Applications of Sensory Analysis for Water Quality Assessment

Julia Frances Byrd

Thesis submitted to the faculty of Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Environmental Engineering

Daniel L. Gallagher
Andrea M. Dietrich
Gregory D. Boardman

18 December, 2017
Blacksburg, VA

Keywords: Coal ash, taste and odors in drinking water, earthy/musty chemical odorants, sensory analysis, principal component analysis

Applications of Sensory Analysis for Water Quality Assessment

Julia Frances Byrd

ACADEMIC ABSTRACT

In recent years, communities that source raw water from the Dan River experienced two severe outbreaks of unpleasant tastes and odors in their drinking water. During both T&O events, which lasted for several months, strong 'earthy', 'musty' odors were reported, but the source was not identified. The first T&O event began in early February, 2015 and coincided with an algal bloom in the Dan River, visible under ice near the Danville raw water intake. The algal bloom was thought to be the cause of the odors, but after the bloom dissipated, unpleasant odors persisted until May 2015. The second T&O in October, 2015 did not coincide with observed algal blooms.

On February 2, 2014 approximately 39,000 tons of coal ash from a Duke Energy coal ash pond was spilled into the Dan River near Eden, NC. The spill occurred upstream from of the raw water intake for the city of Danville. As there were no documented T&O events before the spill, there is concern the coal ash adversely impacted water quality and biological communities in the Dan River leading to the T&O events. In addition to the coal ash spill, years of industrial and agricultural activity in the Dan River area may have contributed to the T&O events.

The purpose of this research was to elucidate causes of the two T&O events by evaluating sensory, chemical, and biological data collected monthly from August, 2016 to September, 2017 from twelve sites along the Dan and Smith Rivers. The Smith River was included as a control for the coal ash in the Dan River.

There were no reported T&O events during the project but the results from Flavor Profile Analysis characterized earthy/musty odors were present with a mean intensity of 3.9 (weak) in column water and 6.4 (weak to moderate) in interstitial waters. No temporal or spatial trends for odors were observed. Seven earthy/musty chemical odorants were at times detected by solid-phase microextraction gas chromatography/mass spectrometry. The intensity of odors was mainly driven by geosmin, but no relationship between strong odors and odorants was observed. Acitnomycetes, associated with earthy/musty odorants were detected, but not corresponding with elevated levels of odorants.

Applications of Sensory Analysis for Water Quality Assessment

Julia Frances Byrd

GENERAL AUDIENCE ABSTRACT

In recent years, communities that source water from the Dan River experienced two severe and unprecedented outbreaks of unpleasant tastes and odors (T&O) in their drinking water. During both odor events strong 'earthy', 'musty' odors were reported, but the source was not identified. The first event began in early February, 2015 and coincided with an algal bloom in the Dan River. The algal bloom was thought to be the cause, but after the bloom dissipated, odors persisted until May 2015. The odors returned in October, 2015 but did not coincide with an algal bloom.

On February 2, 2014 approximately 39,000 tons of coal ash from a Duke Energy coal ash pond was spilled into the Dan River near Eden, NC. As no documented odor events occurred before the spill, there is concern the coal ash adversely impacted the water quality in the Dan River leading to the odor events.

The purpose of this research was to elucidate causes of the two odor events and provide guidance to prevent future problems. Monthly water samples were collected from August, 2016 to September, 2017 from twelve sites along the Dan and Smith Rivers. Multivariate analyses were applied to look for important factors.

There were no reported odor events during the project but sensory analysis characterized earthy/musty odors present. No temporal or spatial trends of odors were observed. Seven earthy/musty odorants commonly associated with odor events were detected.

Acknowledgements

I would like to thank Dr. Daniel Gallagher, my advisor on this journey. Thank you for taking a chance on me. Thank you for second and third chances to prove myself, thank you for answering a million questions, and thank you for patience and having a healthy sense of humor. I also would like to give special thanks to Dr. Andrea Dietrich who always had an answer, or a different way to approach the problems. Thank you to Dr. Board for welcoming me on this research project and providing so many learning opportunities.

Thank you to Mrs. Jody Smiley and Ms. Julie Petruska for their patience, guidance, and kindness to me as I learned to navigate instruments in the laboratory. Many of my experiments would have failed without them both!

Special thanks to my research partner, Keegan Waggener, who kept me sane, safe, and smelled a million samples. Thanks your hard work and putting up with me. Thank you to the wonderful willing participants who helped us with sampling, smelling, cleaning and keeping spirits high, Dalia Rahka, Wenchuo Yao, Mark Cheng, Chase Altizer, Bharani Rajasekaran, and Nevetha Ramesh.

Thank you to my family and friends and to my greatest love, Alastair Colquhoun (the driver). I would not be here today without your love and guidance.

Table of Contents

Chapter 1: Characterization and Evaluation of Odors, Odorants, and Biological Communities on the Dan and Smith Rivers

1. Introduction 1
2. Literature Review 1
3. Methods & Materials 9
4. Results 17
5. Discussion 32
6. Conclusions 33

Chapter 2: Multivariate Analyses and Statistical Testing

1. Introduction 34
2. Literature review 35
3. Methods & Materials 35
4. Multivariate Results 44
5. Discussion 57
6. Conclusions 58
7. References 59

List of Figures

Figure 1. Map of research area and sampling locations....10
Figure 2.Odor descriptors from FPA for all samples....18
Figure 3. Descriptors of odors in Column (left) and Interstitial (right) water samples by FPA....19
Figure 4. Levelplot of column water mean max odor intensity....20
Figure 5. Levelplot of interstitial mean max odor intensity....21
Figure 6. Column and Interstitial water samples prepared for FPA analysis....22
Figure 7. Mean concentrations of earthy/musty odorants for Column (left) and Interstitial (left) by site....23
Figure 8. log_{10} OAVs of earthy/musty odorants....24
Figure 9. Odor Activity Values of earthy/musty odorants in Column (left) and Interstitial (right) by site....25
Figure 10. Regression of mean max odor intensity by log OAVs of different earthy/musty odorants....26
Figure 11. Regression of sum of odorants (Log_{10}) versus odor intensity for all water samples....27
Figure 12. Regression Tree of OAV influence on odor intensity (FPA) in column water....28
Figure 13. Regression Tree of OAV influence on odor intensity (FPA) interstitial water....29
Figure 14. Levelplots Actinomycetes (log_{10}) in column CFU/L (left) and interstitial CFU/g (right) samples....30
Figure 15. Actinomycetes versus Geosmin & 2- MIB (ng/L)....31
Figure 16. Map of research area and sampling locations....36
Figure 17. Cluster dendogram of column water samples containing all data results from sensory, chemical, and biological analyses....45
Figure 18. Cluster dendograms of interstitial water samples containing all data results from sensory, chemical, and biological analyses....46
Figure 19. Probabilistic principal component analysis of all data samples....47
Figure 20. Probabilistic PCA biplot for all water samples....47
Figure 21. Probabilistic principle component analysis of column water samples....49
Figure 22. Biplot of PPCA for column water samples....50
Figure 23. Probabilistic principle component analysis of interstitial water samples....51
Figure 24. Biplot of PPCA for interstitial water samples....52
Figure 25. Principle component analysis of column water with missing values removed....53
Figure 26. PCA biplot of water column samples....54
Figure 27. Principle component analysis of interstitial water samples with missing values removed....55
Figure 28. PCA biplot for interstitial water samples....56

List of Tables

TABLE 1. ODOR THRESHOLD CONCENTRATIONS OF COMMON T&O COMPOUNDS IN DRINKING WATER 3
TABLE 2. DATES OF SITE TOUR AND SAMPLING TRIPS 9
TABLE 3. FLAVOR PROFILE ANALYSIS DATA WORKSHEET EXAMPLE 12
TABLE 4. FLAVOR PROFILE ANALYSIS INTENSITY SCALE 13
TABLE 5. GC/MS ODORANTS AND QUANTIFICATION 14
TABLE 6. MOST FREQUENTLY USED ODOR DESCRIPTORS 19
TABLE 7. SAMPLES ANALYZED FOR EARTHY/MUSTY COMPOUNDS 22
TABLE 8. SUMMARY OF LINEAR REGRESSION FOR SUM ODORANTS INFLUENCE ON ODOR INTENSITY 27
TABLE 9. DATES OF SITE TOUR AND SAMPLING TRIPS 36
TABLE 10.FLAVOR PROFILE ANALYSIS DATA WORKSHEET EXAMPLE 39
TABLE 11. ODOR PROPERTIES FOR REFERENCE STANDARDS USED IN FPA SENSORY TRAINING 39
TABLE 12. FLAVOR PROFILE ANALYSIS INTENSITY SCALE 40
TABLE 13. GC/MS ODORANTS AND QUANTIFICATION 41
TABLE 14. SUMMARY OF ANALYTICAL METHODS AND ANALYTES MEASURED 42
TABLE 15. PRIMARY CONTRIBUTING SAMPLES FOR PPCA 48
TABLE 16. PRIMARY COLUMN WATER SAMPLES CONTRIBUTING TO (PPCA) COMPONENTS 49
TABLE 17. PRIMARY CONTRIBUTING INTERSTITIAL WATER SAMPLES (PPCA) 51
TABLE 18. PRIMARY COLUMN WATER SAMPLES CONTRIBUTING TO PRINCIPLE COMPONENTS 53
TABLE 19. PRIMARY INTERSTITIAL WATER SAMPLES CONTRIBUTING TO PRINCIPLE COMPONENTS 55

Chapter 1: Characterization and Evaluation of Odors, Odorants, and Biological Communities on the Dan and Smith Rivers

1. Introduction

In recent years, communities that source potable water from the Dan River have experienced severe outbreaks of strong, unpleasant tastes and odors (AWWA) in their drinking water. The two most recent T&O events lasted for several months during which strong 'earthy', 'musty' odors were reported in the water but the source was not identified. The first T&O event began in early February, 2015 and coincided with an algal bloom in the Dan River that was observed under the ice near the Danville raw water intake. Initially the algal bloom was thought to be the cause of the odors, but even after the bloom dissipated, the unpleasant odors persisted until May 2015. The second T&O was first noted in October, 2015 and did not coincide with any observed algal blooms. The second T&O event was mitigated to an extent by additional water treatment process at the Danville water treatment plant. In September 2015, the city of Danville upgraded the existing powdered activate carbon (PAC) feeder system, due to insufficient capacity of the older system to provide the larger dose required. Although to date there has not been a subsequent T&O in the area, and there were no detected public health risks, there is a great desire to determine the cause of the odors in order to prevent future T&O outbreaks.

2. Literature Review

2.1. EPA recommendations for taste and odors in drinking water

The Environmental Protection Agency (EPA, 2017) set two types of drinking water standards to protect public health and provide safe, acceptable drinking water. The Primary Maximum Contaminant Levels (PMCLs) are federally enforceable standards that set mandatory limits on contaminants in drinking water that present a risk to human health. The second type of standard, the National Secondary Drinking Water1 Regulations (NSDWRs) (EPA, 2017) are designed to assist public water suppliers with guidelines to manage aesthetic water quality such as color, taste, and odors. The NSDWRs provides a list of Secondary Maximum Contaminant Levels (SMCLs) which are not known to harm human health, but may cause the water to appear discolored or taste or smell bad. The SMCL for odor in drinking water is based on the Threshold Odor Test (AWWA, 2011), a dilution method to analysis the odor in a water sample. Threshold Odor Number (TON) is the ratio of diluted sample with odor free water at which odor is just detectable and calculated by Equation 1.

Equation 1

$$TON = \frac{A + B}{A}$$

Where:
A = mL sample
B = mL odor-free water

The SMCL sets the maximum value Threshold Odor Number (TON) at 3; the level at which odor is just barely detectable in the water sample at 60° C (EPA, 2017 #42). TON is also not effective for assessing odors in drinking water because odors descriptors are not part of the method, nor does it correlate with consumer complaints (AWWA, 2002; Dietrich, 2015 #44)

Drinking water providers are not bound by federal law to monitor for taste and odor (AWWA, 2011), but consumers often judge drinking water quality based on aesthetics, offensive smelling or tasting water. These issues may cause consumers to lose confidence in the quality of tap water (EPA, 1979; AWWA, 2002; Dietrich, 2015; Pierce, 2017). Some states opted to enforce SMCLs and require treatment of tastes and odors in drinking water.

AWWA conducted a survey of 168 water utilities which revealed that 77% of responding utilities take T&O issues very seriously, 47% recently updated treatment systems to better handle T&O, and about one third (29%) employed preventative measures, such as algaecides (Ömür-Özbek, 2012). The same survey indicated that about half (48%) of the utilities employ sensory methods to monitor taste and odor issues and 80% of those utilities employ the TON method as their main sensory test while other methods such as Flavor Profile Analysis (Standard Method 2170) used less often.

2.2. Origins and occurrences of T&O events in drinking water

The origin of contaminants that give rise to T&O events in drinking water can come from both natural and anthropogenic sources (Piet, 1978). Chemical identification of the odorant causing a T&O event is critical for managing T&O issues but can be difficult due to sometimes complex mixtures of odorants that may be present at various concentrations (Khiari, 1995). Odorous compounds that have been identified as causes of T&O in drinking water are summarized below in Table 1, as well as their commonly used odor descriptor(s), odor threshold concentrations (OTC), and any known sources. Threshold values (OTC) are generally defined as the lowest concentration at the probability of detection is 0.5, meaning that 50% of a general population can detect the odor. Reported threshold values vary as individuals participating in sensory analysis studies have different olfactory capacities; thus, some people may detect odors at higher or lower levels than others.

Table 1. Odor Threshold Concentrations of Common T&O Compounds in Drinking Water

Compound	**Odor descriptor**	**Aqueous Odor Threshold ng/L**	**Sources**	**References**
Haloanisoles				
2,4-dichloroanisole	Musty	400	Biomethylation of chlorophenol	Saxby, 1992
2,6-dichloroanisole	Musty	40	Biomethylation of chlorophenol	Saxby, 1992
2,3,4-trichloroanisole	Musty	0.2-2	Biomethylation of chlorophenol	Zhang, 2016
2,4,6-trichloroanisole	Musty; Moldy	0.03-10	Biomethylation of chlorophenol	Suffet, 1999;Diaz, 2005
2,4,6-tribromoanisole	Musty; Moldy	0.15-10	Biomethylation of bromophenols	Zhang, 2016
Halophenols (precursors)				
2,4,6-trichlorophenol	Phenolic	n.a.	Fungicides, herbicides, wood pallets	
Methoxypyrazines				
3-isopropyl-2-methoxypyrazine	Sour; Muddy	2-16	Actinomycetes	Deng et al., 2010; Peng et al., 2014
3,5-dimethly-2-methoxypyrazine	Fungal; Musty	0.4	Bacteria, fungi	Czerny, 2000
3-isobutyl-2-methoxypyrazine	Vegetation	1-16	Actinomycetes	Deng et al., 2010; Peng et al., 2014
Other Earthy/Musty				
Geosmin	Earthy; Musty	1-10	Cyanobacteria, fungi, actinomycetes	Gerber, 1983; Suffet, 1999; Piriou,2009; Zhang, 2005; Deng et al.,2010; Peng et al., 2014
2-methylisoborneol	Earthy; Musty; Moldy	1-15	Cyanobacteria, fungi, actinomycetes	Suffet, 1995; Zhang, 2005; Deng, 2011 ; Peng, 2014
Other chemicals				
2,3 butanedione	Sweet; Buttery	50	Paper manufacturing	Diaz, 2005
2-ethyl-4-methyl-1,3-dioxolane	Sweet; Medicinal	$5*10^5$	Resin manufacturing	Noblet, 1999
methyl tert-butyl ether	Waxy; Citrus; Plastic	$1.5*10^4$	Gasoline additive	Stocking, 2001

Iodinated trihalomethanes				
Iodoform	Medicinal; Sweet	30- 100	Disinfection byproduct	Cancho, 2001
Bromodiiodomethane	Medicinal; Sweet; Solvent	100-200	Disinfection byproduct	Cancho, 2001
Sulfides				
dimethyl disulfide	Septic; Garlic; Putrid	200-330	Decaying vegetation, algae	Gaudagni et al., 1978; Saxby, 1992; Deng et al., 2010
dimethyl trisulfide	Garlic; Swampy	10	Decaying vegetation, algae	Buttery et al., 1976; An, 2012; Peng et al., 2014; Guo et al, 2015
Aldehydes				
2*t*-hexenal	Grassy	$5.0*10^4$		Rashash et al, 1997; Ömür-Özbek, Dietrich, 2008
2*t*-heptanal	Sweet; Fruity; Grassy	$4.0*10^4$		Rashash et al., 1997
2*t*-octenal	Sweet; Melon; Oily	$<6.0*10^3$		Rashash et al., 1997

Severe T&O outbreaks can affect water treatment plant operations and cause shortages as was seen in 2007 when Lake Taihu experienced foul odors that resulted in cutting off the water supply to 2 million residents (Ma, 2013). The presence or absence of taste and odor in drinking water is not directly related with its toxicity, but for consumer satisfaction water should be both toxicologically safe and aesthetically acceptable (Young, 1999).

A review of taste and odor events in the drinking water in Barcelona, Spain from 1990-2004 found that industries along the Llobregat River and its tributaries were the primary sources of objectionable odorants (Boleda, 2007). Wastewater treatment effluents were identified as a source of malodorous compounds to drinking water supplies in Northwest Spain. Towns in Galicia whose drinking water source is the Barces River experienced T&O events of several months during which strong earthy, “sour dish cloth”, and potato bin odors were present in the drinking water. The compounds producing the odor events were identified using closed-loop stripping analysis (CLSA) with sensory-GC and GC-MS as alkyl-methoxypyrazines including 3,5-dimethly-2-methoxypyrazine (MDMP), 3-isopropyl-2-methoxypyrazine, and 3-isobutyl-2- methoxypyrazine (Deng, 2010). This was the first odor event that MDMP was reported as the responsible compound. The source of the alkyl-methoxypyrazines was traced to effluent from a wastewater treatment plant upstream of the drinking water plant raw water intake (Ventura, 2010).

Industrial discharge and spills have also released odorous compounds into drinking water sources leading to taste and odor events. When a sweet, buttery odor in raw and treated waters in Barcelona, Spain was detected intermittently from winter 2000 to summer 2002 SPME and sensory-

GCMS identified the compound 2,3 butanedione in the Anoia River, a tributary of the drinking water source to Barcelona. The origin of the diacetyl was a paper factory located along the Anoia River (Diaz, 2005).

In another odor event that occurred in the Ohio river the odorous compound 2-ethyl-4-methyl-1,3-dioxolane was formed by reactions with chemicals in the rinse water from a resin manufacturer. In this incident 2EMD, which is hydrophilic in nature, was not completely removed by wastewater and drinking water treatment processes and odors described as sweet medicinal or sickeningly sweet were smelled far downstream of the point of discharge and in the drinking water (Noblet, 1999).

Unsuspecting sources such as materials used in distribution systems can cause T&O issues, as well as potential health risks (Tomboulian, 2004; Skjevrak, 2003; Heim, 2007). Coatings and additives used in system materials can migrate into water where reactions with chlorine, other leached additives, fungi, bacteria, or algae in the system can form malodourous compounds. Tomboulian identified material and chemical sources of system components that can impart T&O into drinking water. For example, pipes and liners made of cement/concrete can be a source of 2,4,6- tribromoanisole; 2,4,6-tribromophenol; 2,4,6- trichloroanisole; 2,4,6- trichlorophenol; and many of chemicals of varying organoleptic properties. (Skjevrak et al.,2003) found that volatile organic compounds can migrate from high-density polyethylene and cross bonded polyethylene pipes to the water in contact. VOCs with odorous characteristics such as methyl tert-butyl ether (Stocking, 2001; Suffet, 2007) were identified in water samples from PEX pipes and may be the contributor to strong reported odors (high TON values). Heim and Dietrich reported HDPE pipes imparted perceptible 'waxy/plastic/citrus' odors in water with and without disinfection treatment.

Disinfection processes in water treatment are also a source of taste and odor issues in drinking water. The most common consumer complaints to water utilities are "chlorinous", and "earthy-musty" odors (Ömür-Özbek, 2012). Disinfection may be the cause of both odors as chlorination can stimulate actinomycetes to produce the "earthy-musty" compound 2-MIB in distribution systems (Abbaszadegan, 2015). Abbaszadegan showed that during a 12-day, non-chlorinated cycle of water spiked with actinomycetes, there were no changes in actinomycetes nor 2-MIB concentrations. Three cycles of chlorination were tested for both cast iron and PVC systems. The first shock of chlorine (1 mg/L) did not stimulate actinomycetes to produce 2-MIB. However, at the start of the second chlorination cycle, 2-MIB levels significantly increased in both systems, 2.5 to 22.7 ng/L in the cast iron and 2.3 to 20.5 ng/L in the PVC. The levels of 2-MIB remained elevated through the chlorine depletion period. This was the first report of intermittent chlorine to stimulate 2-MIB production by actinomycetes in drinking water distribution systems.

The composition of water before the chlorine addition influences the disinfection by-products (DBP) formation some of which are compounds with low odor thresholds, such as iodinated trihalomethanes (ITHMs) with OTCs in the range of 0.03-0.2 μg/L (Table 1), and often not effectively removed by water utilities. Due to mining activities upstream of the Llobregat river, the raw water entering the Barcelona water treatment plant is highly saline, leading to trace amounts of iodinated ITHMs forming during the post stabilization process (Cancho, 2001). Even at low concentrations ITHMs can impart a medicinal odor in water. Cancho used flavor profile analysis (FPA) and gas chromatography coupled with olfactrometry (GCO) to determine the odor threshold concentrations of ITHMs in drinking water and found the general trend that more iodinated compounds were more odorous (lower OTCs).

In a study of 22 cities in China, DBPs, including chloroanisole, bromoanisole, and chloro-bromoanisole, were identified as important odorants imparting earthy-musty, off flavors in municipal tap water (Zhang, 2016). Zhang reported that generally there are three types of sources of haloanisoles

in water: biotransformation of halophenol by actinomycetes and fungi that are present in water distribution system; biotransformation by biofilm from halophenol; and the presence of materials that contain the precursor. These T&O compounds can form in the distribution system and delivered to consumers with no treatment which can cause serious problems for water utilities.

Zhang (2005) reported the concentration and occurrence of haloanisoles, including 2,4,6-trichloroanisole; 2,3,6-trichloroanisole (2,3,6- TCA); 2,3,4 -trichloroanisole (2,3,4 -TCA); and 2,4,6-tribromoanisole, (2,4,6-TBA) were higher in summer than in winter. In both summer and winter 2,3,6-TCA showed the highest average concentrations of 16.90 and 2.21 ng/L respectively and 2,4,6-TBA had the lowest concentrations of 0.325 ng/L (summer) and 0.22 ng/L (winter). These occurrences and concentrations of haloanisoles were higher than their odor threshold concentrations (OTC) which necessitates effective removal. Zhang suggests that technology and design of the water distribution system should be considered to reduce the formation of haloanisoles to undetectable levels.

Investigations of taste and odor events in Sweden revealed that 2,4,6-TCA was responsible for earthy/musty flavors in drinking water (Nystrom, 1992). In the four T&O events in three towns the TCA concentrations increased in the distribution systems after leaving the water treatment plant. In one such event the TCA concentrations at the water utilities were below the detection limit of 0.1 ng/L, although higher than the 0.03 ng L^{-1} OTC, whereas 2,4,6-TCA levels in the distribution system had increased to 1.7 ng/L. The likely cause of 2,4,6-TCA in drinking water are chlorinated phenols including, 2,4,6 trichlorophenol (2,4,6-TCP), which may be biomethylated to the corresponding anisoles in the distribution system.

