station Ag1 throughout the year. Group A includes any image that was translucent, green, and often formed colonies. Group B were filaments or microbes that were straight exterior. Group C microbes that formed tighter colonies or microbes that were darker and less translucent. Group D were round microbes. Group E were large microbes. Group F were blurry or unidentifiable images.

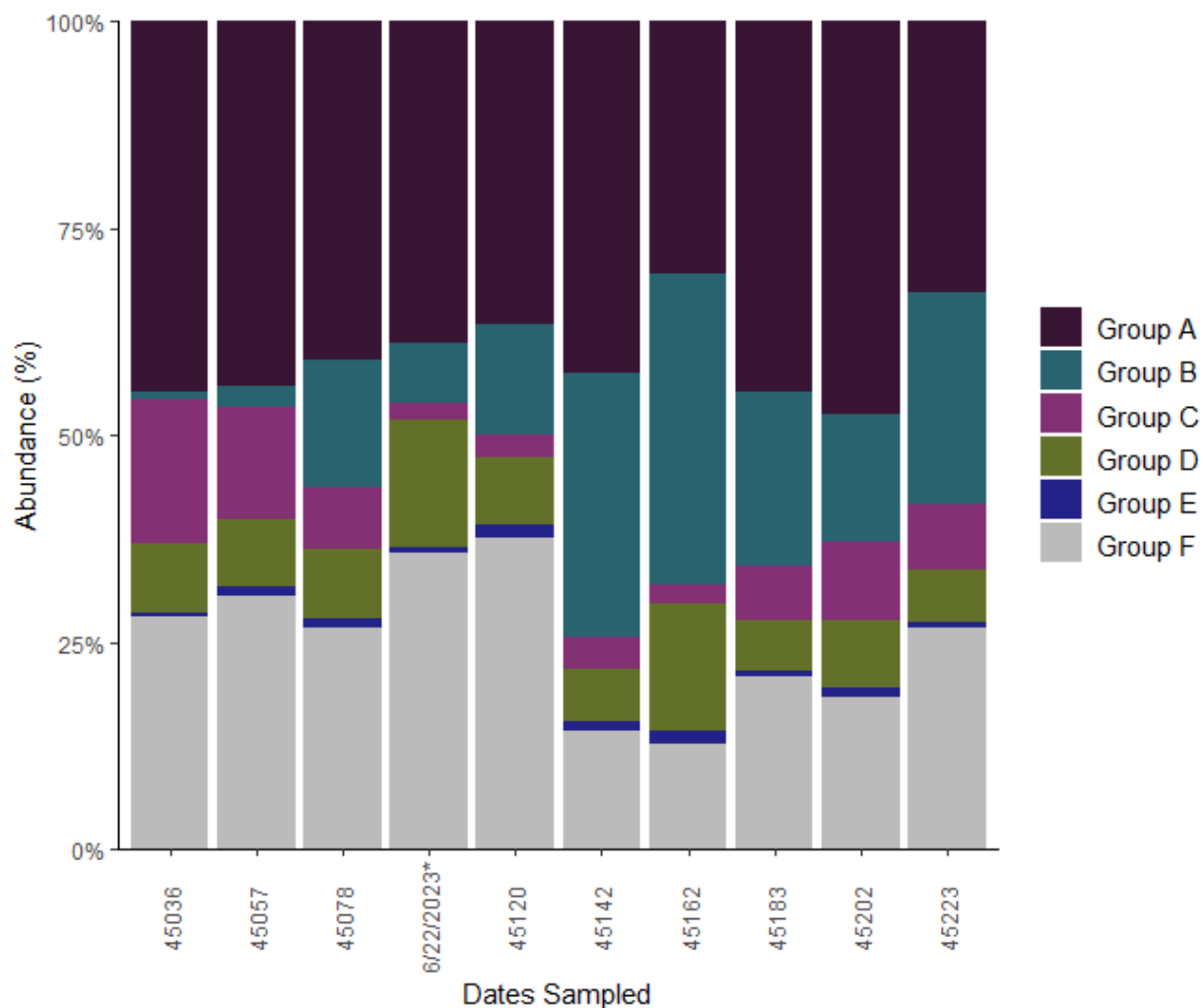


Figure 33: Abundance of microbes captured by the FlowCAM at station Ag2. Bar plot showing the abundance of microbes, 15 μm or larger, from station Ag2 throughout the year. Sample date 6/22/23 captures microbes 10 μm or larger. Group A includes any image that was translucent, green, and often formed colonies. Group B were filaments or microbes that were straight exterior. Group C microbes that formed tighter colonies or microbes that were darker and less translucent. Group D were round microbes. Group E were large microbes. Group F were blurry or unidentifiable images.