2.3. Precursors and Production of Common T&O Compounds

The *O*-methylation of 2,4,6-TCP is critical for biosynthesis of 2,4,6-TCA (Zhang, 2016). Zhang reported the production of 2,4,6-TCA increased by about 1.4 fold when 2,4,6-TCP was increased from 0.01 to 0.2 mg/L. The study also found that higher temperatures (35 °C) are conducive for the biosynthesis of 2,4,6-TCA which is consistent with T&O problems often occurring in late summer and early fall when higher water temperatures are favorable for microbial growth. Additionally, algae species and cyanobacteria can produce 2,4,6-TCA which may contribute to the higher formation potential (FP) in lake water than in river or tap water. Zhang suggests that due to eutrophication and static flow, microorganisms such as cyanobacteria and algae in lakes can produce many T&O compounds including 2,4,6 TCA, geosmin and 2-methlylisoborneol (2-MIB).

The majority of biologically caused T&O events in drinking water are caused by the microbial production of (-) geosmin and (-) 2-MIB which are highly potent, earthy, musty, moldy metabolites. Geosmin and 2-MIB are problematic for water utilities due to their low OTC (Table 1), and resistance to natural degradation, boiling and conventional treatment process. Many aquatic microorganisms synthesize geosmin and 2-MIB, including actinomycetes and fungi, however, among algae in fresh water lakes, reservoirs and rivers they are largely produced by photoautotrophic blue-green algae (cyanobacteria) (Watson, 2007). Cyanobacteria are an important source of volatile organic compounds (VOCs) such as the terpenoids geosmin and 2-MIB (Watson, 2003). Watson purports the objectionable smell of geosmin and 2-MIB in drinking water is indirectly beneficial because although they are not known to be toxic to humans, the cyanobacteria responsible for their production also produce toxins, which could account for the infrequent poisonings by cyanotoxins.

Watson compared combined data for geosmin and 2-MIB and found that any allelopathic activity they exhibit occurs at levels orders of magnitude higher than their odor detections by humans. Geosmin and 2-MIB are chiral molecules with a (+), and a more potent (up to 10 times stronger) (-)

enantiomer (Piriou, 2009). Because the (-) enantiomer is biologically produced their presence in drinking water maybe elicit consumer complaints at levels much lower than estimated OTCs used by drinking water industry, which are determined using commercially available racemic mixes (Watson, 2016;Watson, 2016;Rashash, 1997).

Many efforts have been made to trace the biological producers of geosmin and 2-MIB, which have been largely unsuccessful due to several key factors, including lack of standard units for reporting (Watson, 2007). Watson emphasizes that reporting production and production rates should be expressed as per volume/weight of cells or per unit chlorophyll, which provides insight to the relative capacity of different taxa to metabolize geosmin/2-MIB and for comparing different studies. Both are produced by several cyanobacterial and/or other taxa which can occur either simultaneously or at different temporal and spatial locations in the same water body. Watson (2016) shows that reported *per cell* production of both geosmin and 2-MIB can vary among and within species. Watson adds that some of the reported differences reflect issues with species misidentification and quantification; additionally, advances in molecular and biochemical diagnostics mean that taxonomy of cyanobacteria are often revised.

Surface water blooms are often considered the source of odors in water, but many of the known cyanobacterial producers are nonplanktonic (Watson, 2007). Often overlooked but important source of T&O outbreaks are benthic cyanobacteria, which can release VOCs when detached mats decompose, particularly in shallow areas with low flow (Watson, 2016). Watson presents several methods often used to control benthic algae, including desiccation after reservoir drawdown and physical removal. Watson argues that in addition to environmental and health risks, the use of algaecides can trigger the release of geosmin/2-MIB from lysed cells and results in development of resistant strains.

T&O events have been associated with eutrophication leading to dense blooms of cyanobacteria and eukaryotic algae, which are primary causes of foul odors in source waters (Watson, 2003; Watson, 2016). In the case of Lake Taihu in China, strong septic, musty, and marshy odors triggered great public concern and led to the investigation of dense blooms of cyanobacteria as a source (Ma, 2013). Ma found that cyanobacterial decomposition led to anoxic water conditions, increased nutrient loading and released volatile organic compounds in to the water. The results showed that dimethyl sulfide (DMS) and dimethyl-trisulfide (Deng, 2010) were the dominant T&O compounds causing the septic odor in the anoxic decay phase. Additionally, geosmin and 2-MIB were identified as dominant in the live growth phase of cyanobacterial blooms, which indicates their role in the earthy, musty odors.

The occurrence of odors and algal growth in source water is critical to understanding and controlling T&O issues. The Yellow River in China serves as the drinking water source for six major cities: Lanzhou, Yinchuan, Hohhot, Zhengzhou, Jinan and Dongying. An alternative management strategy is to use settling reservoirs to reduce turbidity before water treatment, which has also created an ideal location for algal and cyanobacteria growth and increased fishy odors in the winter time (Li, 2016). The investigation of odor events in six cities along the Yellow River during winter used FPA to characterize the odor profile of the water and GC-time-of-flight mass spectrometry analysis to identify specific odorant compounds. Earthy/musty odorants were identified in four cities at concentration levels above their OTCs, geosmin at 2.26 and 6.04 ng/L, and 2-MIB at 16.72 and 21.12 ng/L with a positive correlation for their sum odor activity values (OAVs) and earthy/musty odor intensity. Odor activity values are calculated by dividing concentration of the odorous compounds by their odor threshold value (Grosch, 1994; Jelen et al., 2013). OAVs were used to evaluate the contribution of odor from different compounds. Four septic/swampy sulfides were analyzed and only DMDS was detected in the source waters in three locations, all at levels exceeding the OTC (30 ng/L).

The source of fishy odor in the waters remains elusive, although there is suspicion that aldehydes including hexanal, heptanal and nonanal may be responsible. Li (Li, 2016) reports the higher concentrations of the algae, *Cryptophyta Dinobryon,* could be responsible for the fishy odors perceived in two cities. However, another city showed different results, having low algal density but strong fishy odors. In accordance with previous work, Li found that geosmin and 2-MB were responsible for the earthy/musty odors, and dimethyl disulfide (DMDS) for the rancid/swampy odors likely caused by the growth and decay of cyanobacteria.

2.4. Objectives to identify the cause of T&O in the Dan River

The occurrences of T&O events are variable and hard to predict. The literature on T&O events elucidates that biological processes can produce or transform chemicals in the aquatic system into odorous compounds detectable by humans at very low concentrations. There are many different biological process which may cause or contribute to these issues. The objectives of the research outlined in this report are to characterize the biological communities in the Dan River watershed and monitor for shifts in community structures, with a focus on groups that are known to be associated with the production of earthy/musty odorants. SPME GC/MS, a common and efficient analytical method for identification and quantification of chemical odorants in river water. The sensory analysis method, Flavor Profile Analysis, has been used effectively with SPME GC/MS in past research on T&O events in source waters. Monthly river water samples were analyzed for water their biological and ecological composition, specific earthy/musty odorous organics, and sensory characteristics. The results from these analyses will help characterize the temporal and/or spatial trends under normal conditions, and identify potential locations where odorous compounds originate.

3. Methods & Materials

3.1. Water Sampling

Dates of the site tour and monthly sampling trips are summarized below in Table 2. Samples were collected monthly from August, 2016 until September, 2017; a total of thirteen trips. There was not a trip in April, 2017 due to unsafe river conditions caused by several days of heavy rain in the area. In the event of an outbreak of odors, extra sampling trips would have been included, but this did not occur during the project timeline.

Table 2. Dates of site tour and sampling trips

Days and month	Year
July 15 site tour	2016
August 16-17	2016
September 14-15	2016
October 4-6	2016
November 15-17	2016
December 16-18	2016
January 10-12	2017
February 3-5	2017
March 21-22	2017
May 16-17	2017
June 13-14	2017
July 10-11	2017
August 7-8	2017
September 19-20	2017

Samples of water are collected at twelve sites along the Dan and Smith Rivers either from a boat or on a bridge, as well as two drinking water treatment plants that source water from the Dan River, Danville and South Boston. At boat sites, testing for water quality and biological constituents are performed whereas from the bridge sites, only water quality is characterized. Figure 1 shows the boat sites as triangles, bridge sites as circles, and the drinking water treatment plants are rectangles. A description of the sites and methods used at those sites follows.

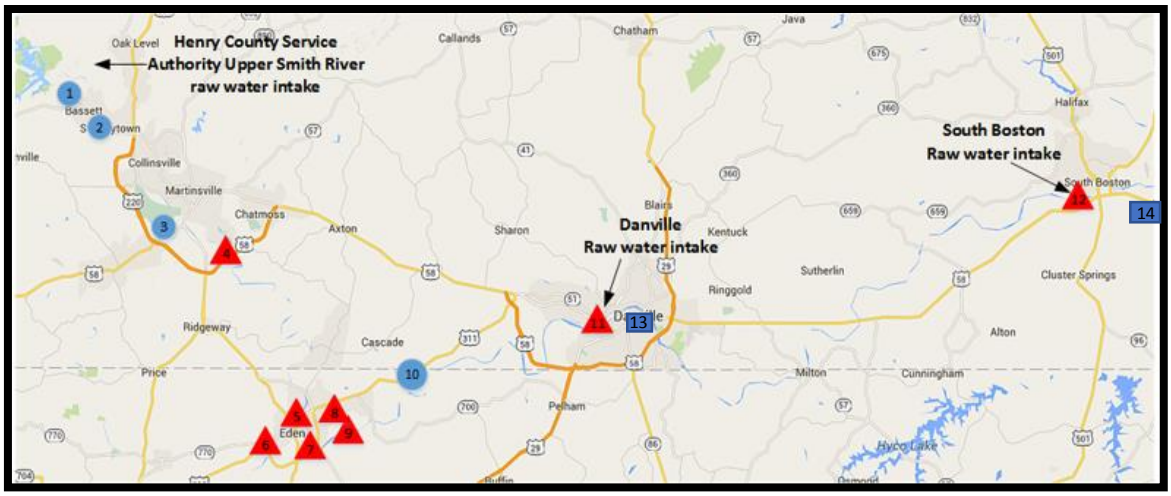

Figure 1. Map of research area and sampling locations

3.1.1.Bridge Sampling Sites

Water column samples are collected by lowering a bucket attached to a rope into the river at its center. Interstitial water and sediments are collected using an Ekman dredge at the center, left and right sides of the river to make a composite sample. Additionally, samples of raw influent water and finished water are collected at the Danville and South Boston drinking water treatment plants.

S-1 Bassett This site is the furthest upstream site on the Smith River, located at the bridge crossing the river along Bullocks drive in Bassett, VA. Sediment and interstitial water samples are usually taken by hand from both sides of the riverbank because the rocky riverbed is not conducive for using the Ekman dredge.

S-2 Fieldale This site is located at the bridge crossing the Smith River along South Daniels Creek road in Fieldale, VA. Sediment and interstitial water samples are usually taken by Ekman dredge here.

S-3 Martinsville This site is located at the bridge crossing the Smith River along Rives road in Martinsville, VA. Sediment and interstitial water samples are usually taken by Ekman dredge here.

S-10 Stateline Bridge This site is located at the bridge on the Dan River border between Virginia and North Carolina. The bridge links the State Line Bridge Road in Virginia to the Berry Hill Bridge Road in North Carolina. Sediment and interstitial water samples are usually taken by hand due to the high elevation of the bridge, which makes capturing water difficult through Ekman dredges.

3.1.2.Boat Sampling Sites

The boat sites along the Dan and Smith Rivers are sampled along a 1000-meter reach. Water samples and field measurements are collected center-stream at the upstream end of each reach. Starting at the upstream end of each reach, interstitial water and algae samples are collected near the shore at 100 meter intervals. The starting point along the left or right bank is determined at random by the toss of a coin. From there, the next two samples are collected along the same bank, then the opposite bank is sampled for two intervals, alternating banks after every second interval until the 1000-meter reach is completed. Interstitial water and algae are from each 100-meter segment composited along each 1000-meter reach.

Field sheets are kept on which are recorded the station number, date and time of collection, and a table indicating from which stream bank the sediment and algae samples are collected along each 100-meter segment.

S-4 Smith River Sports Complex The sample reach is between the Martinsville Sewage Treatment Plant and the Smith River Sports Complex. The start of the 1000-meter reach is approximately 1100 meters downstream of the launch point at the Martinsville Sewage Treatment Plant. The entire reach is downstream of the treatment plant outfall. This reach is sampled by canoe because there is no launch point for a larger boat within a reasonable distance to the reach. There are also two shallow rapids that would not be navigable in a larger boat.

S-5 Smith River This is the furthest downstream site on the Smith River and the reach sampling ends just before the confluence with the Dan River. The reach at begins just downstream of the Kings Highway bridge. Due to a set of rapids at approximately 300 meters below the bridge, and the configuration of the boat used to-date, sampling has begun 300 meters downstream of the bridge along this reach.

S-6 Dan River The furthest upstream site on the Dan River and upstream of the coal-ash spill. The sample reach begins near an unknown building on the left bank at 835 N. Bridge St, Eden, NC (from Google Maps), approximately 300-meters downstream of the Old State Highway 87 bridge. The reach ends just upstream of the confluence with the Smith River.

S-7 Dan River This site is also upstream of the coal-ash spill but downstream of the confluence with the Smith river. The sample reach begins approximately 340-meters downstream of the Leaksville Boat Ramp, and 940-meters downstream of the confluence of the Dan and Smith Rivers. The reach starts opposite a small tributary that enters the river on the right bank. The Eden Sewage Treatment plant outfall enters the river from the left bank at approximately the halfway point of the reach. The reach ends just upstream of the S. Van Buren Rd. Bridge.

S-8 Industrial Culvert This is not a river site but a low flowing stream which enters the Dan River approximately 1030-meters downstream of the Duke Energy boat ramp. Grab samples are collected at this site, rather than sampling along a 1000-meter reach. The water samples are collected from a pipe running through a stone dam that was constructed next to the river. Interstitial water and algae samples are collected approximately 100-meters upstream of the dam, at a road crossing.

S-9 Dan River The sample reach begins approximately 1900-meters downstream of the Duke Energy boat ramp, where Town Creek enters the Dan River along the right bank. The reach ends approximately 1200-meters upstream of the S. Fieldcrest Road bridge.

S-11 Dan River upstream of the Schoolfield Dam, Danville, VA The sample reach begins 1000-meters upstream of the boat ramp at Abreu Grogan Park and ends at Schoolfield Dam in Danville.

S-12 Dan River at South Boston, VA The sample reach approximately 3300-meters upstream of the boat ramp at the South Boston Sewage Treatment Plant. The reach ends at the South Boston Water Treatment Plant intakes, on the left side of the river.

3.2. Flavor Profile Analysis (FPA)

3.2.1. Flavor Profile Analysis sessions

Flavor Profile Analysis was conducted according to Standard Method 2170 for characterizing odor descriptions and intensities of each collected sample. Column and interstitial water samples were collected head space free in 500mL glass bottles were stored at 6 °C for less than 48-hours before FPA

analysis. 200 mL of sample water was poured into 500 mL Erlenmeyer flasks and heated in a water bath at 45˚ C for one hour before 4 to 6 trained panelists analyzed each sample.

Each sensory analysis session takes place in an odor free room with four to eight trained panelists who are advised not to wear perfume and not to eat or drink for at least an hour before each session. Panelists calibrate themselves for taste and odor before each session by tasting a sour reference standard (FPA = 4) and drinking taste free (Dasani) water throughout the FPA session.

Erlenmeyer flasks (500 mL) containing 200 mL of odor free water in a water bath at 45 ^{o}C are used by panelists as a calibration aid before they smell actual samples. Four to six river water samples are analyzed in each session. Panelists wait at least two minutes between smelling samples. Panelists are instructed to use a fragrance-free paper towel while holding the bottom of Erlenmeyer flasks, remove the ground glass stopper, and smell the odor free water or river water samples. River water samples are labeled with a numeric code so that panelists do not know the location or type of sample.

Panelists recorded odor descriptions and intensities for each sample on worksheets, an example is shown below in Table 3. Panelists often listed several descriptors and intensity ratings to describe each odor present in a sale. For example, in Table 3, Sample 1 had an “earthy” aroma that was moderate (6-8) intensity, as well as a strong (12)” swampy” odor; however, Sample 2 only had a weak (4) “musty” odor. Panelists are asked not to discuss their experiences with each sample during the analysis. After all panelists have characterized each sample, a discussion follows to reach a consensus of descriptors for each sample.

Table 3. Flavor Profile Analysis data worksheet example

Sample	Descriptor	Intensity
1	Musty	4
	Earthy	6
2	Swampy	8
	Earthy	4
3	Musty	4

3.2.2.Selection and training of sensory analysts

Prior to the start of sampling, FPA panelists were trained to recognize using reference standards that represent tastes and odors that are found in drinking waters. Training was offered July 26, 2016 to thirteen panelists, six females and seven males, on Virginia Tech campus in Blacksburg, VA.

Training was conducted according to protocol from the handbook, *Flavor Profile Analysis: Screening and Training of Panelists* (Association, 1993). In the 8-hour training session, panelists tasted different solutions of sucrose, sodium chloride, caffeine, and citric acid to identify the four basic tastes. Panelists smelled different reference standards to recognize different odors, including geosmin, 2-methylisoborneol, dimethyl trisulfide, and hexanal (Table 1).

3.2.3.Institutional Review Board

The research study protocol for Flavor Profile Analysis was approved by the Institutional Review Board for Research Involving Human Subjects at Virginia Tech (IRB 16-671). Informed consent was obtained in writing in accordance with the IRB protocol. Subjects reported being 18 years or older, healthy, not pregnant, and not consuming food or beverages or smoking one hour before the sensory evaluation/training. The training panel was composed of 13 members (6 females) of ages 21-65. During a FPA session, 4-7 of the trained panelists participated.

3.2.4.Chemicals and reagents

Several compounds were purchased and used as odor references for the training (Table 5).High purity n-hexanal (CAS 66-25-1), (+/-) geosmin (CAS 19700-21-1), 2-methylisobrneol (CAS 2371-42-8), and dimethyl trisulfide (CAS 624 – 92-0) were purchased from Sigma Aldrich (St. Louis, MO) and used as reference standards (Table 1). The taste training involves practice sessions using the four basic taste reference standards: sweet-sucrose, salty-sodium chloride, sour–citric acid, and bitter–quinine. Sucrose (CAS 57-50-1) and sodium chloride (CAS 7647-14-5) were food grade products purchased at a grocery store. Food grade citric acid (CAS 77-92-9) and quinine (CAS 130-65-0) were purchased from Sigma Aldrich. The taste free water was Dasani bottled water purchased at a grocery store. Panelists tasted different concentrations of reference standards to recognize the intensity scale associated with each reference. The objective of this training is to enable panelists to assess any taste and odor at various intensities. The strength of a taste or odor is scored on a six-point scale as shown below in Table 2 and Table 4.

Table 4. Flavor Profile Analysis Intensity Scale

Word Scale	Numerical Scale	Example
Taste/Odor Free	**0**	Pure water
Threshold	**T**	Detect a trace of taste or odor but cannot describe
Very Weak	**2**	Detect a very weak taste or odor and can provide description
Weak	**4**	Detect and describe one or more weak tastes or odors. “Weak” is analogous to sweetness of canned fruits.
Moderate	**8**	Detect and describe one or more moderate intensity tastes or odors. “Moderate” is analogous to the sweetness of soda pop.

Strong	**12**	Detect and describe one or more strong tastes or odors. "Strong" is analogous to sweetness of jelly.

3.3. Gas Chromatography/Mass Spectrometry (GC/MS) Coupled with Head-space Solid Phase Microextraction (SPME)

3.3.1.Chemicals

Concentrated analytical standards of geosmin, 2-methylisoborneol, 2-methoxy-3,5-dimethylpyrazine, and 2,4,6-trichloroanisole, were purchased as 100 µg/L analytical standards (Sigma Aldrich, St. Louis, MO) and diluted to 20 µg/L with methanol (HPLC grade, EMD Millipore Corporation, Darmstadt, Germany) for use as stock standards. The 2,4,6-tribromoanisole standard is prepared by dissolving 97% pure solid in methanol. Standards are prepared by injecting stock standards in deionized water (≥16 MΩcm). The Carrier gas is ultra-high purity helium (AirGas, Radnor, PA), which passes through a three-stage filter before use.

3.3.2.Instrumentation

The gas chromatograph used is a Thermo Focus GC equipped with an Rxi-5Sil capillary column (30 m length, 0.25 mm internal diameter, 0.50 µm 5% polydimethylsiloxane film, Restek, Bellefonte, PA). Detection is provided by a Thermo DSQ II quadrupole mass spectrometer.

The manual injection solid phase microextraction fiber (Restek, Bellefonte, PA) is one cm in length with a 50/30 µm divinylbenzene/carboxen/polydimethylsiloxane coating.

3.3.3.Procedure

Samples are taken directly from the sampling sites in 60 mL glass volatile organic analysis vials with no head-space, and stored at ≤ 4 °C for less than two weeks before analysis. To prepare for analysis, 20 mL of sample is removed so that appropriate headspace is provided. Before an analytical batch, SPME fibers are conditioned in the GC inlet for 30 minutes at 270 °C with the GC oven set to 250 °C. Samples are then heated in a water bath to a temperature of 60 °C, and then the SPME fiber is inserted through the septum and allowed to equilibrate for 30 minutes.

The SPME fiber is then inserted into the GC injector which is heated to 240 °C and set to a splitless, 1.2 mL/min helium flow. The fiber remains in the injector for 5 minutes to desorb and clean. The column's initial temperature is set to 40 °C, which is held for 5 minutes. The temperature is then ramped at 8 °C per minute to 160 °C, and then at 20 °C per minute to 260 °C, which is held for one minute. The entire GC analysis takes 26 minutes. The MS transfer line was maintained at a temperature of 250 °C. Mass spectrometry is performed with selective ion monitoring for the compounds listed in (Table 5).

Table 5. GC/MS Odorants and Quantification

Odorant	CAS #	Retention Time, min	Selective Ion Monitoring, m/z
Geosmin	19700-21-1	21.37	41,112,125
2-Methylisoborneol	2371-42-8	17.55	95,107,108

2-Methoxy-3,5-dimethylpyrazine	92508-08-2	14.47	54,109,138
2,4,6-Trichloroanisole	87-40-1	19.36	195,197,210
2,4,6-Tribromoanisole	607-99-8	23.16	329,344,346

3.3.4. Quantification

Chromatography and spectrometry were processed using Xcalibur software (Thermo Fisher, Waltham, MA). Calibration is achieved using external standards with concentrations ranging from 0.5 to 250 ng/L. Linearity is difficult to achieve over the entire range for all odorants, so calibration curves are generated for low and high range. The lower range extends from 0.5 ng/L to 10 ng/L, and the higher range covers 10 ng/L to 250 ng/L.

3.4. Standard Methods for Known Odor-Associated Microorganisms

Relative changes in numbers of actinomycetes and fungi were quantified using plate counting techniques and changes in amount of algal growth were analyzed by monitoring chlorophyll concentrations in all river water, and sediment samples according to standard methods (APHA, 2012). Sediment samples were collected from the same contained as interstitial waters. Actinomycetes were enumerated on Actinomycete Isolation agar (AIA) (Method 9520; HiMedia, Mumbai, India.) and fungi were enumerated on rose-bengal agar (RBA) (Method 9610; HiMedia, Mumbai, India.). Briefly, water, sediment, and soil samples were transported to VT laboratory on ice and aliquots of each sample type were serially diluted in deionized water. 100 uL of the appropriate dilution step were spread on the appropriate agar and plates are incubated at 28 °C for AIA and 26 °C for RBA for 5 days. Colony forming units (CFUs) were counted immediately following incubation and data was presented as CFU/100mL water or CFU/g sediment. Algal biomass was quantified by extracting chlorophyll from sediment or water filters using vortexing and freezing in acetone (Method 10200). Extracted chlorophyll was quantified on a fluorometer - spectrophotometer (Thermo Fisher Inc., Waltham, MA, USA). Data were reported as mg Chl/g sediment or mg Chl/mL water.

3.5. Statistical Data Analysis

Multivariate analyses and statistical tests were performed using R version 3.4.2 (09/28/2017). Both qualitative analysis and formal statistical tests were used to evaluate the sensory, chemical, and biological data from the Dan and Smith rivers as summarized in the previous sections. The sensory data, FPA, were first evaluated using the R text mining package *(tm)* to identify and summarize (Table 6) the frequency each descriptor is reported. Visual representations (Figure 2) of the relative importance of each descriptor were created with the word cloud generating package, where larger text represents the most frequently used descriptor. Each of the 6 to 8 panelists might use one or several descriptors and intensities for each sample, the FPA data often had many different observations of odor for each sample.

The mean max FPA odor intensity is a derived value used in statistical analyses as a representative value of the overall odor intensity reported for each sample. The mean max FPA intensity was determined by first selecting the highest odor intensity reported by each panelist for a sample, then the mean of those intensities was taken to represent the overall intensity of odor detected in each sample. Levelplots that were created using the rasterVis and lattice packages to show the mean max FPA odor intensity for each sample (Figure 4 and Figure 5). On the horizontal axis of the levelplots are columns for each month of sampling, rows for each site on the vertical axis. Each square in the levelplot

contains the mean max odor intensity value for a sample and color coded such that the redder squares are samples with strong odor intensity, and blue squares have weak odors. For example, in Figure 4 the square in the first column, August 2016, and first row, Site 1 on the Smith River, is light blue, so the odor intensity for that column water sample can be determined using the scale on the left side of the graph as a value of about 3, a weak odor intensity. Levelplots are also used to show other chemical and biological data for which the columns and rows may represent a different variable, but the color scale will be the same such that red squares are always indicative high concentrations or values, and blue squares have low values.

Results from SPME-GC/MS quantification of earthy/musty odorants, geosmin (GSM), 2-MIB, TCA, TBA, MDMP, are also visualized using levelplots in Figure 7. The columns in these plots contain the monthly mean concentration of an odorant and rows represent each site. In Figure 7 the levelplot on the left shows that the interstitial water sample for geosmin (GSM) was high, colored red, for site 1 relative to site 3, in blue. Table 7 provides a summary of the number of samples analyzed by SPME GC/MS, number of samples that odorants were detected above and below method detection limit. The results presented here include the odorants quantified below detection limit as this value is taken as the best estimate for the instrument. IPMP and IBMP were also measured, but are excluded in the following analyses due to low frequency of detection.

Odor Activity Values (OAV) were calculated for each odorant by dividing the concentration measured by its respective odor threshold concentration. OAVs equal to one indicate the detected odorant was present at a concentration detectable for about 50% of the population; OAVs greater than one are much more likely to detected by humans. The derived OAVs were used to evaluate and compare the amount of odor contributed by each compound to the overall odor intensity for each sample. The levelplots in Figure 9 are structured similarly to Figure 7, except columns contain the monthly mean OAVs of each odorant for each site, in rows 1 to 12. In Figure 8 the last column in the plot on the left shows TCA in column water at site 6, the dark red square, had the highest mean monthly OAV. In several graphs the OAVs were transformed to a $\log_{10}$ scale before they were plotted. The distributions of all OAVs shown in box plots (Figure 8) are on a $\log_{10}$ scale.

The influence of the OAVs on the overall odor profile was evaluated by plotting $\log_{10}$ OAVs against mean maximum FPA odor intensity in scatterplots (Figure 10).Formal statistical tests were performed using the glm package to fit generalized linear regressions, in which mean max FPA odor intensity is the response variable, to $\log_{10}$ OAV for each odorant (Figure 10). Alpha value of 0.05 was used to determine a significant difference in slope of the line.

For further evaluation of OAVs and odor intensity, regression trees were created (Figure 12 and Figure 13) using the rpart package. Regression trees are similar to multiple regression models, in that Briefly, regression trees are fitted by stating that the continuous response variable, mean max FPA odor intensity, is to be estimated as a function of all explanatory variables: GSM, 2-MIB, TCA, TBA, and MDMP (Crawley, 2007).

The tree model shown in Figure 12 was fitted using binary recursive portioning; the data are sequentially split, at nodes, on the coordinate axes of the explanatory variables, odorant OAVS. The variable explaining the greatest deviance in the response variable is the root node at the top of the tree. From the root node the data are successively split, at nodes, for which the explanatory variable at the node maximally distinguishes the response variable in the left and right branches. The process is repeated recursively until no further reduction in deviance is obtained by a split, or too few data remain. For each sub-group which explains the greatest deviance. Displayed on the tree model at each node are the following: node number (1) and variable label (GSM); threshold value of that variable used to create

the split (GSM=0.27); the number of cases going into the split (n=152); and the mean value of the response variable, mean max FPA odor intensity, within that node. The mean value of the response variable (FPA odor intensity) above and below the node (GSM=0.27) are on the right and left branches of the tree.

Levelplots were also used to show the biological data results for Actinomycetes (CFU/L) in Figure 14. Due to the large range of results, the data were transformed on a $\log_{10}$ scale to better visualize variations in the results. Actinomycetes results ($\log_{10}$ CFU/L) were also plotted, explanatory variable, against geosmin and 2-MIB (ng/L), response variables, in Figure 15.

4. Results

4.1. Odor characteristics

The odor characteristics of the Dan and Smith rivers were evaluated by Flavor Profile Analysis (FPA) as described in the methods section 3.2 of this chapter. Panelists analyzed the odors of each sample and provided individual descriptions and intensities for each sample. Often more than one descriptor was used to characterize the aroma of the sample, as shown in Table 3. The data worksheets from all panelists were collected after each FPA session then the descriptors and intensity were recorded for each sample. The FPA method provides a comprehensive sensory evaluation of each site for all 13 months from which seasonal and spatial trends, as well as differences for the rivers were examined using R Studio for statistical analysis.

The main goal of the FPA data was to characterize the odors for each Dan and Smith River site to establish a baseline and identify river sites with more intense odors under normal conditions. As the FPA method uses both odor descriptors and an intensity rating the data provides information on what type (descriptors) and how much (intensity rating) of odors are present. In total 2915 descriptors and intensities were recorded for all river water samples over 13 months of sampling. Although the FPA panelists were trained on sensory analysis, as humans have unique olfactory senses and interpretations, which lead to a variety of descriptors in the data. Of 2915 observations there were 109 unique descriptors, so despite the differences in human odor perception, there was some uniformity in the data.

The FPA data was first analyzed for frequently used odor descriptors. Word clouds were created in Figure 2 and Figure 3. Descriptors of odors in Column (left) and Interstitial (right) water samples by FPA Figure 3 to visually assess the frequency of different descriptors for the samples. The word clouds show odor descriptors for all water samples in Figure 2, as well as separated by sample type, column water on the left and interstitial water on the right Figure 3. The frequently used descriptors are in larger text, such as “earthy”, and smaller words like “moldy” are less often recorded. As mentioned previously, sometimes panelists use different words to describe similar odors, for example “musty” is used often whereas the similar odor, “moldy” was used less often. It is apparent from the three figures that “earthy” and “musty” odors are present regardless of water type. Also there are certain odors that are unique to each sample type such as, “swampy” and “smoky” for interstitial waters, as well as “grassy” and “chlorinous” in the column water.

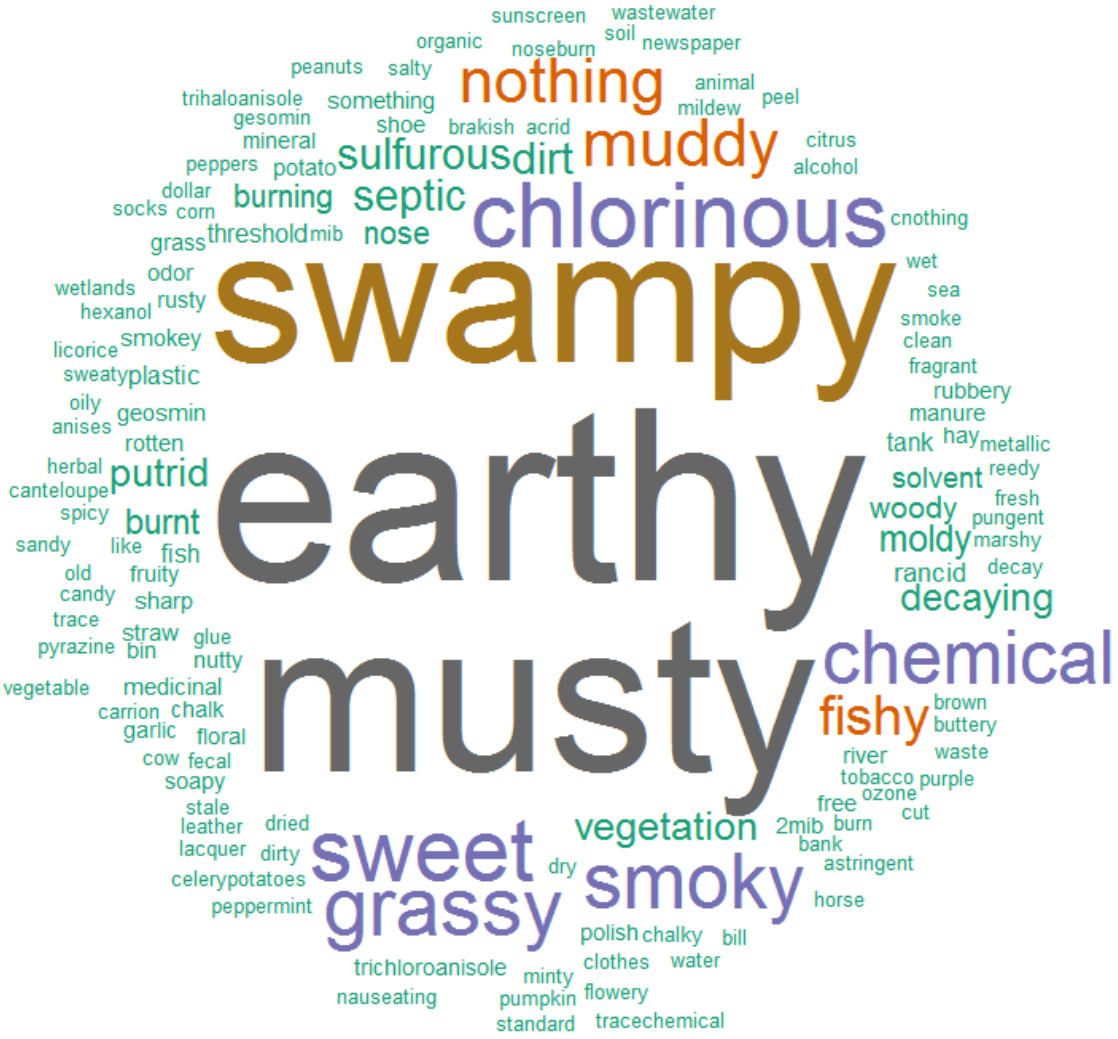

Figure 2.Odor descriptors from FPA for all samples

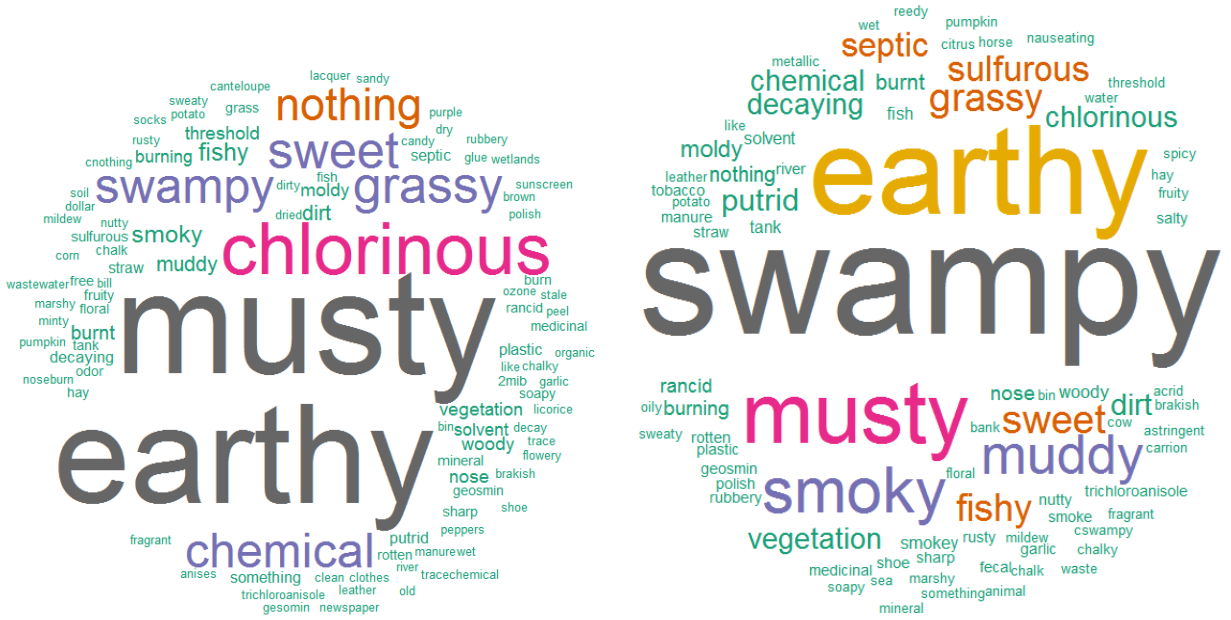

Figure 3. Descriptors of odors in Column (left) and Interstitial (right) water samples by FPA

The five most frequent descriptors were: “earthy” (493); “musty” (424); “chlorinous” (153); “swampy” (404); and “smoky” (144). The FPA data was then subset into categories of column water or interstitial water samples, to distinguish if certain descriptors are more common in one water type. The frequency of these descriptors in each water type, column or interstitial are summarized below in Table 6. Although “earthy”, “musty” are commonly used to describe interstitial and column water, column water is described as “chlorinous” about three times more often than for interstitial water. Similarly, “swampy” and “smoky” odors in interstitial water samples are described almost five times more often than in column water.

Table 6. Most frequently used odor descriptors

Water type	“Earthy”	“Musty”	“Chlorinous”	“Swampy”	“Smoky”
Column	239	258	117	75	25
Interstitial	254	166	36	329	119
Total number of descriptors	2915				

As several panelists analyzed each sample, often attributing more than one descriptor and intensity for a given sample, simplification of the FPA data was needed so that each sample corresponded with a single value for odor intensity. Instead of using the mean intensity of all odors described by for each panelist, or an overall mean of the intensities of a sample for the session, a mean of maximum intensities was calculated. Briefly, for each sample the maximum odor intensity described by each panelist was selected and the mean of those values was used to calculate a new variable, the mean maximum odor intensity. This value was used in the following analyses to represent the amount of odor present at each site for each month of sampling.

The mean maximum FPA odor intensity value can be visually evaluated temporally and spatially using the levelplots show in Figure 4 and Figure 5. The columns on the horizontal axis are each month of sampling, August 2016 to September 2017, and the rows represent river sites in ascending order, Site 1 is the most upstream site on the Smith River and Site 12 is furthest downstream on the Dan River. Each square contains the mean max FPA odor intensity for the sample collected at each site, for each month. Each square in the levelplot represents the mean maximum odor intensity and color coded such that blue squares have a weaker odor intensity; pink squares have moderate odor intensities; and the red squares have the strongest odors. Variations in odor intensity for each site and month are shown in levelplots for both column water (Figure 4. Levelplot of column water mean max odor intensity and interstitial water (Figure 5)

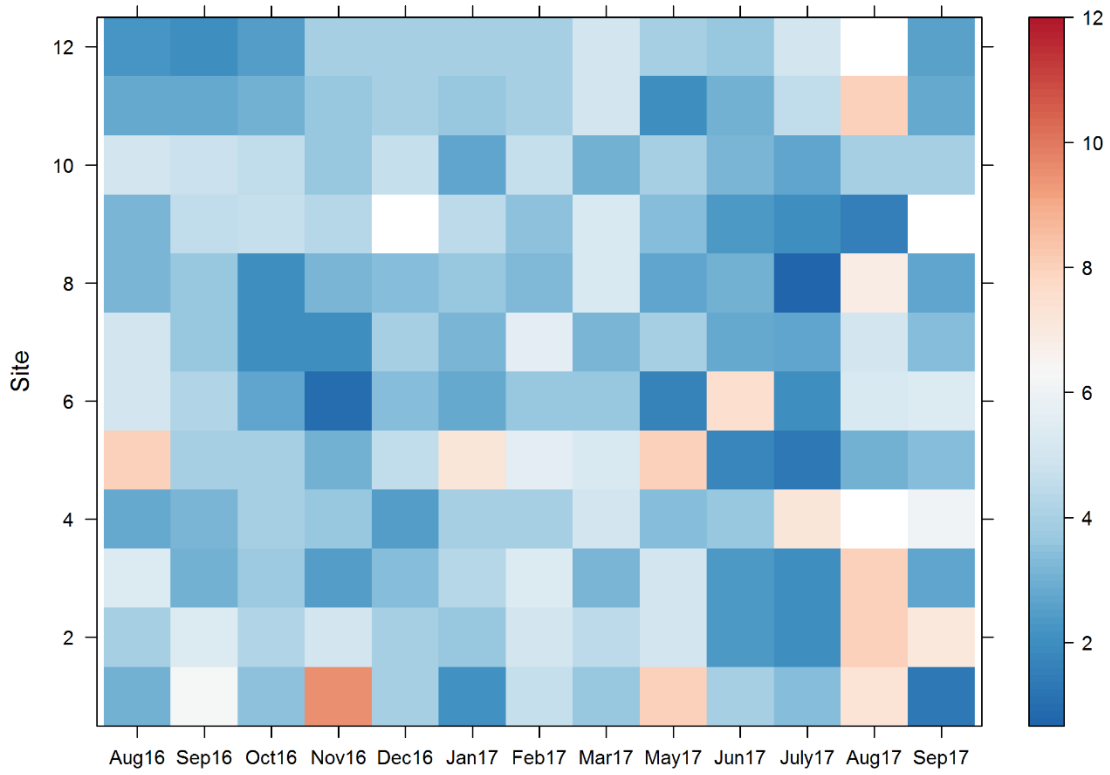

Figure 4. Levelplot of column water mean max odor intensity

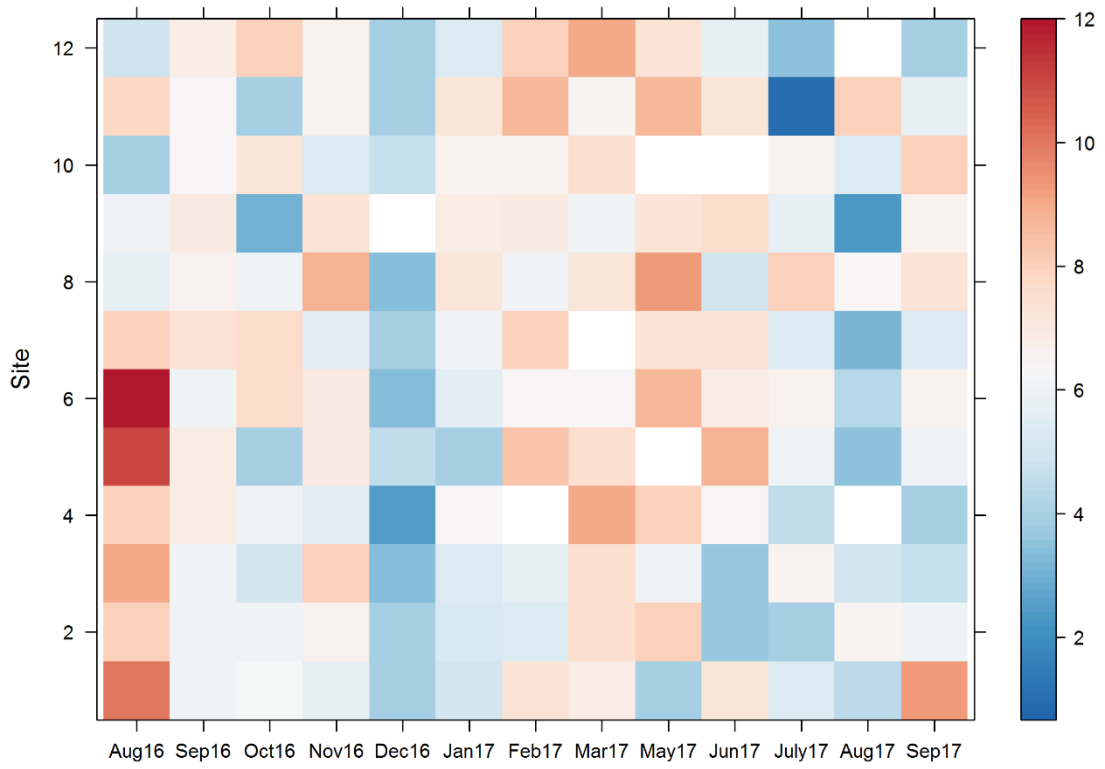

Figure 5. Levelplot of interstitial mean max odor intensity

Comparing monthly results, between columns, shows no obvious seasonal trends in which a particular month, or several months, experienced exceptionally strong odors. The highest odor intensity value for column water was at 9.5, moderate to strong, and occurred at site 1 in November of 2016. The highest odor intensity value for interstitial waters was 12, strong, in August, 2016 at site 6 on the Dan River.

The results displayed in Figure 4 and Figure 5 show that overall interstitial waters have a stronger odor intensities than column water samples. This difference was expected as interstitial samples are collected from the river bed and unfiltered for FPA analysis. Water samples prepared for

FPA analysis in Figure 6 show the column water (furthest left) to be relatively clear, while the interstitial water samples (middle and right side) are quite turbid.

Figure 6. Column and Interstitial water samples prepared for FPA analysis

4.2. Odorants by SPME GC/MS

Odorants were frequently detected by SPME- GC/MS near or below estimated method detection limits (Table 7. Samples analyzed for earthy/musty compounds. Mean concentrations of geosmin, 2-MIB, TCA, and TBA are shown in levelplots below (Figure 7). IPMP and IBMP were detected in collected samples very infrequently and are not included in the levelplots below. Overall, mean odorant concentrations overall were higher in interstitial waters than their corresponding column water. Geosmin was the dominant odorant detect in interstitial water samples, especially in the upper Smith River sites (sites 1 and 2).