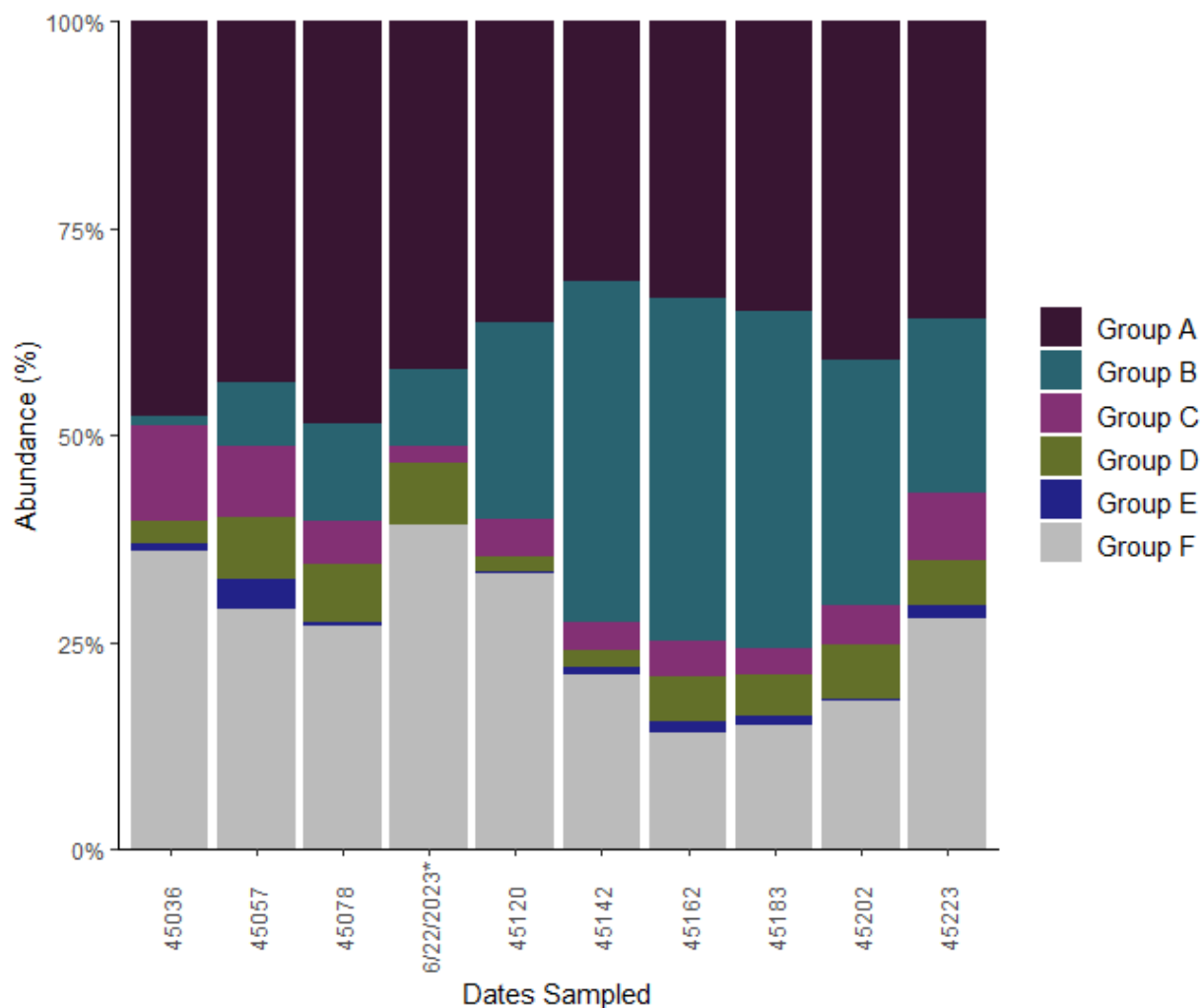


Figure 34: Abundance of microbes captured by the FlowCAM at station D1. Bar plot showing the abundance of microbes, 15 μm or larger, from station D1 throughout the year. Sample date 6/22/23 captures microbes 10 μm or larger. Group A includes any image that was translucent, green, and often formed colonies. Group B were filaments or microbes that were straight exterior. Group C microbes that formed tighter colonies or microbes that were darker and less translucent. Group D were round microbes. Group E were large microbes. Group F were blurry or unidentifiable images.