Table 7. Samples analyzed for earthy/musty compounds

Number of samples, n	GSM	2-MIB	MDMP	IBMP	IPMP	TCA	TBA
Total samples	347	347	347	347	347	347	347
Samples analyzed for odorants	321	321	321	173	200	321	321
Odorants not detected	85	237	138	158	191	187	146
Odorants detected, < MDL	8	0	125	5	1	100	120
Odorants detected, > MDL	228	84	58	10	7	34	55

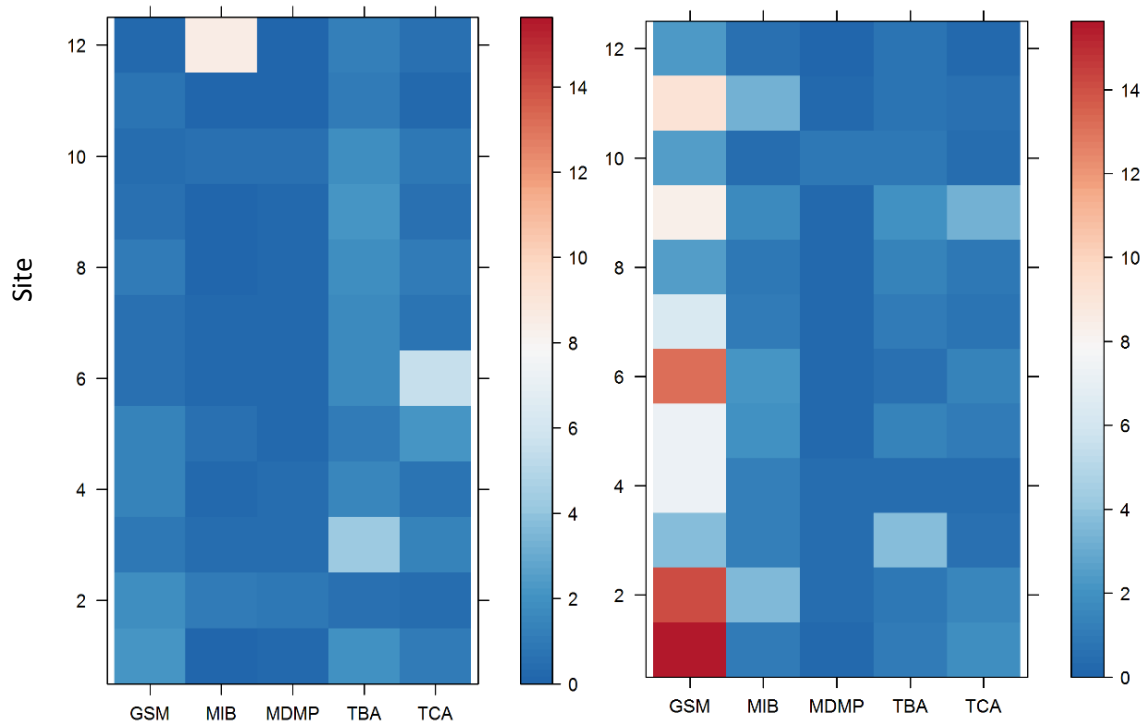

Figure 7. Mean concentrations of earthy/musty odorants for Column (left) and Interstitial (left) by site

4.3. Odorants Activity Values (OAVs)

Odor activity values, (OAVs), are calculated by dividing the concentrations of different odorants, detected in river water samples, by their respective odor threshold concentrations (Table 1). OAVs can be used to analyze the contribution of the odorants to the odor profile of a sample (Guo et al., 2016; Li, 2016). Because there are a range of reported odor threshold concentrations in water, the geometric mean of the reported values was used to determine the OAVs; for example, the OTC for geosmin is 1-10 ng/L, so the geometric mean of the two values is 3.16.

The distribution of OAVs for detected odorants in all samples are shown below in Figure 8 transformed on a logarithmic scale. The number of samples that each odorant was detected in both interstitial and column water samples is indicated on the axis above the boxplot. The median OAVs for geosmin, 2-MIB, and MDMP are close in range, 0.63, 0.65, and 0.77 respectively. However, the ansioles have higher medians, 2.0 for TCA and 1.2 for TBA. These values should be interpreted with care as the number of samples each odorant was detected is very different. In Figure 8, TCA appears to be a major driver of odor activity, but TCA was not detected by SPME- GC/MS in the samples as often as geosmin. This may be due to the detection limits of SPME GC/MS, which can identify geosmin at lower concentrations than is possible for detecting haloanisoles (Table 7).

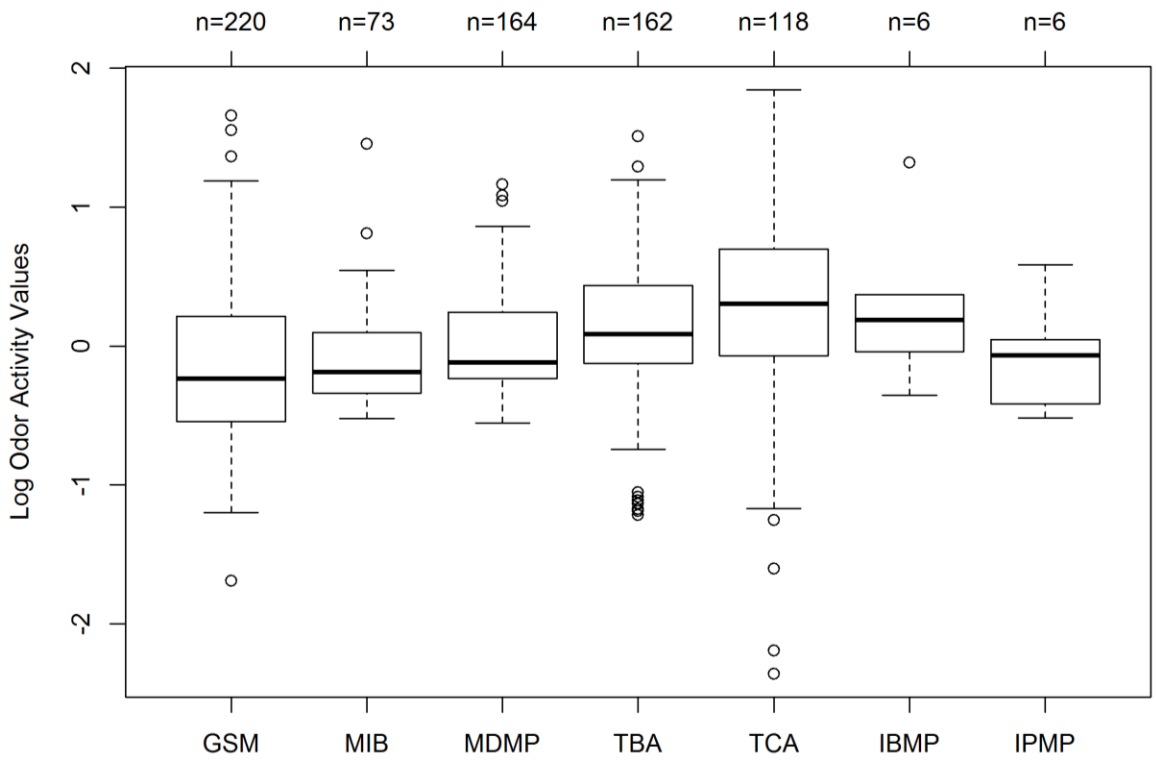

Figure 8. $\log_{10}$ OAVs of earthy/musty odorants

Low OAVs were detected in column water samples, with the exception of TCA (70.2), TBA (32.3), and 2-MIB (28.5) at sites 6, 3, and 12, respectively. The OAV for geosmin in interstitial waters was high at several sites, the highest at sites 1 (45.8) and 2 (35.6) on the upper Smith, as well as sites 6 (15.5) and 11 (23.2) on the Dan River. The OAVs were also high for TBA (19.5) and 2-MIB (6.5) at Smith river sites, and on the Dan River at site 9 TCA (59.4). Although there is not a particular site, nor odorant with a consistently high OAVs, often one or more compounds were detected at the same site, which could impact the overall odor intensity of the sample.

The OAVs of the mean concentrations of earthy/musty odorants at each site, for 13 months of sampling are shown in Figure 9 as levelplots for column (left) and interstitial water samples (right). Columns contain the OAVs for geosmin, 2-MIB, MDMP, TBA, and TCA at each site. IPMP and IBMP were detected in collected samples very infrequently and are not included in the levelplots below.

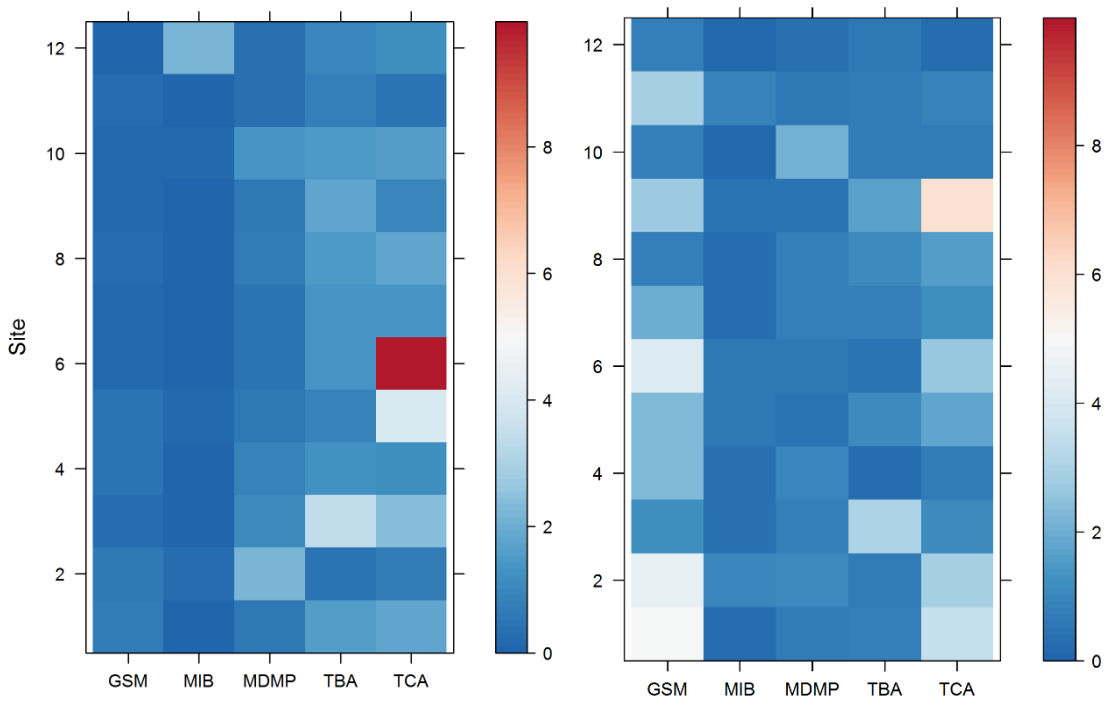

Figure 9. Odor Activity Values of earthy/musty odorants in Column (left) and Interstitial (right) by site

The relationship between mean max odor intensity and OAVs detected in each sample were evaluated using regression analysis techniques. Shown in Figure 10 are mean max odor intensity values (FPA) plotted against log transform OAVs for geosmin and 2-MIB (left) and TCA, TBA, MDMP (right). The regression line for each relationship is also displayed in each graph color coded for each odorant. Additional linear regression analyses were performed to evaluate the influence of each compound's OAV on FPA at each site. The results of the regression showed no statistically significant relationships, with the exception of interstitial water samples at site 4 ($p=0.008$).

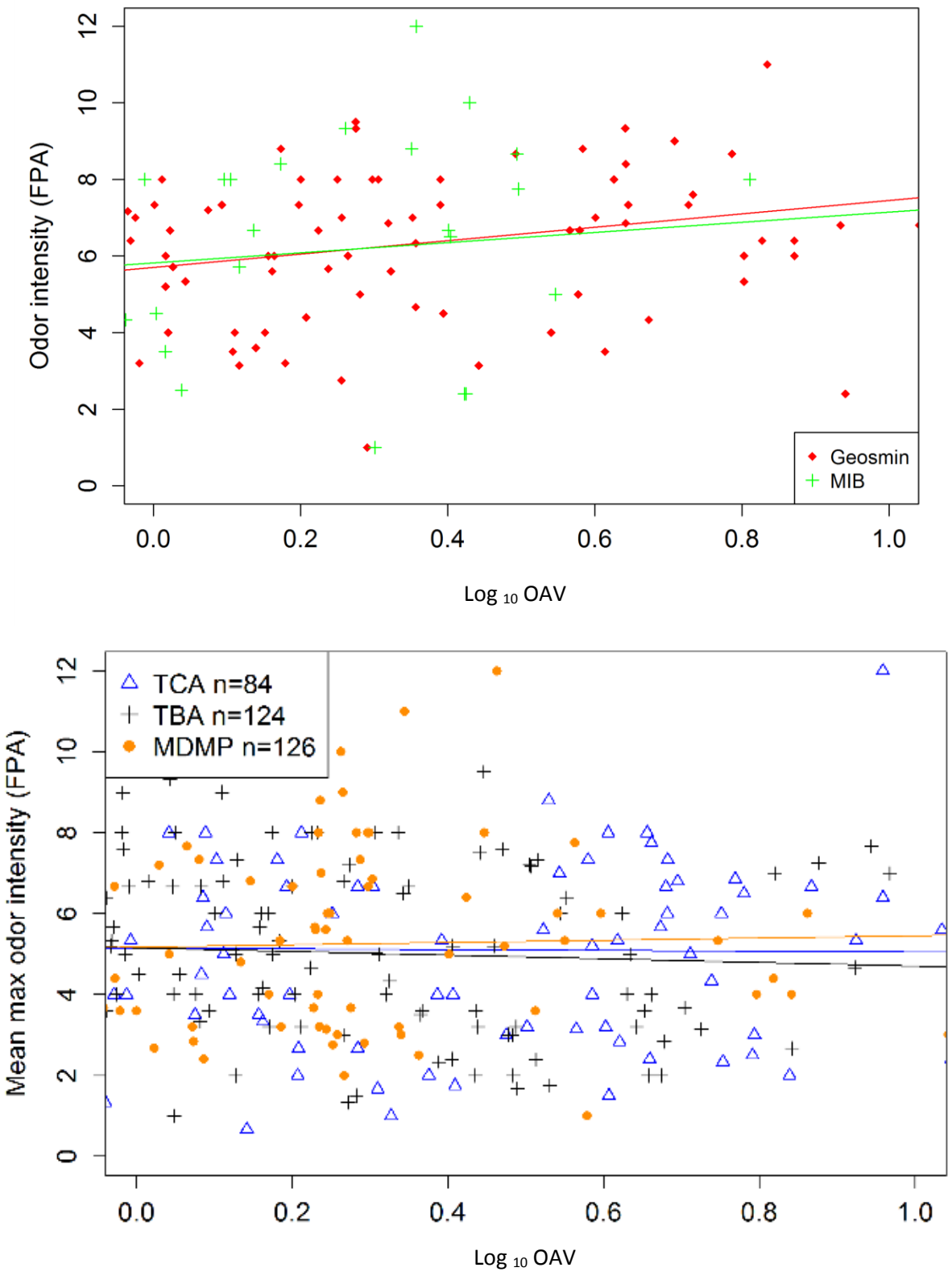

Figure 10. Regression of mean max odor intensity by log OAVs of different earthy/musty odorants

Multiple odorants were often detected in one sample, so the overall odor intensity for a given sample may be the combined effect of all odorants. Two other variables were derived to evaluate the relationship between odor intensity and the total amount of odorants in each water samples; sum of all OAVS, and sum of all odorants detected (ng/L). The sum of the OAVs and sum of all detected odorants, transformed on a $\log_{10}$ scale, for each water sample are plotted against mean max odor intensity below in Figure 11. The regression line modelling the influence on odor intensity by the summation of the odorants are also included in the graph. The results of the linear regression, summarized in Table 8, show model significance ($p>0.05$), and predict that mean max odor intensity can increase by 0.026 and 0.035 for each increase in the sum of OAVs and sum of odorants detected (ng/L). The data included in the model (Figure 11) are from both interstitial and column water for all months and sites sampled.

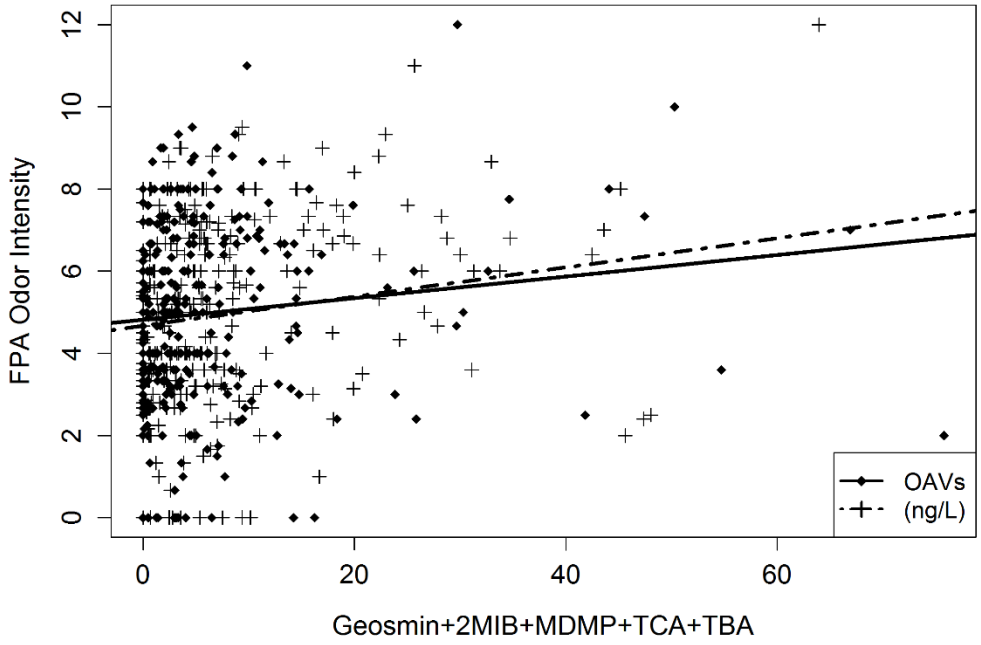

Figure 11. Regression of sum of odorants (Log_{10}) versus odor intensity for all water samples

Table 8. Summary of linear regression for sum odorants influence on odor intensity

Variable	Slope (estimated increased FPA odor Intensity per unit increase in X variable)	p-value	R^2	Mean
Odor Activity Values	0.026	0.047	0.01	5.77
Detected Amounts (ng/L)	0.036	2.4e-06	0.07	8.53 (ng/L)

Tree regression analyses were performed to evaluate which odorant OAVs were the most important explanatory variables in terms of in odor intensity. The tree regression models were built by fitting all odorant OAVs to explain the response of FPA odor intensity for column water (Figure 12) and interstitial water samples (Figure13). No distinction was made between different sites; rather the purpose of the analysis was to elucidate which odorant OAVs, and at what level, can drive changes in the odor intensity.

The tree regression shown in Figure 12 contains all odorant OAV and FPA results for column water samples only. The root node for the regression tree is with a mean odor intensity of 3.9. The model indicates that geosmin is the main driver of odor intensity in the column water. The branch to the right evaluates the data where geosmin is greater than 0.27 OAV; the left branch evaluates the data where geosmin OAV is lower. On the right branch, when geosmin's OAV exceeds 0.27, a third node is created where the mean FPA odor intensity is 4.5. This node can be further split based on TBA at 2.9 OAV. The interpretation is that when the OAVs for geosmin is greater than 0.27 and TBA OAV is less than 2.9, the mean odor intensity increases to 4.9. Two of the terminal nodes on the right side of the root, higher geosmin OAV conditions, have the two highest mean odor intensity. The highest mean odor intensity, 5.8, occurs when geosmin and TCA are greater than 0.27 and 0.16, respectively, but TCA is lower than 2.9. The shorter summary is that the combined effect of geosmin, higher than 0.27, plus TCA, higher than 0.16, and without TBA, in a sample results in the odor intensity of 5.8 (slightly more than weak odor) on the FPA scale.

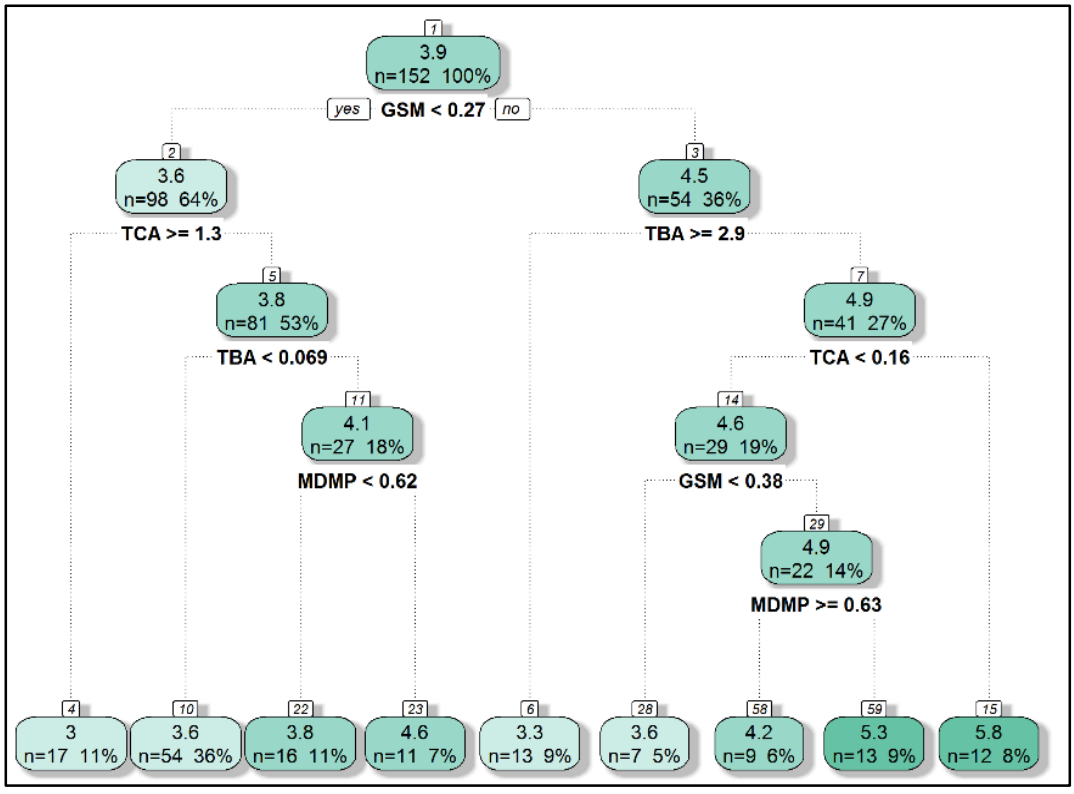

Figure 12. Regression Tree of OAV influence on odor intensity (FPA) in column water

The odorant OAVS and odor intensity results for interstitial water samples are evaluated via tree regression and the model is shown below in Figure 13. Geosmin is also the most important variable driving the odor intensity, but the mean odor intensities are higher for interstitial waters. The overall

mean is 6.4. Again, geosmin is the primary driver for the changes in FPA, with a split occurring at an geosmin OAV of 2. Similar to the column water in Figure 12, geosmin is the key driver and TBA is the next important variable. When geosmin OAV is higher than 2 and the mean odor intensity increases to 7.2. Again there is an unusual trend where the terminal node with the highest mean odor intensity (8.6) occurs for high geosmin, 2-MIB greater than 0.52, but TBA is less than 1.1. When 2-MIB OAV is lower than 0.52 the mean FPA decreases to 6.9, suggesting that in the absence of TBA, 2-MIB may be adding to the overall odor intensity.

Another unexpected trend is shown in the model under high geosmin and TBA greater than 1.1; a terminal node is reached and the mean odor intensity decreased to 5.4, lower than for low TBA conditions. This may be explained by examining the left branch of the root node, low geosmin conditions for which TBA is the most important explanatory variable. The terminal node for high TBA but low geosmin conditions has one of the highest mean odor intensities of 6.8. The model shows that, similar to the column water, geosmin and TBA are key drivers of FPA odor intensity for interstitial water.

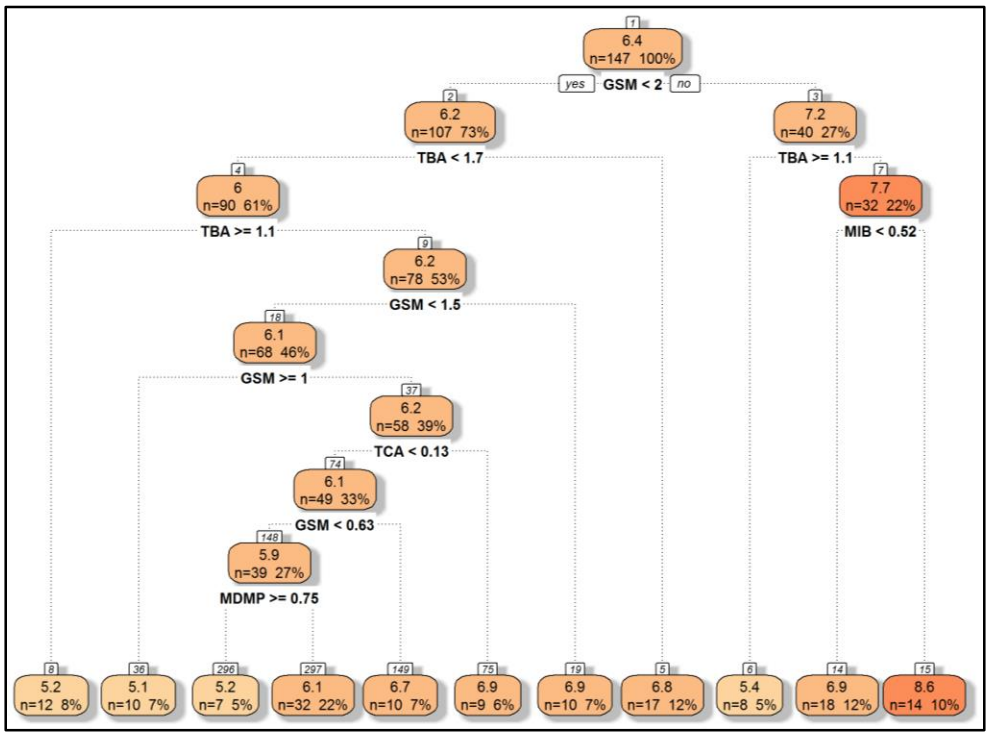

Figure 13. Regression Tree of OAV influence on odor intensity (FPA) interstitial water

4.4. Biological population characteristics

Actinomycetes concentrations (column water, CFU/L; interstitial water, CFU/g) were analyzed in both interstitial and column water samples at each site and results are shown in the levelplot below (Figure 14). The columns contain the concentrations of actinomycetes on a log 10 scale for each month for each site. At the time this report was written the results from September 2017 were not available. Actinomycetes levels in the column waters were highest in August of 2016 and 2017 at sites 10 (10810 CFU/L) and 12 (9910 CFU/L) on the Dan River. There were no sites where this range was detected until August, 2017 at site 6 (15385 CFU/L) on the Dan River.

The results from interstitial water samples show higher overall levels of actinomycetes, but did not follow similar temporal nor spatial trends as observed in the column water samples. The highest levels of actinomycetes were detected at site 11 in October, 2016 (20,000 CFU/g) through January, 2017 (41,153 CFU/g), which was the highest for all samples analyzed. However, the column water at site 11 did not follow the same trend. In March, 2017 at site 11, actinomycetes jumped from very low in February (3,076.9 CFU/L) to (6,923 CFU/L), the highest level relative to column water samples for that month, but this increase did not correspond with an increase in the interstitial waters.

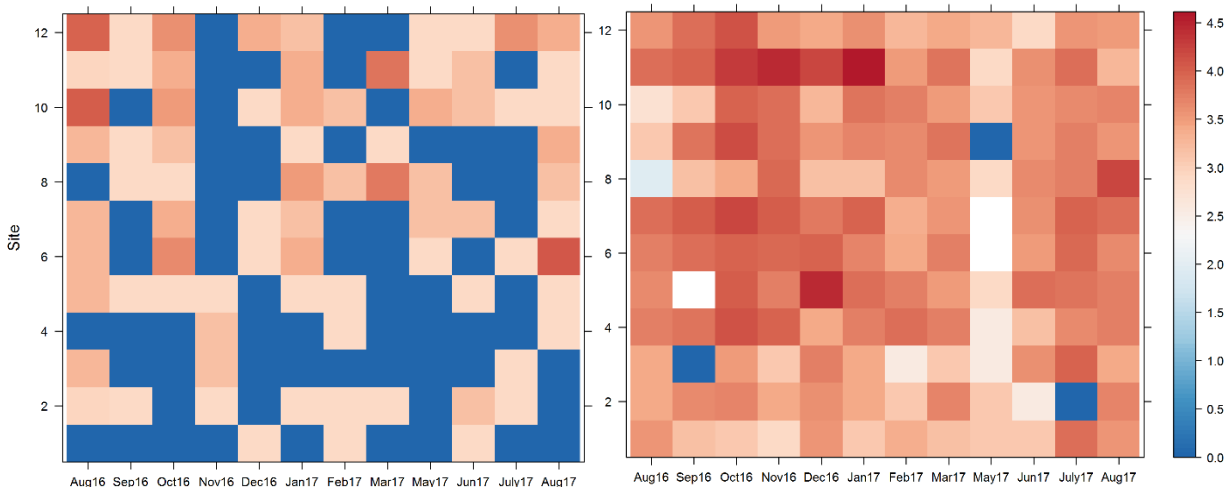

Figure 14. Levelplots Actinomycetes (log_{10}) in column CFU/L (left) and interstitial CFU/g (right) samples

The relationship between geosmin, 2-MIB and actinomycetes in all water samples was evaluated to identify if an increase in actinomycetes corresponded with an increase in the odorants. The results of this analysis are shown in (Figure 15), where actinomycetes concentrations are log_{10} transformed on the vertical axis, and geosmin and 2-MIB in ng/L are given on the horizontal axis. No general trends are obvious in this plot, except that low concentrations of the odorants corresponds with low actinomycetes, as shown by the points on the lower left of the graph. There are several extreme values for geosmin (144.7 and 112.8 ng/L) and 2-MIB (110 ng/L) that do not correspond to high levels of actinomycetes. Actinomycetes were not detected in the samples when the odorants were detected at extreme values.

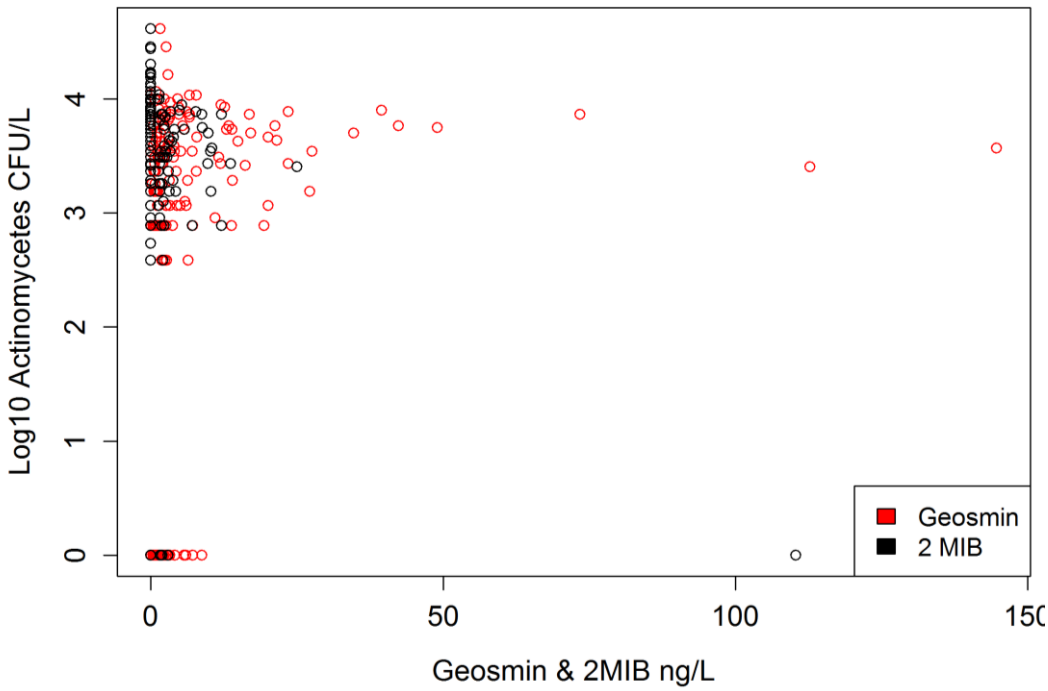

Figure 15. Actinomycetes versus Geosmin & 2- MIB (ng/L)

5. Discussion

The motivation of this research was to identify the source, or sources, of the unprecedented earthy/musty T&O problems that had previously occurred in the Dan River in southwest Virginia. Beginning in August, 2016 column and interstitial waters were collected from twelve different locations over about 50 miles of the Dan and Smith Rivers. Water quality, biological, and sensory analyses were performed for each sample with the goal of characterizing each river under 'normal' conditions, meaning not during a taste and odor event. Although no considerable T&O events happened during the time of this study, the research design and results provide guidance for the local communities, as well as other communities, in the event of future T&O problems.

The odor characteristics of the water sampled from the Dan and Smith Rivers were frequently described as "earth"," musty", and "swampy" for interstitial water samples (Figure 3). These results from Flavor Profile Analysis are consistent with the odors that were reported to be present in the drinking water during the two T&O events in 2015. The intensity of the odors differed between samples types, column water often had low to no odor intensity, while interstitial waters commonly had moderate to strong odors. This was not surprising as the interstitial water samples came from water at the river bed and were often highly turbid, while column water samples were usually quite clear.

On several occasions interstitial waters had high or low odor intensity, but not a corresponding intensity in column water at the same site. This was the case in August 2017 where the column water samples from the upper Smith River, sites 1-3, had the highest odor intensities relative to other sites that month (Figure 4). However, the interstitial waters at the sites were not graded as relatively high (Figure 5). Furthermore, there were no clear odor intensity trends in terms of months or river sites which might help in a prediction of where or when odors would be elevated.

Several organic odorants with earthy/musty properties that have been identified in the literature for causing T&O events were detected in samples from both rivers (Table 5). Odorant Activity Values were derived to evaluate the contribution of each odorant to the odor profile of the water samples. When odorants were detected in the water samples, they were at or very close to their odor threshold concentrations. The levels detected during this project are not reflective of a T&O event, for which much higher levels are reported. For example, Li reports that during a T&O event in the Yellow River 2-MIB and geosmin were detected in the ranges of (5.77-21.12 ng/L) and (2.26- 9.73 ng/L) respectively (Li, 2016). The median 2-MIB and geosmin detected from the Dan and Smith river samples were, respectively, 0 and 1.05 ng/L. Comparing odorants at each site showed that TCA was high at site 6 in the column water (32.4 ng/L), but not high in the interstitial water samples, and the mean maximum odor intensity was a 2 on the FPA scale (very weak).

The influence of OAVs on odor intensity was evaluated by tree regression analysis (Figure 12 and Figure 13) showing that geosmin was the key driving variable in terms of odor intensity for both water types. The tree regressions predict that when certain OAVs are present in the same samples, such as geosmin and 2-MIB, the highest mean odor intensity is expected. However, formal statistical testing did not support that high odorant OAVs resulted in high odor intensity. Instead, the linear regression revealed no significant change in odor intensity by odorant OAVs (Figure 10). These results were not entirely surprising as odorants were not often detected above their odor threshold concentrations. The odors characterized by FPA were frequently weak for column water samples, and the higher intensity odors in the interstitial waters could be caused by odorous compounds which were not measured. For example, 'swampy' was a descriptor often reported for

interstitial waters; however, this research focused on earthy/musty odorants, so odorants such as sulfides, which are characterized as 'swampy', were not included in SPME GC/MS analysis.

The approach for tracking the source of earthy/musty odorants was to monitor the microbial communities along the rivers and identify where odorants could be produced biologically; such as by actinomycetes who are known to be associated with of geosmin and 2-MIB in freshwater systems. Levelplots of actinomycetes show that the detected levels varied monthly and by site with no consistent trends, other than being at higher levels in the interstitial samples (Figure 14). Column water samples from the Dan River were often higher in actinomycetes levels than the Smith River, which suggests a potential production of geosmin and 2-MIB at several sites in the Dan River. However, high levels of actinomycetes did not correlate with high levels of either odorant (Figure 15). Perhaps water quality conditions enabled biological communities to produce large amounts of odorants that caused the T&O events of 2015, but without a similar event during this study, it is not possible to speculate on the role that microbes might have played.

6. Conclusions

- "Earthy" and "musty" odors were detected in both column and interstitial water samples from the Dan and Smith Rivers (Figure 2 and Figure 3).
- Odors were often reported as moderate to strong (Table 3) in interstitial water, and weak or absent in column water samples.
- No river site had consistently higher odor intensities relative to other sites during the thirteen months of sample collection (Figure 4. Levelplot of column water mean max odor intensity and Figure 5).
- The highest odor intensities in the interstitial water samples were seen in August, 2016 for sites 1 to 5 (Smith River) and site 6 (Dan River). This trend did not occur in subsequent months.
- Site 5 near Eden, NC on the Smith River had higher odor intensities in the column waters more often than other river sites.
- Earthy/musty odorants were detected in both column and interstitial water; most often detected at or below their odor threshold concentrations (Table 1, Figure 7, and Figure 8).
- Geosmin most influenced odor intensity for both column and interstitial water (Figure 12 and Figure 13).
- No statistical relationship between odorant activity values and odor intensity detected (Figure 10).
- Actinomycetes were detected in both column and interstitial water in both the Dan and Smith Rivers (Figure 14); no statistical relationship was seen between level of actinomycetes and geosmin or 2-MIB (Figure 15).

Chapter 2: Multivariate Analyses and Statistical Testing

1. Introduction

In recent years, communities that use the Dan River as a source water have experienced two severe outbreaks of strong, unpleasant tastes and odors (AWWA) in their drinking water. The T&O events both lasted for several months during which strong 'earthy', 'musty' odors were reported in the water, but the source was not identified. The first T&O event began in early February, 2015 and coincided with an algal bloom in the Dan River that was observed under the ice near the Danville raw water intake. Initially the algal bloom was thought to be the cause of the odors, but even after the bloom dissipated, the unpleasant odors persisted until May 2015. The second T&O was first noted in October, 2015 and did not coincide with any observed algal blooms. The second T&O event was mitigated to an extent by additional water treatment process at the Danville water treatment plant. In September 2015, the city of Danville upgraded the existing powdered activate carbon (PAC) feeder system, due to insufficient capacity of the older system to provide the required dose for odorant removal. Although to date there has not been a subsequent T&O in the area, and there were no detected public health risks, there is a great desire to determine the cause of the odors in order to prevent future T&O outbreaks.

Years of industrial and agricultural activity in the Dan River area may have contributed to the water quality issues that stimulated the recent T&O events. Industrial discharges into source waters have been identified as sources of T&O compounds, such as 'sweet', 'buttery' diacetyl from paper manufacturing (Boleda, 2007). Lumber, wood preservatives, and textile manufacturing discharges may continue odorant precursors, such as halophenols which can be microbial methylated into extremely odorous haloanisoles (Nystrom, 1992; Hill, 1995; Zhang, 2016). he addition of phosphorus, usually the limiting growth nutrient in fresh waters, can have major effects on algae, bacteria, and fungal communities, thereby contributing to downstream T&O issues.

In addition to the legacy of industrial pollutants, approximately 39,000 tons of coal ash was spilled into the Dan River near Eden, NC from a Duke Energy coal ash pond. The spill began on February 2, 2014 and caused visible grey discoloration of the Dan River from the spill site to far downstream. The city of Danville withdraws its raw drinking water downstream of the Duke Energy site and responded quickly to the coal-ash spill. The utility effectively managed to provide drinking water that never exceeded state and federal limits for finished drinking water. The Dan River and Danville's treated water were monitored for coal-ash constituents daily after the spill; later the frequency of sampling decreased as results from treated water did meet drinking water regulations.

As historical records do not show any past algae, fungi, or bacteria related T&O events before the two in 2015, there is not one obvious starting point to draw conclusions about the odor sources. However, drawing on research of other earthy/musty T&O events provides guidance by which this project is designed. The objectives include analyzing water quality and ecological communities in the Dan River to evaluate the impact of industrial and anthropogenic sources that may have resulted in the taste and odor events. The Smith River, which begins at Philpott Dam and flows through several towns, joins the Dan River in Eden, NC at the Towns' Creek confluence. The Smith River provides a control to compare the impacts of the coal-ash spill on the Dan River. Interestingly, the Smith River was reported

as having the strongest odors during the two, severe T&O events of 2015 (Johnson, 2016, personal communication).

2. Literature review

Long-term, water quality monitoring is a good approach to better understand the chemical and biological variations of a river system, but it also generates a large data set which can be difficult to interpret (Vega, 1998; Feher, 2016) Multivariate statistical techniques are valuable for evaluating temporal/spatial trends, and interpreting large, complex water quality and ecological data sets generated at different locations. Applications of multivariate techniques, including cluster analysis (CA) (Vega, 1998), principal component analysis (PCA), and factor analysis (FA) can help to interpret complex data and possibly identify factors that are most influential (Everitt, 2011).

Cluster and principle component analyses are often used as an accepted unbiased method that can elucidate associations between different variables or samples. Vega et al. (1998) explains that hierarchical agglomerative cluster analysis shows groupings of samples based on their inter-sample similarity, which provides an overall comparison of the similarity between samples in the data set. PCA is useful for reducing the dimensionality of the data set by explaining correlations between a large amount of variables in terms of a smaller number of underlying factors, principle components, while not compromising much information (Crawley, 2007).

Vega et al used PCA, CA, and analysis of variance (Vega, 1998) to investigate impacts of pollution events and seasonal trends on the water quality on the Pisuerga river, a river in the Duero basin in Spain. Water samples were collected every three months from three sampling locations for two and a half years and were analyzed for 22 physical and chemical parameters. Exploratory analyses, including boxplots, ANOVA, PCA and CA were used to try distinguishing sources of variation in the water quality. The combination of these methods allowed for the identification of spatial and temporal trends affecting the river water, including pollution from domestic wastewater discharges and seasonal changes in the flow rate, such as high flow rates during the winter months.

3. Methods & Materials

3.1 Water Sampling

Dates of the site tour and monthly sampling trips are summarized below in Table 9. Samples were collected monthly from August, 2016 until September, 2017 with a total of thirteen trips. There was not a trip in April, 2017 due to unsafe river conditions caused by several days of heavy rain in the area. In the event of an outbreak of odors, extra sampling trips would have been made, but this did not occur during the project timeline.

Table 9. Dates of site tour and sampling trips

Days and month	Year
July 15 site tour	2016
August 16-17	2016
September 14-15	2016
October 4-6	2016
November 15-17	2016
December 16-18	2016
January 10-12	2017
February 3-5	2017
March 21-22	2017
May 16-17	2017
June 13-14	2017
July 10-11	2017
August 7-8	2017
September 19-20	2017

Samples of water are collected at twelve sites along the Dan and Smith Rivers either from a boat or on a bridge, as well as two drinking water treatment plants that source water from the Dan River, Danville and South Boston. At boat sites, testing for water quality and biological constituents are performed whereas from the bridge sites, only water quality is characterized. Figure 16 shows the boat sites as triangles, bridge sites as circles, and the drinking water treatment plants are rectangles. A description of the sites and methods used at those sites follows.

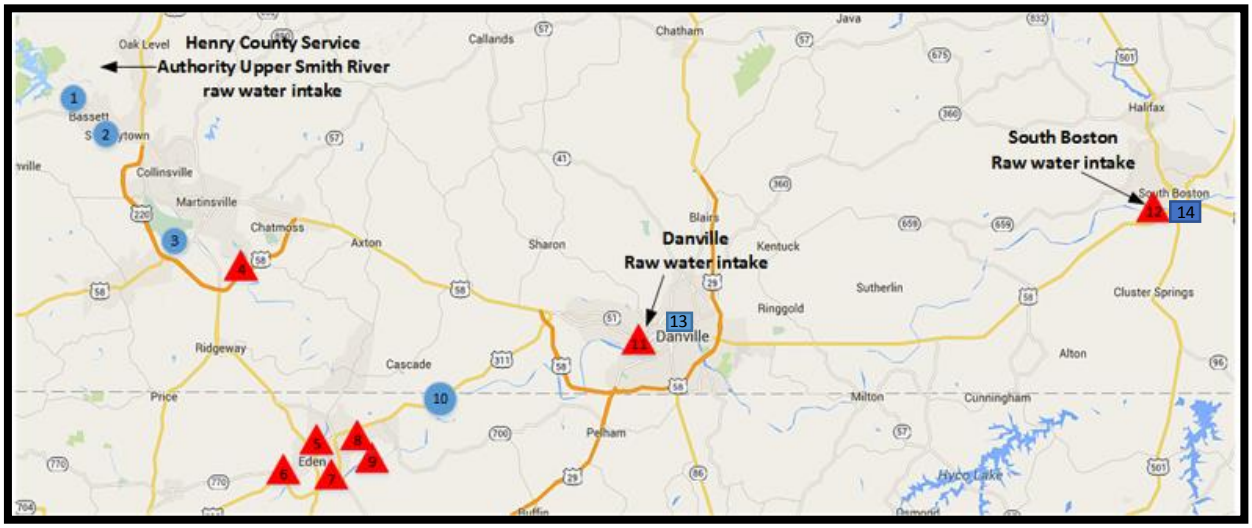

Figure 16. Map of research area and sampling locations

3.1.1.Bridge Sampling Sites

Water column samples were collected by lowering a bucket attached to a rope into the river at its center. Interstitial water and sediments were collected using an Ekman dredge at the center, left and right sides of the river to make a composite sample. Additionally, samples of raw influent water and finished water were collected at the Danville and South Boston drinking water treatment plants.

S-1 Bassett This site is furthest upstream site on the Smith River, located at the bridge crossing the river along Bullocks drive in Bassett, VA. Sediment and interstitial water samples were usually taken by hand from both sides of the riverbank because the rocky riverbed was not conducive for using the Ekman dredge.

S-2 Fieldale This site is located at the bridge crossing the Smith River along South Daniels Creek road in Fieldale, VA. Sediment and interstitial water samples were usually taken by Ekman dredge here.

S-3 Martinsville This site is located at the bridge crossing the Smith River along Rives road in Martinsville, VA. Sediment and interstitial water samples were usually taken by Ekman dredge here.

S-10 Stateline Bridge This site is located at the bridge on the Dan River border between Virginia and North Carolina. The bridge links the State Line Bridge Road in Virginia to the Berry Hill Bridge Road in North Carolina. Sediment and interstitial water samples were usually taken by hand due to the high elevation of the bridge, which made capturing water difficult through Ekman dredges.

3.1.2.Boat Sampling Sites

The boat sites along the Dan and Smith Rivers are sampled along a 1000-meter reach. Water samples and field measurements were collected center-stream at the upstream end of each reach. Starting at the upstream end of each reach, interstitial water and algae samples were collected near the shore at 100 meter intervals. The starting point along the left or right bank was determined at random by the toss of a coin. From there, the next two samples were collected along the same bank, then the opposite bank was sampled for two intervals, alternating banks after every second interval until the 1000-meter reach was completed. Interstitial water and algae were from each 100-meter segment composited along each 1000-meter reach.

Field sheets were kept on which are recorded the station number, date and time of collection, and a table indicating from which stream bank the sediment and algae samples were collected along each 100-meter segment.