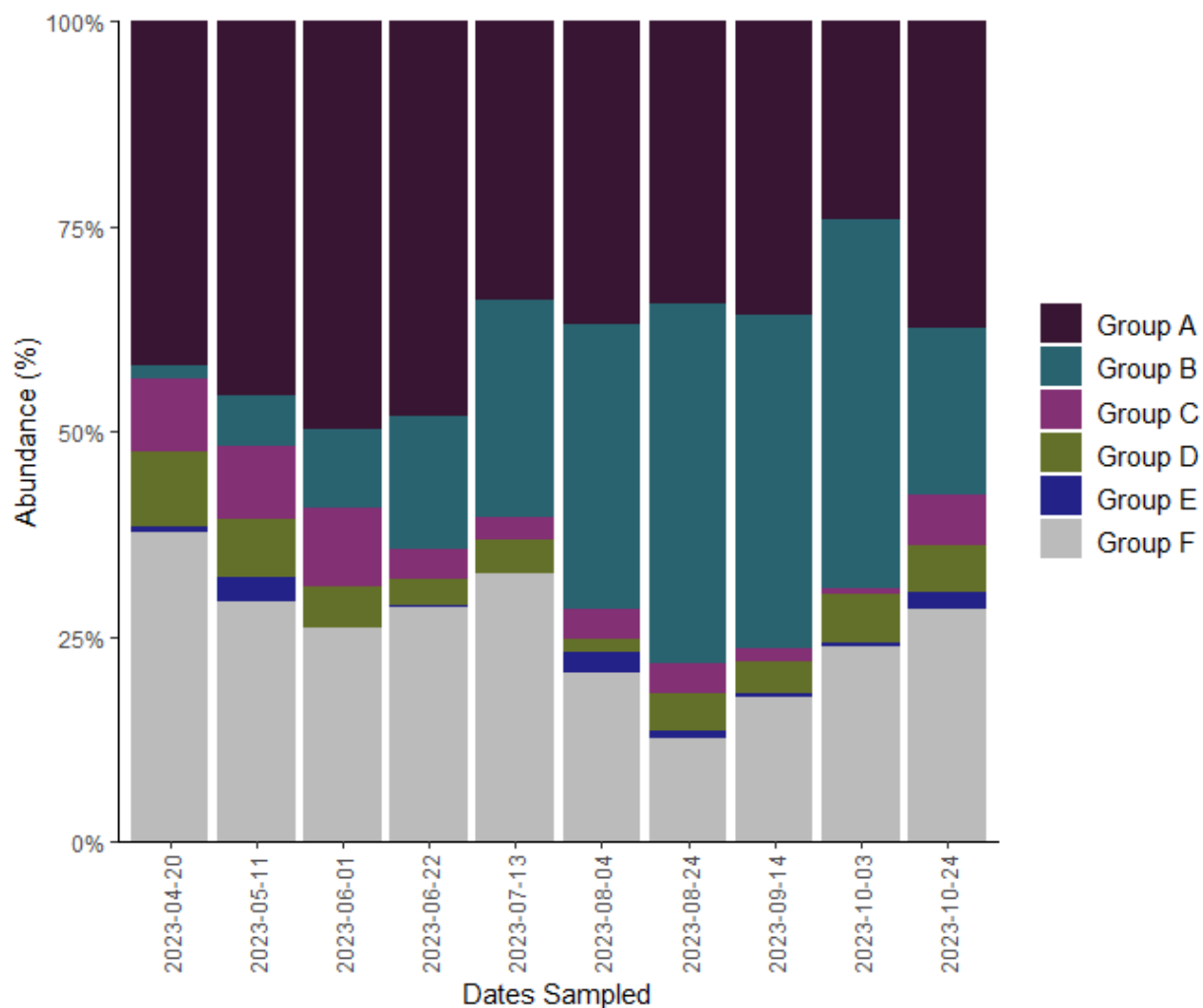


Figure 35: Abundance of microbes captured by the FlowCAM at station D2. Bar plot showing the abundance of microbes, 15 μm or larger, from station D2 throughout the year. Group A includes any image that was translucent, green, and often formed colonies. Group B were filaments or microbes that were straight exterior. Group C microbes that formed tighter colonies or microbes that were darker and less translucent. Group D were round microbes. Group E were large microbes. Group F were blurry or unidentifiable images.

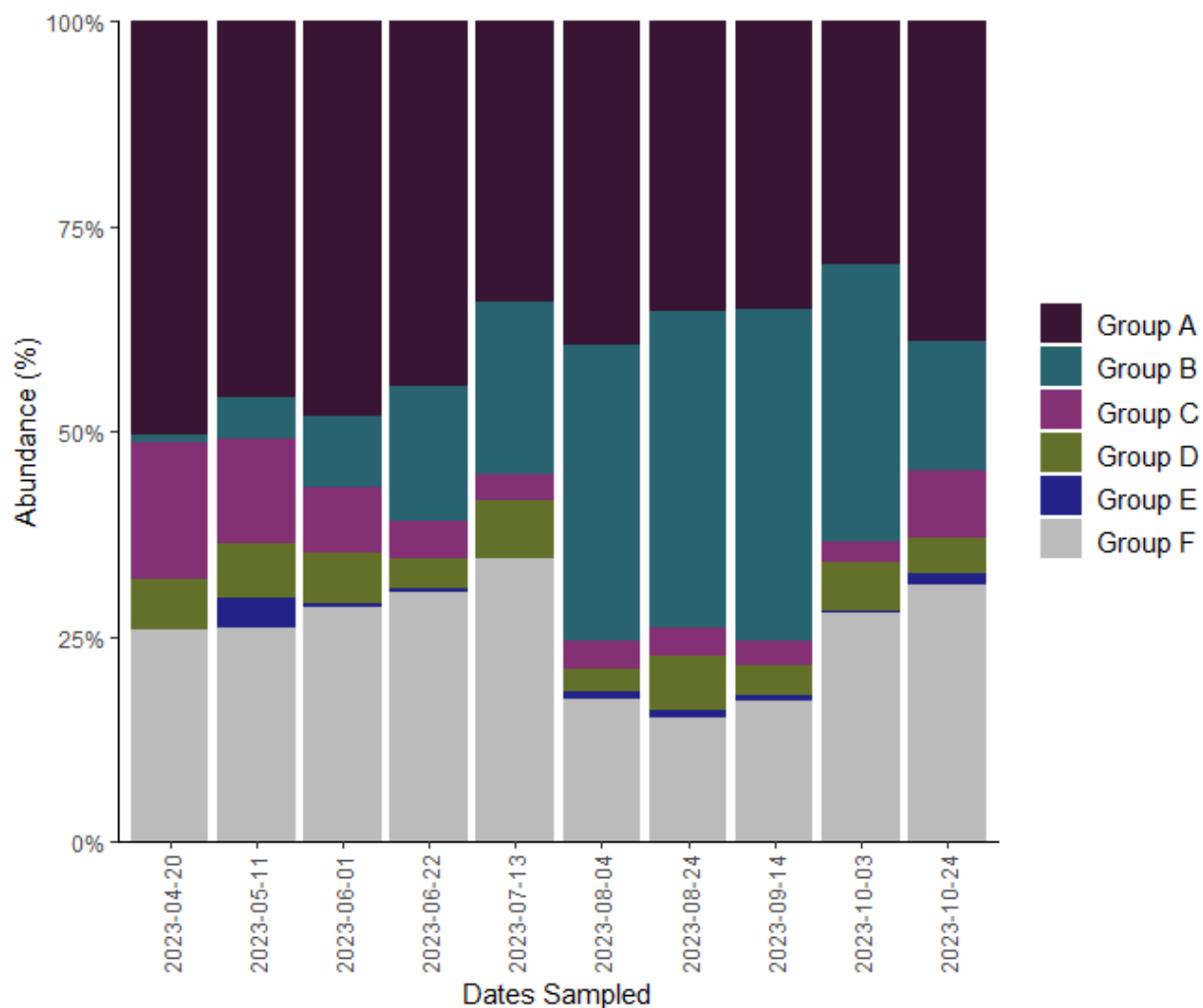


Figure 36: Abundance of microbes captured by the FlowCAM at station C1. Bar plot showing the abundance of microbes, 15 μm or larger, from station C1 throughout the year. Group A includes any image that was translucent, green, and often formed colonies. Group B were filaments or microbes that were straight exterior. Group C microbes that formed tighter colonies or microbes that were darker and less translucent. Group D were round microbes. Group E were large microbes. Group F were blurry or unidentifiable images.