S-4 Smith River Sports Complex The sample reach is between the Martinsville Sewage Treatment Plant and the Smith River Sports Complex. The start of the 1000-meter reach is approximately 1100 meters downstream of the launch point at the Martinsville Sewage Treatment Plant. The entire reach is downstream of the treatment plant outfall. This reach was sampled by canoe because there is no launch point for a larger boat within a reasonable distance to the reach. There were also two shallow rapids that would not be navigable in a larger boat.

S-5 Smith River This is the furthest downstream site on the Smith River and the reach sampling ended just before the confluence with the Dan River. The reach at begins just downstream of the Kings Highway bridge. Due to a set of rapids at approximately 300 meters below the bridge, and the configuration of the boat used to-date, sampling began 300 meters downstream of the bridge along this reach.

S-6 Dan River The furthest upstream site on the Dan River and upstream of the coal-ash spill. The sample reach began near an unknown building on the left bank at 835 N. Bridge St, Eden, NC (from Google Maps), approximately 300-meters downstream of the Old State Highway 87 bridge. The reach ended just upstream of the confluence with the Smith River.

S-7 Dan River This site is also upstream of the coal-ash spill but downstream of the confluence with the Smith river. The sample reach began approximately 340-meters downstream of the Leaksville Boat Ramp, and 940-meters downstream of the confluence of the Dan and Smith Rivers. The reach started opposite a small tributary that enters the river on the right bank. The Eden Sewage Treatment plant

outfall enters the river from the left bank at approximately the halfway point of the reach. The reach ended just upstream of the S. Van Buren Rd. Bridge.

S-8 Industrial Culvert This is not a river site but a low flowing stream which enters the Dan River approximately 1030-meters downstream of the Duke Energy boat ramp. Grab samples were collected at this site, rather than sampling along a 1000-meter reach. The water samples were collected from a pipe running through a stone dam that was constructed across the culvert. Interstitial water and algae samples were collected approximately 100-meters upstream of the dam, at a road crossing.

S-9 Dan River The sample reach begins approximately 1900-meters downstream of the Duke Energy boat ramp, where Town Creek enters the Dan River along the right bank. The reach ends approximately 1200-meters upstream of the S. Fieldcrest Road bridge.

S-11 Dan River upstream of the Schoolfield Dam, Danville, VA The sample reach begins 1000-meters upstream of the boat ramp at Abreu Grogan Park and ends at Schoolfield Dam in Danville.

S-12 Dan River at South Boston, VA The sample reach approximately 3300-meters upstream of the boat ramp at the South Boston Sewage Treatment Plant. The reach ends at the South Boston Water Treatment Plant intakes, on the left side of the river.

3.2. Flavor Profile Analysis (FPA)

3.2.1. Flavor Profile Analysis sessions

Flavor Profile Analysis was conducted according to Standard Method 2170 for characterizing odor descriptions and intensities of each collected sample. Column and interstitial water samples were collected head space free in 500mL glass bottles were stored at 6 °C for less than 48-hours before FPA analysis. 200 mL of sample water was poured into 500 mL Erlenmeyer flasks and heated in a water bath at 45° C for one hour before 4 to 6 trained panelists analyzed each sample.

Each sensory analysis session takes place in an odor free room with four to eight trained panelists who were advised not to wear perfume and not to eat or drink for at least an hour before each session. Panelists calibrate themselves for taste and odor before each session by tasting a sour reference standard (FPA = 4) and drinking taste free (Dasani) water throughout the FPA session.

Erlenmeyer flasks (500 mL) containing 200 mL of odor free water in a water bath at 45 °C were used by panelists as a calibration aid before they smell actual samples. Four to six river water samples were analyzed in each session. Panelists wait at least two minutes between smelling samples. Panelists were instructed to use a fragrance-free paper towel while holding the bottom of Erlenmeyer flasks, remove the ground glass stopper, and smell the odor free water or river water samples. River water samples were labeled with a numeric code so that panelists do not know the location or type of sample.

Panelists recorded odor descriptions and intensities for each sample on worksheets, an example is shown below in Table 10. Panelists often listed several descriptors and intensity ratings to describe each odor present in a sale. For example, in Table 10, Sample 1 had an “earthy” aroma that was moderate (6-8) intensity, as well as a strong (12) ”swampy” odor; however, sample 2 only had a weak (4) “musty” odor. Panelists were asked not to discuss their experiences with each sample during the analysis. After all panelists have characterized each sample, a discussion follows to reach a consensus of descriptors for each sample.

Table 10.Flavor Profile Analysis data worksheet example

Sample	Descriptor	Intensity
1	Musty	4
	Earthy	6
2	Swampy	8
	Earthy	4
3	Musty	4

3.2.2.Selection and Training of Sensory Analysts

Prior to the start of sampling, FPA panelists were trained to recognize using reference standards that represent tastes and odors that were found in drinking waters. Training was offered July 26, 2016 to thirteen panelists, six females and seven males, on Virginia Tech campus in Blacksburg, VA.

Training was conducted according to protocol from the handbook, *Flavor Profile Analysis: Screening and Training of Panelists* (AWWA Manual, 1993) In the 8-hour training session, panelists tasted different solutions of sucrose, sodium chloride, caffeine, and citric acid to identify the four basic tastes. Panelists smelled different reference standards to recognize different odors, including geosmin, 2-methylisoborneol, dimethyl trisulfide, and hexanal. The odor propertries of these reference standards are summarized below in Table 11. Threshold values (OTC) are generally defined as the lowest concentration, in water, that about 50% of a general population can detect the odor. Reported threshold values vary as individuals participating in sensory analysis studies have different olfactory capacities, thus some people may detect odors at higher or lower levels than others.

Table 11. Odor properties for reference standards used in FPA sensory training

Compound	Odor descriptors	Aqueous Odor Threshold ng/L	References
Geosmin	Earthy; musty	1-10	Rashash et al., 1997; Suffet, 1999;Piriou, 2009
2-methylisoborneol	Earthy; musty	1-15	Rashash et al., 1997; Suffet, 1999; Piriou, 2009
Dimethyl trisulfide	Garlic; swampy	10	Deng, 2010; Guo, 2015
Hexanal	Grassy	$5.0*10^4$	Rashash et al., 1997

3.2.3.Institutional Review Board

The research study protocol for Flavor Profile Analysis was approved by the Institutional Review Board for Research Involving Human Subjects at Virginia Tech (IRB 16-671). Informed consent was obtained in writing in accordance with the IRB protocol. Subjects reported being 18 years or older, healthy, not pregnant, and not consuming food or beverages or smoking one hour before the sensory evaluation/training. The training panel was composed of 13 members (6 females) of ages 21-65. During a FPA session, 4-8 of the trained panelists participated.

3.2.4.Chemicals and reagents

Several compounds were purchased and used as odor references for the training (Table 11). High purity n-hexanal (CAS 66-25-1), (+/-) geosmin (CAS 19700-21-1), 2-methylisobrneol (CAS 2371-42-8), and dimethyl trisulfide (CAS 624 – 92-0) were purchased from Sigma Aldrich (St. Louis, MO) and used as reference standards (Table 11).The taste training involves practice sessions using the four basic taste reference standards: sweet-sucrose, salty-sodium chloride, sour–citric acid, and bitter–quinine. Sucrose (CAS 57-50-1) and sodium chloride (CAS 7647-14-5) were food grade products purchased at a grocery store. Food grade citric acid (CAS 77-92-9) and quinine (CAS 130-65-0) were purchased from Sigma Aldrich. The taste free water was Dasani bottled water purchased at a grocery store. Panelists tasted different concentrations of reference standards to recognize the intensity scale associated with each reference. The objective of this training is to enable panelists to assess any taste and odor at various intensities. The strength of a taste or odor is scored on a six-point scale (Table 12).

Table 12. Flavor Profile Analysis Intensity Scale

Word Scale	Numerical Scale	Example
Taste/Odor Free	**0**	Pure water
Threshold	**T**	Detect a trace of taste or odor but cannot describe
Very Weak	**2**	Detect a very weak taste or odor and can provide description
Weak	**4**	Detect and describe one or more weak tastes or odors. "Weak" is analogous to sweetness of canned fruits.
Moderate	**8**	Detect and describe one or more moderate intensity tastes or odors. "Moderate" is analogous to the sweetness of soda pop.
Strong	**12**	Detect and describe one or more strong tastes or odors. "Strong" is analogous to sweetness of jelly.

3.3.Gas Chromatography/Mass Spectrometry (GC/MS) Coupled with Head-space Solid Phase Microextraction (SPME)

3.3.1.Chemicals

Concentrated analytical standards of geosmin, 2-methylisoborneol, 2-methoxy-3,5-dimethylpyrazine, and 2,4,6-trichloroanisole, were purchased as 100 µg/L analytical standards (Sigma Aldrich, St. Louis, MO) and diluted to 20 µg/L with methanol (HPLC grade, EMD Millipore Corporation, Darmstadt, Germany) for use as stock standards. The 2,4,6-tribromoanisole standard is prepared by dissolving 97% pure solid in methanol. Standards were prepared by injecting stock standards in deionized water ($\geq$16 MΩcm). The Carrier gas is ultra-high purity helium (AirGas, Radnor, PA), which passes through a three-stage filter before use.

3.3.2.Instrumentation

The gas chromatograph used is a Thermo Focus GC equipped with an Rxi-5Sil capillary column (30 m length, 0.25 mm internal diameter, 0.50 µm 5% polydimethylsiloxane film, Restek, Bellefonte, PA). Detection is provided by a Thermo DSQ II quadrupole mass spectrometer.

The manual injection solid phase microextraction fiber (Restek, Bellefonte, PA) is one cm in length with a 50/30 µm divinylbenzene/carboxen/polydimethylsiloxane coating.

3.3.3. Procedure

Samples were taken directly from the sampling sites in 60 mL glass volatile organic analysis vials with no head-space, and stored at ≤ 4 °C for less than two weeks before analysis. To prepare for analysis, 20 mL of sample is removed so that appropriate headspace is provided. Before an analytical batch, SPME fibers were conditioned in the GC inlet for 30 minutes at 270 °C with the GC oven set to 250 °C. Samples were then heated in a water bath to a temperature of 60 °C, and then the SPME fiber is inserted through the septum and allowed to equilibrate for 30 minutes.

The SPME fiber is then inserted into the GC injector which is heated to 240 °C and set to a splitless, 1.2 mL/min helium flow. The fiber remains in the injector for 5 minutes to desorb and clean. The column's initial temperature is set to 40 °C, which is held for 5 minutes. The temperature is then ramped at 8 °C per minute to 160 °C, and then at 20 °C per minute to 260 °C, which is held for one minute. The entire GC analysis takes 26 minutes.

The MS transfer line was maintained at a temperature of 250 °C. Mass spectrometry is performed with selective ion monitoring for the compounds listed in (Table 5).

Table 13. GC/MS Odorants and Quantification

Odorant	CAS #	Retention Time, min	Selective Ion Monitoring, m/z
Geosmin	19700-21-1	21.37	41,112,125
2-Methylisoborneol	2371-42-8	17.55	95,107,108
2-Methoxy-3,5-dimethylpyrazine	92508-08-2	14.47	54,109,138
2,4,6-Trichloroanisole	87-40-1	19.36	195,197,210
2,4,6-Tribromoanisole	607-99-8	23.16	329,344,346

3.3.4.Quantification

Chromatography and spectrometry were processed using Xcalibur software (Thermo Fisher, Waltham, MA). Calibration is achieved using external standards with concentrations ranging from 0.5 to 250 ng/L. Linearity is difficult to achieve over the entire range for all odorants, so calibration curves were generated for low and high range. The lower range extends from 0.5 ng/L to 10 ng/L, and the higher range covers 10 ng/L to 250 ng/L.

3.4. Field Measurements

3.4.1.DO, pH, temperature, and specific conductance**:**

YSI Model 556 multimeters were used to collect the dissolved oxygen, pH, temperature and specific conductance measurements. The meter is calibrated at the beginning of the work day, and some calibrations were checked again at each station. The probe is lowered to one foot below the water surface. The readings on the display were allowed to equilibrate, and the values were recorded on the field data sheet for the station.

3.4.2.Total Alkalinity**:**

Alkalinity is expressed as "P" (phenolphthalein) alkalinity or as "T" (total) alkalinity. Both types were determined by titration with sulfuric acid standard solution to an end point evidenced by the color change of an indicator solution. The total alkalinity includes all carbonate, bicarbonate, and hydroxide alkalinity. Although phenolphthalein testing was included in sampling procedures, no carbonate (e.g. "P"

alkalinity) was detected and is not included in the data results sections that follow; the total alkalinity was predominately bicarbonate.

3.5. Inductively-Coupled Plasma – Mass Spectrometry (ICP-MS)

Water collected from the sampling trips is filtered for ICP-MS by means of a syringe fitted with a 0.45 µm to obtain 10 mL samples. The ICP-MS analysis is performed with a Thermo Electron X-Series Instrument in accordance with Standard Method 3125-B. The instrument is used to quantify levels of various metals, and non-metals (Na, Mg, Al, Si, Ca, V, Cr, Fe, Mn, Co, Ni, Cu, Zn, As, Mo, Ag, Cd, Sn, Pb). Calibration standards were prepared in a matrix of 2% trace metal grade nictric acid by volume.

3.5.1.Reagents

The reagents used for this analysis were: 1. Argon and 2. Nitric acid, HNO_3 (Nitric Acid, CAS 7697-37-2, TraceMetal™ grade; Fisher Scientific; specific gravity 1.41); nitric acid, 1 + 1; nitric acid, 2%; nitric acid, 1%.

3.6. Ion Chromatography (IC)

Samples for anion analyses were filtered using 0.45 µm filters (Millipore™, nitrocellulose membrane) and then analyzed using the Dionex Ion Chromatography System (ICS) 1600 with the guard column, Dionex IonPac™ (AG9-HC, 4 x 50 mm), and analytical column, Dionex IonPac™ (AS9-HC, 4 x 250 mm). Samples were placed in 40 mL volatile organic analysis vials with polypropylene caps and low-bleed 0.125" PTFE/silicone septa (VOA, 0.135" septa for E-series vials).

3.6.1.Reagents

The standards used for this analysis were at concentrations of 0.1, 0.5, 1, 5, 10, 25 mg/L for each anion, which were prepared from 1000 µg/L Anion Standards (SPEX CertiPrep, H_2O[7732-18-5]). The eluent is 9 mM sodium carbonate prepared using Fisher S263-3kg from a 0.5M stock solution. Bromide, chloride, fluoride, nitrate, nitrite, phosphate and sulfate were measured by EPA Method 300.

3.7. Standard Methods for Known Odor-Associated Microorganisms

Relative changes in numbers of actinomycetes and fungi were quantified using plate counting techniques. Variations of algal growth were analyzed by monitoring chlorophyll-a concentrations in both interstitial and column water samples according to standard methods (APHA, 2012). Actinomycetes were enumerated on Actinomycete Isolation agar (AIA) (Method 9520; HiMedia, Mumbai, India.) and fungi were enumerated on rose-bengal agar (RBA) (Method 9610; HiMedia, Mumbai, India.). Briefly, samples were transported to VT laboratory on ice and aliquots of each sample type were serially diluted in deionized water. 100 µL of the appropriate dilution step were spread on the appropriate agar and plates were incubated at 28 °C for AIA and 26 °C for RBA for 5 days. Colony forming units (CFUs) were counted immediately following incubation and data was presented as CFU/100mL. Algal biomass was quantified by extracting chlorophyll-a from water, filtered using vortexing and freezing in acetone (Method 10200). Extracted chlorophyll-a was quantified on a fluorometer - spectrophotometer (Thermo Fisher Inc., Waltham, MA, USA). Data was reported as mg Chl/mL in water.

Table 14. Summary of analytical methods and analytes measured

Method	Analyte	Standard Method
Flavor Profile Analysis (FPA)	Sensory characterization of odors	2170
Headspace solid phase Microextraction gas chromatography mass spectrometry (HS SPME GC/MS)	Geosmin, 2-MIB, haloanisoles, methoxypyrazines	

Agar plate counts	Actinomycetes, fungi, chlorophyll-a	9520,9610, 2012
Ion chromatography (IC)	NO_2^-, NO_3^-, PO_4^{3-}, SO_4^{2-}, Br^-, F^-, Cl^-	EPA 300
Inductively-Coupled Plasma – Mass Spectrometry (ICP-MS)	27Al, 75As, 137Ba, 111Cd, 43Ca, 35Cl, 52Cr, 59Co, 65Cu, 54Fe, 208Pb, 25Mg, 55Mn,95Mo, 60Ni, 31P, 39K, 78Se, 288Si, 107Ag, 23Na, 88Sr, 34S, 238U, 51V, 66Zn	3125-B
Alkalinity by titration	Total alkalinity	2320-B
YSI Model 556 multimeters	DO, pH, temperature (˚C), specific conductance	

3.8. Multivariate data analysis and statistical methods

Multivariate analyses and statistical tests were performed using R version 3.4.2 (09/28/2017). Informal hypotheses tests were performed on the collected date with the aim of graphing to look for differences between the Dan and Smith Rivers, column and interstitial water, spatial and temporal trends.

Hierarchical cluster analysis was performed to evaluate the similarities between each river site based on the sensory, chemical, and biological data from the collected samples. The raw data are first scaled so all variables are standardized with a mean equal to zero, and a standard deviation equal to 1. The scaled data are then used to calculate the Euclidean distance between each pair of individual and arranged into a matrix which serves as the starting point for the clustering analysis. The hclust function in the stats package performed the hierarchical clustering based on a dissimilarity object, the Euclidean distance matrix, and a method for analyzing the object, Ward's method of minimum covariance.

Hierarchical clustering analysis first assigns each object to its own cluster, then proceeds to join the most similar objects in their own cluster; objects are confined to just one cluster, so once an object is assigned to a cluster it cannot be repeated in another group (Everitt, 2011).

For example, Figure 17 contains all chemical, sensory, and biological data results collected from all river sites from August, 2016 to September, 2017. The data were subset for separate analysis of column and interstitial water samples. Figure 17. Cluster dendogram of column water samples shows that the industrial culvert (8), is the most dissimilar site relative to others; and the second break separates the Dan (6-12) and Smith River (1-5) sites. The Dan and Smith River sites cluster together showing their similarity based on the collected data.

Principle component analysis (PCA) is another technique performed on the combined data set as a tool for exploratory analysis. The large amount of variables in the collected data limit the usefulness of other graphical techniques to provide useful information. For example, it would be difficult to elucidate any meaningful relationship that may exist between odorant and actinomycetes, or fungi, and chorlophyll, without graphing each combination for each site and each month of samples. PCA is a technique used often for large datasets to reduce dimensionality by identifying a few variables which account for causing the most variation in the data.

PCA was applied using two methods: Probabilistic PCA from the pcaMethods package, graphics were created using the ggplot2 and ggfortify packages. and the conventional PCA from the stats package. Conventional PCA requires complete cases so missing data were removed before analysis. One problem encountered using conventional PCA is that by omitting incomplete cases there is a possibility of losing a substantial amount of information, which lowers the effective samples size, makes analyses less effective, or possibly biased results (Everitt, 2011).

Probabilistic PCA was chosen to handle missing values as an efficient, and unbiased method to impute data points via interpolation(Grabich, 2014). The method is a dimensionality reduction technique that analyzes the data via a lower dimensional latent space (Tipping, 1999 Missing data occurred when samples were either not tested for a specific constituent or over the detection limit range. For example, the analytical method for biological data, which requires counting units formed after inoculation and incubation, has an upper limit of 300 units; if the number of units formed are more than 300, the sample is too numerous to count and the data point is not available for analysis. Another cause of missing data points was that interstitial waters were not analyzed by the field measurements summarized in section 3.4.1.

The purpose of using a second method for PCA, where missing values were omitted from the data set, was to informally validate the results of the PPCA. The method to perform this analysis was the singular value decomposition which examines the correlations and covariance between individuals. The variables were scaled for unit variance and centered so that each variables mean was subtracted before analysis.

4. Multivariate Results

4.1. Hierarchical Cluster Dendograms

The purpose of the cluster analysis was to evaluate the extent of similarity of the data from each site. The dendogram shown in Figure 17 contains all the data results from sensory, chemical, and biological analyses for the column water samples. The tree-like graphics lists the river sites that are clustered together along the horizontal axis, and the distance at which the cluster was formed along the vertical axis. The dendogram can be interpreted either from the top down, or bottom up. From the top of the graph, the first cluster that formed was site 8, the industrial culvert, separating from the other sites. The next cluster was formed separating the Dan and Smith river sites on the right and left side of the graph. On the right side, Dan River sites, the next most dissimilar object is site 12, the further downstream site. The next clusters that are formed contain the two lower Dan River sites, 10 and 11; then site 6 separates and sites 7 and 9 form a cluster. Similar trends are followed on the left side, Smith River sites, where the lower Smith Rivers sites, 4 and 5 cluster together, and the upper Smith River sites 1,2, and 3 clustered.

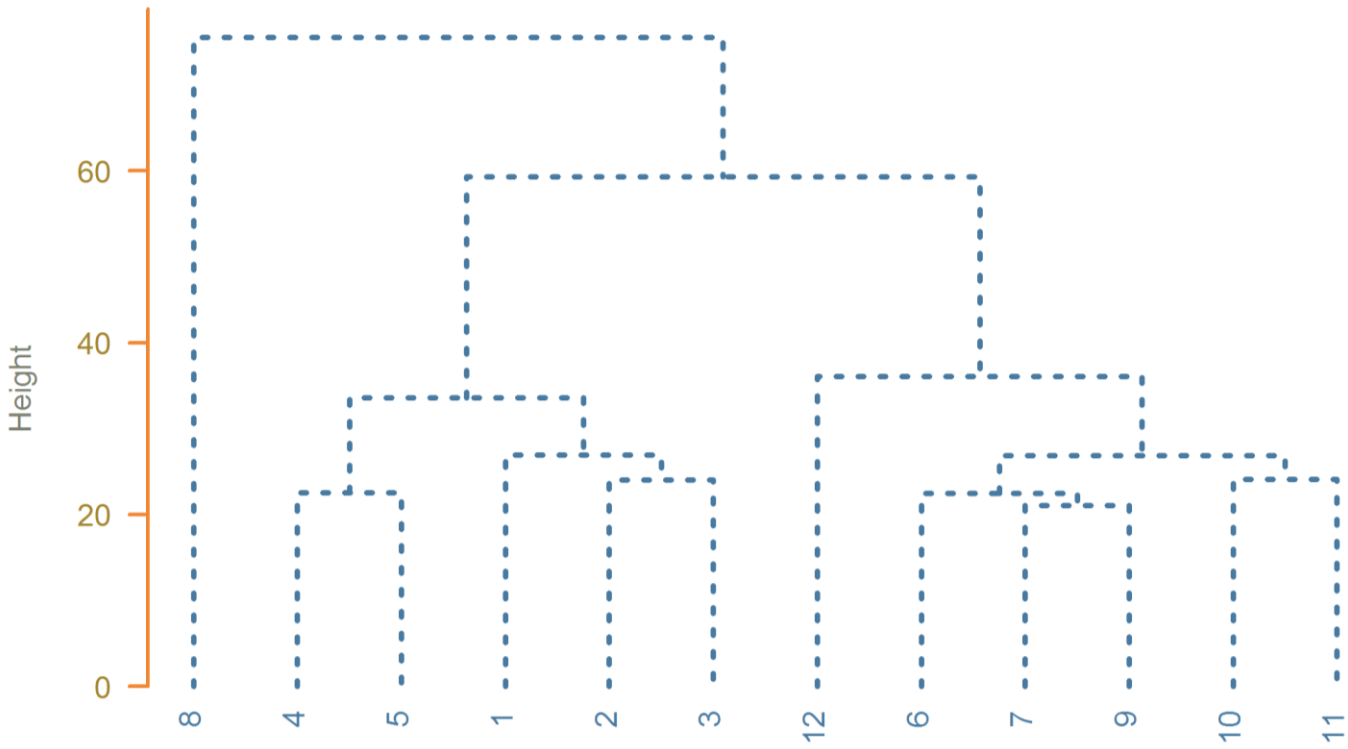

Figure 17. Cluster dendogram of column water samples containing all data results from sensory, chemical, and biological analyses

The cluster dendogram below in Figure 18 is built using the data results from interstitial water samples. Field measurements including dissolved oxygen, pH, temperature, conductivity, and total alkalinity were not measured for interstitial water. Consistent with the column water, the upper Smith River sites 1,2, and 3 form a cluster, and the industrial culvert, site 8, was the first to cluster alone. However, the Smith and Dan Rivers do not separate in terms of the interstitial waters. Instead of forming two clusters, each containing the sites for the respective river, sites 4 and 5 on the lower Smith are more similar to the Dan River sites, 6,7 and 9. The difference in patterns between the two dendograms may be that the additional measurements for column water samples (pH, temperature, etc.) which are not measured for interstitial waters, provide more variables to relate the sites in close proximity and between rivers.

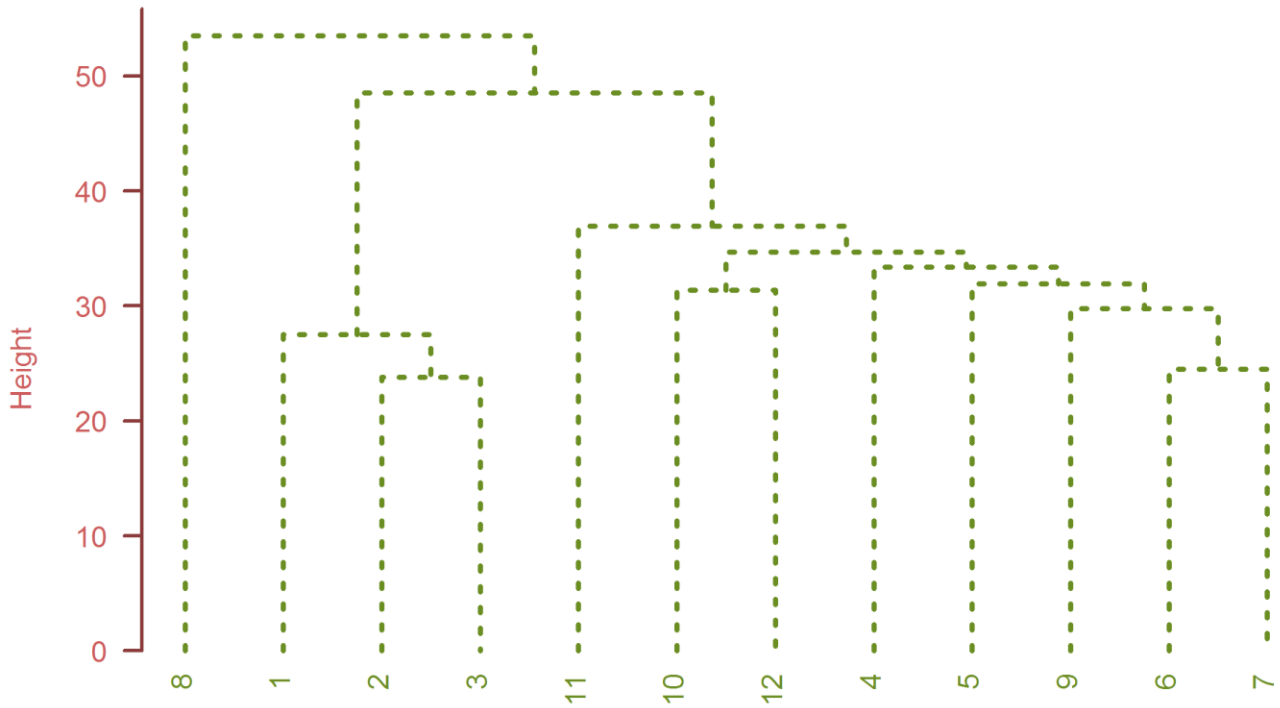

Figure 18. Cluster dendograms of interstitial water samples containing all data results from sensory, chemical, and biological analyses

4.2. Probabilistic principal component analysis

The probabilistic principle component analysis shown below in Figure 19 contains all measured sensory, chemical, and biological data results for samples collected (Table 14). The data were scaled for unit variance and means centered at zero before the analysis. Of the 312 samples with 50 variables included in the analysis, 7.4% of the observations were missing values which were imputed by interpolation and included in the PCA.

Ideally, the first few principle components would explain more than 80% of the variation, in which case the data set could be simplified by focusing on the factors that contribute to the principle components without compromising too much accuracy. For example, if the plot showed one or two clusters of points, which represent each sample collected, this would indicate there are certain factors that make these samples similar to each other, and distinct from others. Identifying which factors are driving the data to different groups, would help focus further analysis on fewer factors that are the most important, without overlooking or missing valuable information in the total data.

However in Figure 19 the only obvious cluster is formed near the origin, with scores close to zero the conclusion is that the majority of the data cannot be explained by either of the principle components. There are a few singular data points that are driving the principle components, one point on the far left is contributing to nearly all of the 19.6 % of the variation for the first principle component. The two principle components calculated cumulatively explained 28.5% of the total variation in the data.

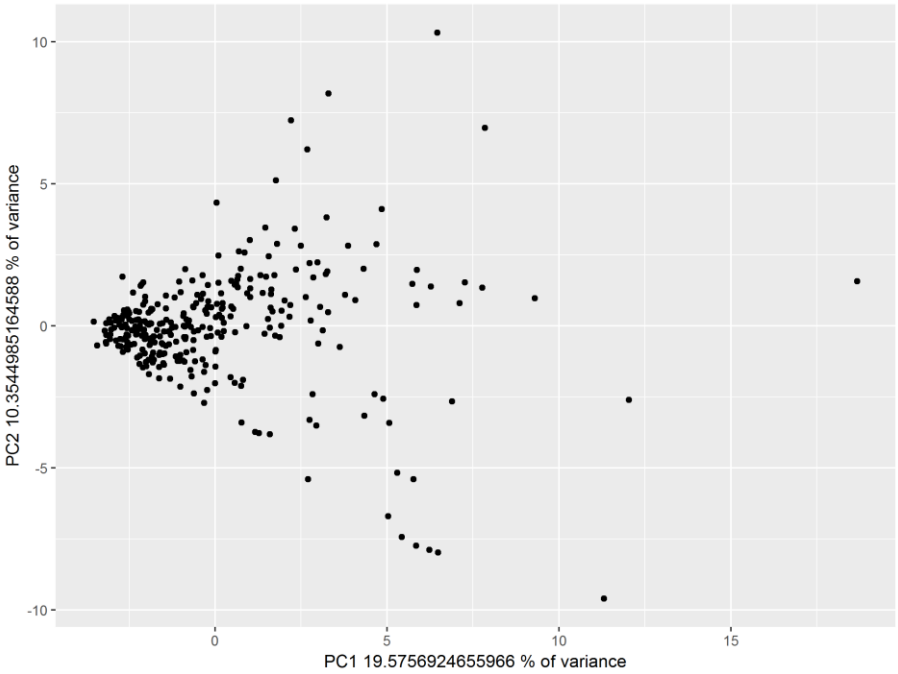

Figure 19. Probabilistic principal component analysis of all data samples

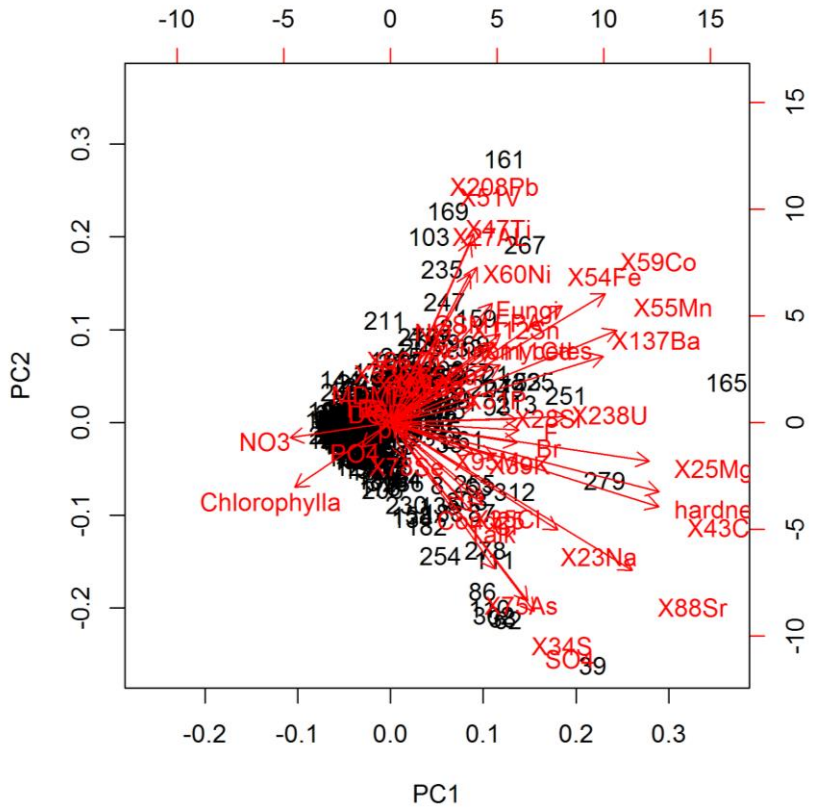

Figure 20. Probabilistic PCA biplot for all water samples

The biplot shown in Figure 20 is a graphic representation of the contributions of each of the variables to the first two principle components and the individual observations on the same scale. In an ideal case, this would represent which variables are causing the data to form distinguished clusters which are similar to each other in terms of certain variable. The variable names are shown on the plot, and their position on the graph indicates their contribution to pc 1 and 2. The variables, which are dispersed horizontally contribute to pc1, these are largely water quality parameters such as nitrate, magnesium, chloride, and from the biological data chlorophyll-a. Variables spread vertically contribute to pc 2 are mostly sulfate and several metals, including lead, vanadium, arsenic. Because pc 1 and 2 only explain a small portion of the overall data, it may be that these variables are significant in just a few samples.

The biplot shows that points, 161,165,169, and 39 are the primary contributors to the first two principle components; these are seen on the PPCA plot (Figure 19) as points that are the most separate from the large cluster near the origin. Again, this indicates that there are a just a few samples which were very unique from the others; but overall there are no large differences between the samples due to certain factors. Table 15 summarizes the four samples that are distinct in the PPCA. Each are interstitial water samples, and two from the Dan River were in February, 2017. Each are interstitial water samples, and two are from the Dan River in February, 2017. Site 8 is the industrial culvert which flows into the Dan River near to site 9, but has not shown to significantly influence the river water. Site 1 is the furthest upstream site on the Smith River.

Table 15. Primary contributing samples for PPCA

Individual	Scores		Water type	River site	Sample date
	PC1	PC2			
39	11.22	-9.58	Interstitial	8 – Culvert	9/2016
161	6.35	10.41	Interstitial	9 – Dan	2/2017
165	18.68	1.58	Interstitial	11 – Dan	2/2017
169	3.19	8.34	Interstitial	1 – Smith	3/2017

The PPCA plot shown in Figure 21 contains 156 measured data points from the column water samples, and 18 values (2.2%) were imputed via interpolation for variables with missing data points. The data were scaled for unit variance and centered before analysis. The first principle component accounts for 25% of the variance of the original data, and cumulatively the first two components account for about 35% of the variance in the original data. The plot of scores show that many data points are clustered near the origin and values close to zero, however there are several points on the far left along the horizontal axis, and near the bottom right, showing these observations are contributing to each principle component. The biplot shown in Figure 22 provides information on which samples are important in terms of pc1, (152, 20, 32) and pc2 (72,73). The biplot (Figure 22) also shows that several metals, and conductivity, are the variables contributing the most to each component. Table 16 provides a summary of the important contributing samples. Samples from the industrial culvert are largely contributing to the principle components, this may be due to the distinct nature of the site, which is low flowing and receives industrial drainage. Additionally, the points on the PPCA plot not clustered around the origin, are mostly disperse, indicating the samples are not closely related by some trends in the variables. If the observations were related by some common variable, for example if all had much higher geosmin concentrations relative to other samples, they would cluster together.

Table 16. Primary column water samples contributing to (PPCA) components

Individual	Scores		River site	Sample date
	PC1	PC2		
20	-15.99	2.87	8 – Culvert	9/2016
32	-16.39	2.11	8 – Culvert	10/2016
72	-1.17	-8.02	12 – Dan	1/2017
73	2.20	4.25	1 – Smith	2/2017
152	-14.93	4.11	8 – Culvert	9/2017

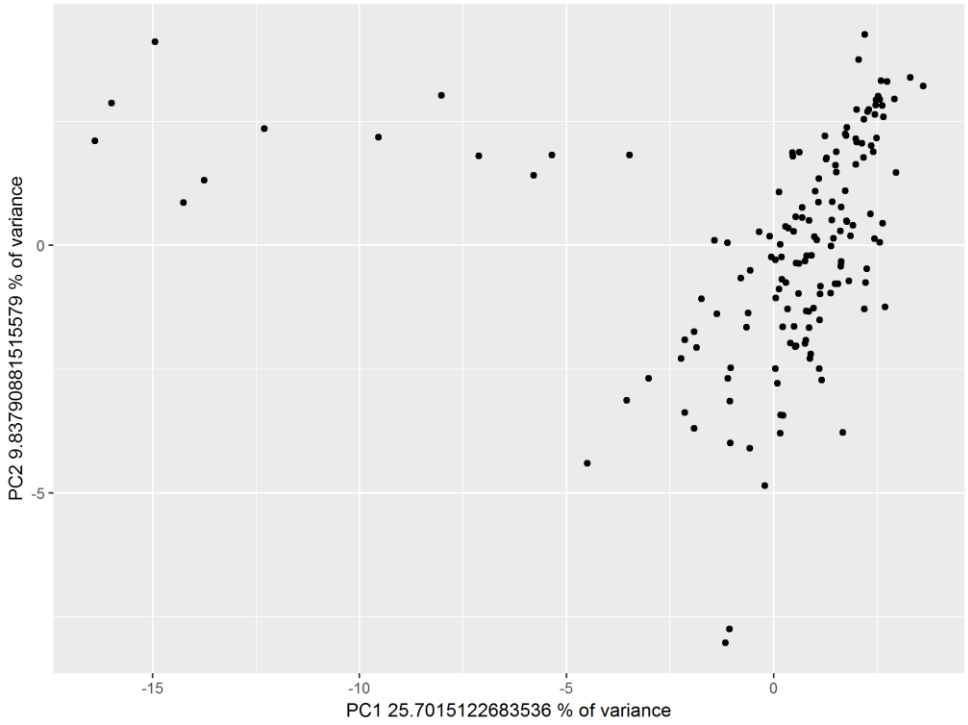

Figure 21. Probabilistic principle component analysis of column water samples

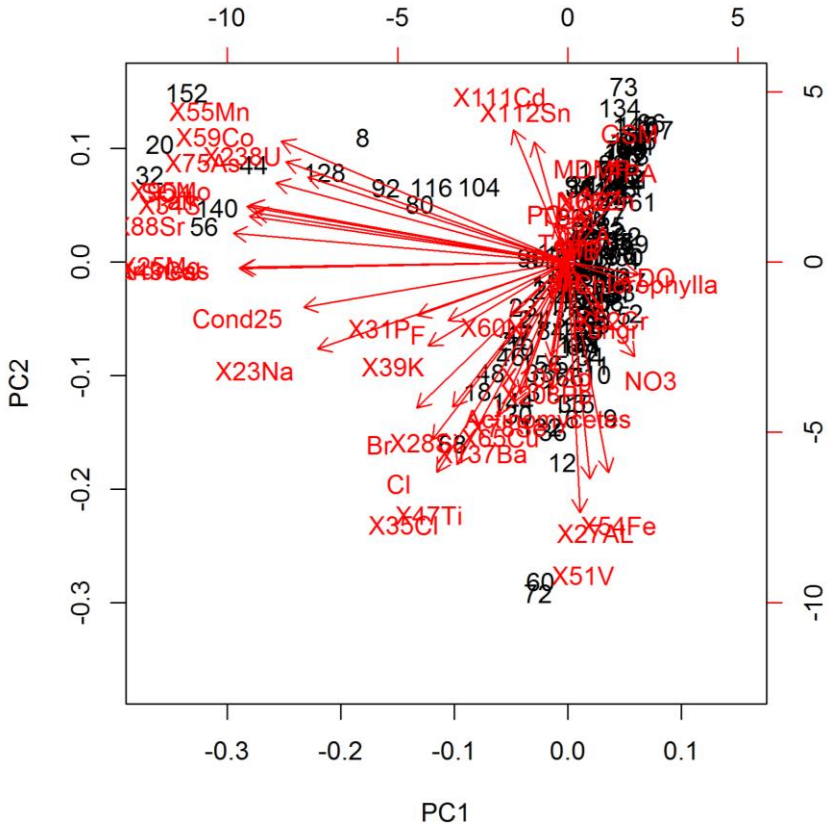

Figure 22. Biplot of PPCA for column water samples

The PPCA scores plot for interstitial water samples are shown below in Figure 23. There were 44 missing values, 2.8% of the data, which were imputed by the analysis. Similar to the column water samples, the first two principle components do not explain a majority of the variation present in the original data. Cumulatively, pc 1 and pc2 account for only 27.4% of the total variance. There are two points on the PPCA which stand out as primary contributors: on the top right, showing a strong contribution to pc1 and the other at the bottom, primarily contributing to pc2. The biplot below (Figure 24) shows these point are 80 and 82 for pc1 and 2 respectively.

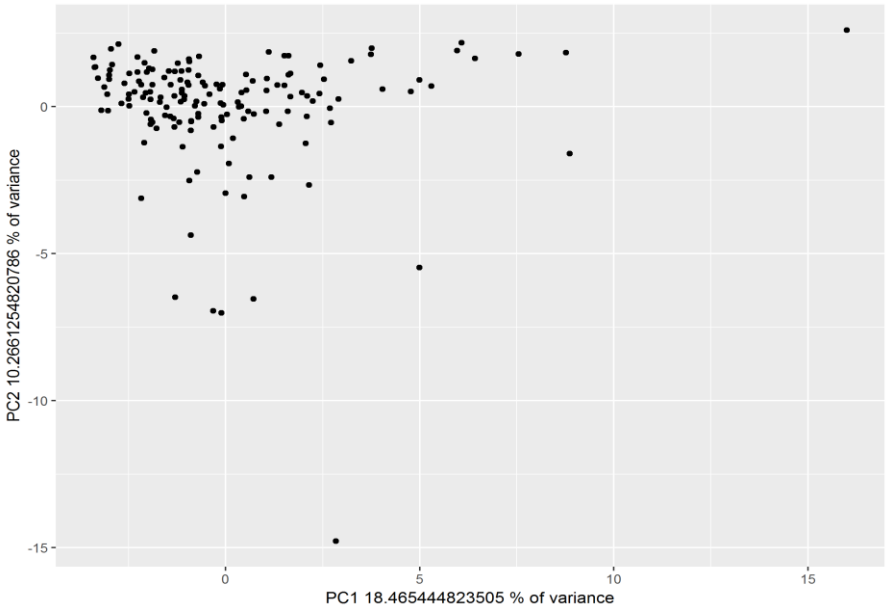

Figure 23. Probabilistic principle component analysis of interstitial water samples

Table 17 provides a summary of the samples which are primary contributors to the principle components show in Figure 23. As seen in the combined probabilistic PCA (Figure 23), two Dan River sites, 9 and 11, in February, 2017, where the primary contributors to the principle components. These samples were the most distinct in terms of their chemical, biological, and sensory data results.

Table 17. Primary contributing interstitial water samples (PPCA)

Individual	**Scores**		**River site**	**Sample date**
	PC1	**PC2**		
80	2.85	-13.94	9 – Dan	2/2017
82	16.01	2.88	11 – Dan	2/2017

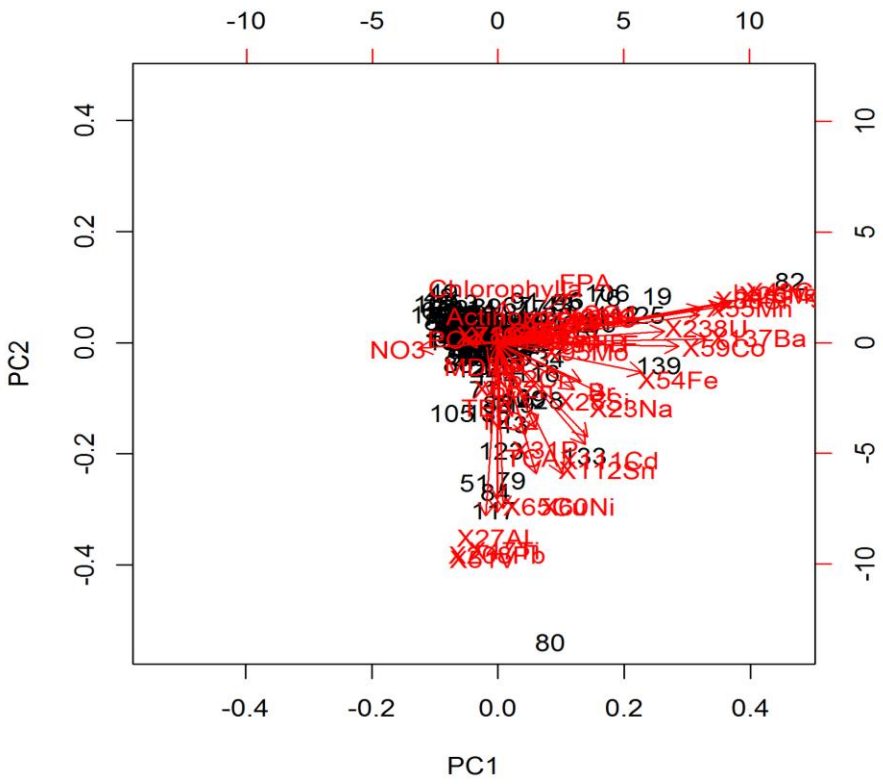

Figure 24. Biplot of PPCA for interstitial water samples

4.3. Principal component analysis

The method used for the principle component analysis of the column water samples, the PCA plot is shown below in Figure 25, does not use interpolation to handle missing values; instead missing values were excluded before the analysis. The data are structured so that each row contains the measured variable from each month of sampling at a particular site. For example, one row would contain all biological, chemical, and sensory data results for May, 2017. Regardless if only one or several values are missing from the results, that site will omitted before the analysis. Although removing the row with missing values reduces the number of observations from which the principle components are determined, one benefit is a more conservative method as estimation is not required.

The PCA plot of column water samples (Figure 25) using the more conservative PCA method shows similar results to the probabilistic PCA method used. Similar to the previous analyses, there are not distinct clusters show in the PCA, rather a few samples that are distinct from others based on their measure parameters. Table 18 summarizes these points labeled 32, 44, 56, 60, and 73 are extreme points on the plot and primary contributors to the components. The samples from the culvert at site 8 are distinct from other river sites as shown in the cluster analysis previously (Figure 17 and Figure 18).