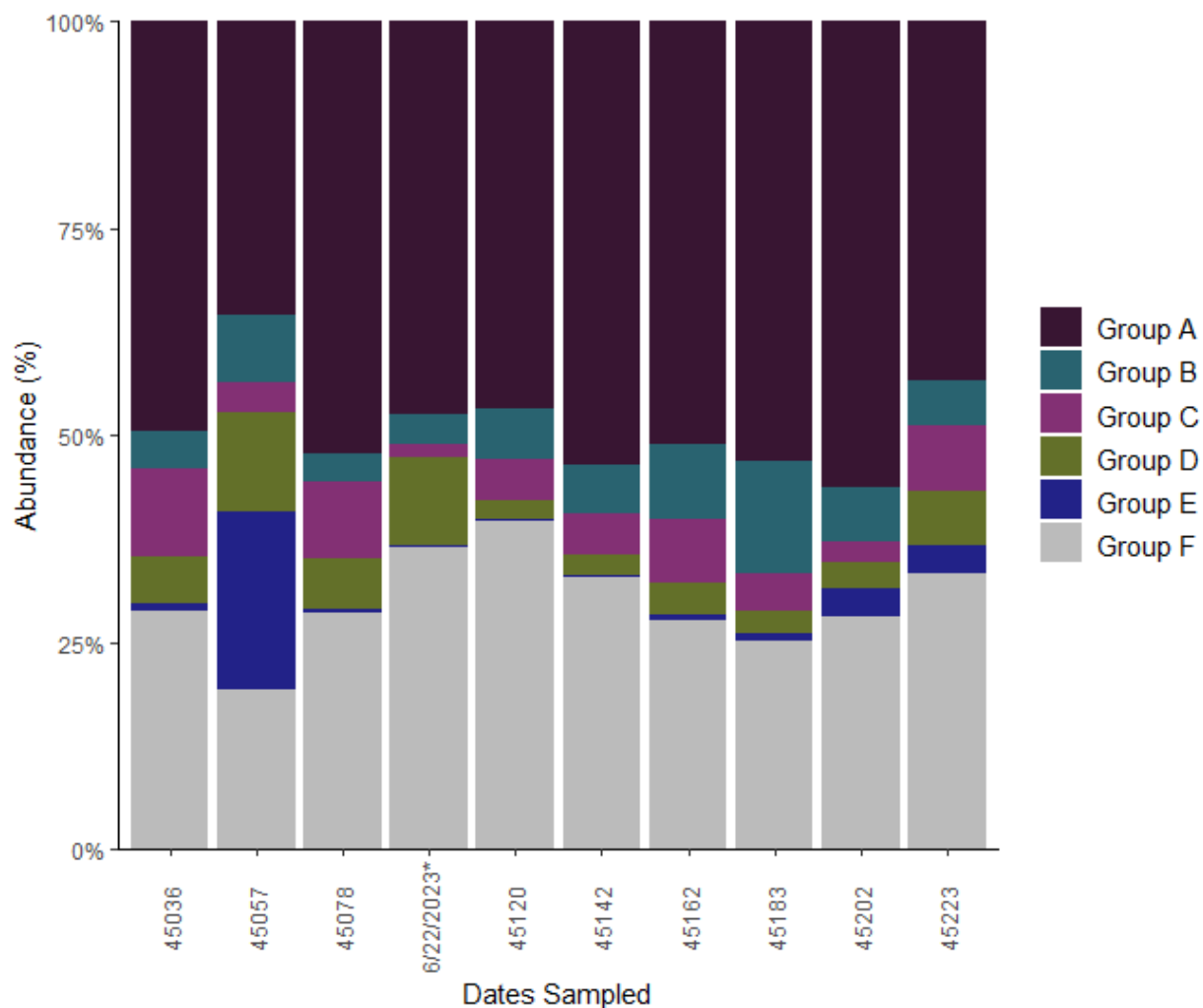


Figure 37: Abundance of microbes captured by the FlowCAM at station C2. Bar plot showing the abundance of microbes, 15 μm or larger, from station C2 throughout the year. Sample date 6/22/23 captures microbes 10 μm or larger. Group A includes any image that was translucent, green, and often formed colonies. Group B were filaments or microbes that were straight exterior. Group C microbes that formed tighter colonies or microbes that were darker and less translucent. Group D were round microbes. Group E were large microbes. Group F were blurry or unidentifiable images.

Vita

David was born in Herne, Germany. He attended Schiller-Schule Bochum until the end of 10th grade, when he moved to the United States. There, he finished high school at Webb School Knoxville, before attending the University of Tennessee. He completed his Bachelor of Science degree in both Microbiology and German. After graduation he joined the Department of Microbiology at the University of Tennessee in Dr. Steven Wilhelm's laboratory. He completed the requirements for the Master of Science degree in July 2024.