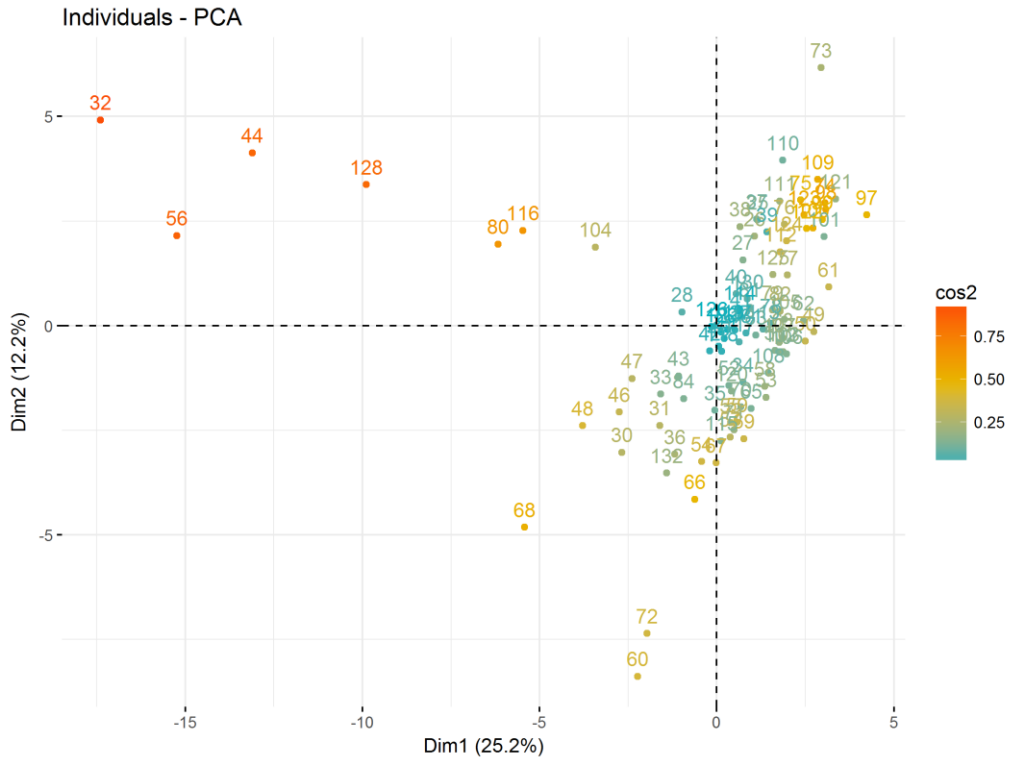

Figure 25. Principle component analysis of column water with missing values removed

The biplot displayed in Figure 26 shows which variables are primarily contributing to each component. These samples are consistent with results from the probabilistic PCA for column water samples (Figure 21 and Table 16). Also, metals are again a primary contribution to the principle components. The consistent dominance of metals could be due to the large number that are measured compared with others measured parameters in the data; metals accounts for 37 of the 50 measured variables in the total data set.

Table 18. Primary column water samples contributing to principle components

Individual	River site	Sample date
32	8 – Culvert	10/2016
44	8 – Culvert	11/2016
56	8 – Culvert	12/2016
60	12 – Dan	12/2016
73	1 – Smith	2/2017

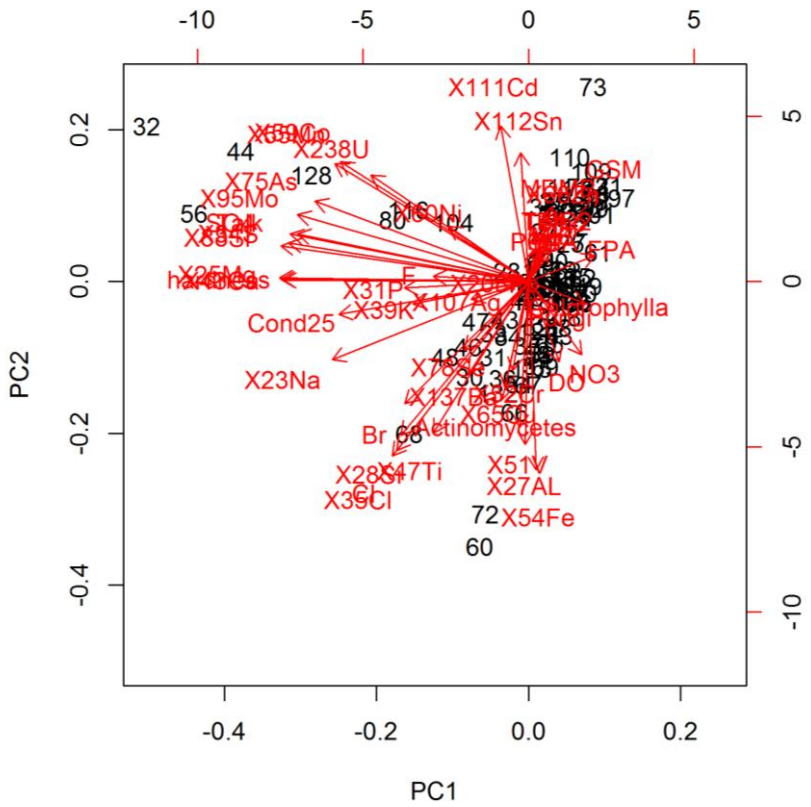

Figure 26. PCA biplot of water column samples

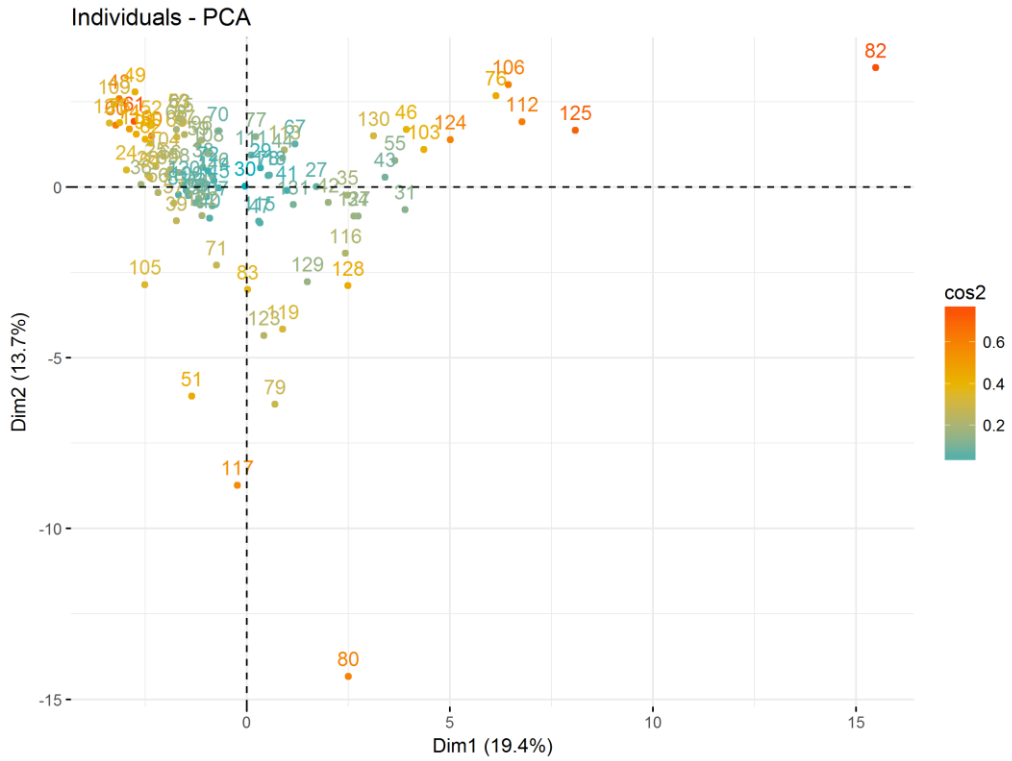

Figure 27. Principle component analysis of interstitial water samples with missing values removed

The PCA plot shown in Figure 27 contains the interstitial water samples for which the cumulative variance explained by the first two principle component is 33% of the total variance. As seen in the previous PCA plots, most of the data points are clustered close to the origin, meaning that the first principle components explain very little to none of the variability of these data. However, consistent with previous analyses, several data points, 82,106, 112, and 125, are located far along the axes. These samples are distinct from other samples, and dominant in terms of the variables contributing to the principle components. The PCA is not able to find variables which explain the majority of the data, but there are a few variables that distinguish several samples which are summarized below in Table 19.

Table 19. Primary interstitial water samples contributing to principle components

Individual	River site	Sample date
80	9 - Dan	2/2017
82	11 – Dan	2/2017
106	11 – Dan	5/2017
112	5-Smith	6/2017
125	6 – Dan	7/2017

The first two samples, sites 9 and 11 from February,2017, were also identified in the probabilistic PCA as primary contributors (Figure 23 and Table 17). The PCA identified three sites which were not observed to be major contributors in the previous analyses including interstitial waters: site 11 from May, 2017; site 5 in June, 2017; and site 6 in July,2017. This could be an example of where the

complete case analysis that is required for PCA has over projected based on the truncated dataset. These sites may not be as extreme nor influential if the excluded observations were also considered.

The biplot in Figure 28 shows the variables that are controlling the first principle component are nitrate and several metals. The second principle component is mainly driven by chlorophyll-a and several metals. The metals may be more dominant in this analysis as compared to the column water samples because the interstitial waters are not measured for temperature, total alkalinity, dissolved oxygen, conductivity, thus there are fewer variables to balance the number of metals measured (37 metal elements of 45 variables).

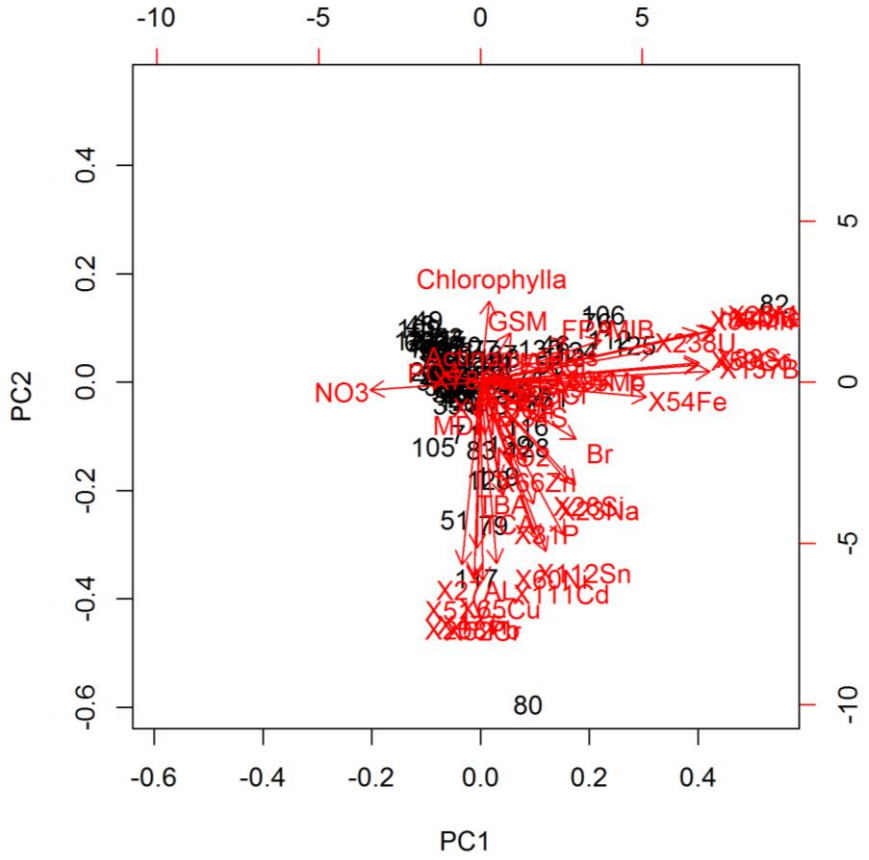

Figure 28. PCA biplot for interstitial water samples

5. Discussion

The goal of this research was to monitor water quality and ecological communities in the Dan and Smith Rivers to evaluate the impacts of industrial and anthropogenic activities that may have negatively impacted this drinking water source. The data and results presented in the previous section are from thirteen months of sampling at twelve different sites along the Dan and Smith Rivers. The sampling time period and numerous sampling locations were needed to better understand the chemical, sensory, and biological characteristics of the river system, but drawing conclusions from the large amount of data are not obvious. Multivariate statistical techniques were used to evaluate temporal and spatial trends, and interpret which measured parameters were influencing changes in the rivers.

Hierarchical agglomerative cluster analyses were used as a starting point to identify groupings of sampling sites, providing an overall comparison of the sites' similarities and differences based on the measured parameters (odor intensity (FPA), odorants (SPME GC/MS), major anions (IC) and cations (ICP-MS), biology, field measurements). The cluster dendogram of column water samples for all sites demonstrated several important trends. First, site 8, the industrial culvert, was the most different from all other sites, which is not a surprising because it is a small stream formed from industrial discharge and runoff. The distance between site 8 on the dendogram from site 9, located just downstream on the Dan River, suggests that measured parameters (Table 14) between the two sites are very different; furthermore, the discharge from site 8 is largely not influencing the downstream Dan River sites. The Dan River sites 6, 7 and 9 are also closely grouped on the dendogram; sites 7 and 9 are shown to be the most similar Dan River sites. The importance of this observation is that there is very little difference in the column water samples measured upstream and downstream (sites 7 and 9) of the coal-ash spill.

The cluster dendogram for interstitial water samples follows some patterns observed in the column water analysis. The industrial culvert, site 8, is the most distinct sampling site. The separation, thus difference based on their measure parameters (Table 14) between site 8 and 9, suggests there is little to no influence from this industrial culvert on the Dan River. Additionally, sites 6, 7 and 9 are grouped together; however, sites 6 and 7 are slightly more distinct from site 9 in terms of interstitial waters. This suggests some dynamic variations in the interstitial waters from the coal-ash spill. Beyond this similar trend, the groupings of river sites based on the interstitial water samples seem to be driven by the site location. The next break in the dendogram separates the upper Smith River sites (1-3) from the other river sites. The lower Smith River sites, 4 and 5, are more similar to the Dan River sites, 6, 7, and 9, and group together. The downstream Dan River sites, 10 and 12, cluster together, while site 11, nearest to the Danville water treatment intake, is the most unique based the measured parameters for the interstitial waters. The cluster dendograms provide some interpretation for how each site and rivers are related, but further analyses are required to elucidate what factors are influencing these relationships.

Two methods of principal component analyses were employed with the purpose of reducing dimensionality and identifying variables that are driving change in the data set. The probabilistic PCA was useful as this method can extend the scope of conventional PVA by imputing missing values into the dataset, allowing all observations to be evaluated, instead of omitting all where data are not available (Bishop and Tipping, 1999). The PPCA of all samples shown in Figure 19, including all results from chemical, biological, and sensory data (Table 14) for both interstitial and column waters, does not find one or a few factors which are key to driving changes in the total data set. Instead of separate clusters of data points, most data points are loosely grouped close to the origin on the graph. There are a few points that stand out from others, suggesting these samples were the most different from other collected samples. These samples were identified (Table 15) as two site on the Dan river collected in February, 2017 and the industrial culvert sample from September, 2016. These results were consistent

with the conventional PCA method used to evaluate the dataset (Figure 25 and Figure 27). Separate PCA were conducted for interstitial and column waters because field data (pH, dissolved oxygen, temperature, specific conductance) are not measured for the interstitial waters; thus, these missing factors were omitted for the PCA. The conventional PCA method shows a similar trend to the PPCA, the majority of samples cannot be distinguished by a few factors; however, the samples also identified by the PPCA were the most different (Table 19). Water quality parameters and measured chlorophyll-a were largely the distinguishing factors from other samples; however, these same factors are not largely driving change for other collected samples. There were no few identified factors that were effectively distinguishing among a majority of the dataset.

One reason for the lack of distinction amongst the data may be that there were no outstanding events, such as a T&O event, during the project timeline. The collected samples therefore reflect the Dan and Smith Rivers under normal conditions. The PCA methods successfully identified the industrial culvert samples as the most dissimilar from others, which was not surprising as this is not a river site. Furthermore, if the samples from the culvert site were indistinguishable from river samples, it would have suggested improper sample collection, measurements, and data analyses. Although the two rivers are distinct in terms of their water quality and biological communities, there were no extraordinary shifts in these parameters across river sites or over thirteen months of sampling, that may have helped explain what released the strong earthy/musty odors in the Dan and Smith Rivers.

6. Conclusions

- Cluster dendograms show the Dan and Smith Rivers' column water are distinct in terms of measured biological, chemical, and sensory factors.
- The industrial culvert is dissimilar from all other sample sites, and its discharges are not largely influencing the Dan River downstream.
- Water quality and biological parameters, such as nitrate, magnesium, chloride and chlorophyll-a, were factors that distinguished a few samples. However, these factors were not highly variable for the majority of the collected samples.
- The principal component analyses could not reduce the dimensionality of the large dataset by identifying a smaller set of factors that were separated certain samples from others, such as very different levels of metals detected in certain sites, or an increase in levels of odorants detected in warmer months.
- The PCA performance was likely limited by the fact that there were no T&O or aberrant water quality events during the project timeline.

7. References

Abbaszadegan, M., Yi, M., & Alum, A. (2015). Stimulation of 2-methylisoborneol (MIB) production by actinomycetes after cyclic chlorination in drinking water distribution systems. *Journal of Environmental Science and Health Part a-Toxic/Hazardous Substances & Environmental Engineering, 50*(4), 365-371. doi:10.1080/10934529.2015.987526

Association, A. A. P. H. (2011). Standard Methods for the Examination of Water and Wastewater. In *2150B Odor Threshold Test* Washington, DC: American Public Health Association, American Water Works Association, Water Environment Federation.

Association, A. W. W. (1993). *Flavor Profile Analysis: Screening and Training of Panelists*. Denver, CO.

AWWA, T. O. C. (2002). COMMITTEE REPORT: Options for a taste and odor standard. *Journal (American Water Works Association), 94*(6), 80-87.

Boleda, M. R., Díaz, A., Martí, I., Martín-Alonso, J., Matia, L., Romero, J., & Ventura, F. (2007). A review of taste and odour events in Barcelona's drinking water area (1990–2004). *Water Science & Technology, 55*(5). doi:10.2166/wst.2007.182

Cancho, B., Fabrellas, C., Diaz, A., Ventura, F., & Galceran, M. T. (2001). Determination of the odor threshold concentrations of iodinated trihalomethanes in drinking water. *Journal of Agricultural and Food Chemistry, 49*(4), 1881-1884. doi:10.1021/jf001252m

Crawley, M. J. (2007a). Multivariate Statistics. In *The R Book* (pp. 731-747): John Wiley & Sons, Ltd.

Crawley, M. J. (2007b). Tree Models. In *The R Book* (pp. 685-700): John Wiley & Sons, Ltd.

Deng, X., Liang, G., Chen, J., Qi, M., & Xie, P. (2011). Simultaneous determination of eight common odors in natural water body using automatic purge and trap coupled to gas chromatography with mass spectrometry. *J Chromatogr A, 1218*(24), 3791-3798. doi:10.1016/j.chroma.2011.04.041

Diaz, A., Ventura, F., & Galceran, T. (2005). Determination of odorous mixed chloro-bromoanisoles in water by solid-phase micro-extraction and gas chromatography-mass detection. *Journal of Chromatography A, 1064*(1), 97-106. doi:10.1016/j.chroma.2004.12.027

Dietrich, A. M., & Burlingame, G. A. (2015). Critical Review and Rethinking of USEPA Secondary Standards for Maintaining Organoleptic Quality of Drinking Water. *Environmental Science & Technology, 49*(2), 708-720. doi:10.1021/es504403t

EPA, U. (Producer). (2017, 12/16/2017). Secondary Drinking Water Standards: Guidance for Nuisance Chemicals. *Drinking Water Contaminants - Standards and Regulations*. Retrieved from https://www.epa.gov/dwstandardsregulations/secondary-drinking-water-standards-guidance-nuisance-chemicals#main-content

Everitt, B. H., Torsten. (2011). An Introduction to Applied Multivariate Analysis with R. In (2011 ed.). New York: Springer-Verlan.

Grabich, S., C. Gray, L. Messer, K. Rappazzo, J. Jagai, AND D. Lobdell. (2014). Principle Component Analysis with Incomplete Data: A simulation of R pcaMethods package in Constructing an Environmental Quality Index with Missing Data.

Gerber, N.N. (1983). Volatile Substances from Actinomycetes - Their Roles in the Odor Pollution of Water. *Water Science and Technology, 15 (6-7),115-125.*

Heim, T. H., & Dietrich, A. M. (2007). Sensory aspects and water quality impacts of chlorinated and chloraminated drinking water in contact with HDPE and cPVC pipe. *Water Res, 41*(4), 757-764. doi:10.1016/j.watres.2006.11.028

Hill, J. L., Hocking, A. D., & Whitfield, F. B. (1995). The role of fungi in the production of chloroanisoles in general purpose freight containers. *Food Chemistry, 54*(2), 161-166. doi:https://doi.org/10.1016/0308-8146(95)00021-A

Johnson, A. (2016).

Khiari, D., Suffet, I. H., & Barrett, S. E. (1995). The determination of compounds causing fishy/swampy odors in drinking water supplies. *Water Science and Technology, 31*(11), 105-112. doi:https://doi.org/10.1016/0273-1223(95)00463-W

Li, X., Yu, J., Guo, Q., Su, M., Liu, T., Yang, M., & Zhao, Y. (2016). Source-water odor during winter in the Yellow River area of China: Occurrence and diagnosis. *Environ Pollut, 218*, 252-258. doi:10.1016/j.envpol.2016.06.069

Ma, Z., Niu, Y., Xie, P., Chen, J., Tao, M., & Deng, X. (2013). Off-flavor compounds from decaying cyanobacterial blooms of Lake Taihu. *Journal of Environmental Sciences, 25*(3), 495-501. doi:10.1016/s1001-0742(12)60101-6

Noblet, J., Schweitzer, L., Ibrahim, E., Stolzenbach, K. D., Zhou, L., & Suffet, I. H. (1999). Evaluation of a Taste and Odor Incident on the Ohio River. *Water Science and Technology, 40*(6), 185-193.

Nystrom, A. (1992). Drinking Water Off-flavour Caused by 2,4,6 Tricholoroanisole. *Water Science & Technology, 25*(2).

Ömür-Özbek, P. (2012). *Global Taste and Odor Survery of Water Utilities*. Retrieved from

Piet, G. J., Slingerland, P., Degrunt, F. E., Zoeteman, B. C. J., & Vanderheuvel, M. P. M. (1978). DETERMINATION OF VERY VOLATILE HALOGENATED ORGANIC-COMPOUNDS IN WATER BY MEANS OF DIRECT HEAD-SPACE ANALYSIS. *Analytical Letters, 11*(5), 437-448. doi:10.1080/00032717808067884

Piriou, P., Devesa, R., De Lalande, M., & Glucina, K. (2009). European reassessment of MIB and geosmin perception in drinking water. *Journal of Water Supply: Research and Technology—AQUA, 58*(8). doi:10.2166/aqua.2009.124

Rashash, D. (1997). FPA of Selected Odorous Compounds. *American Water Works Association Journal, 89*(4).

Skjevrak, I., Due, A., Gjerstad, K. O., & Herikstad, H. (2003). Volatile organic components migrating from plastic pipes (HDPE, PEX and PVC) into drinking water. *Water Research, 37*(8), 1912-1920. doi:10.1016/s0043-1354(02)00576-6

Stocking, A. J., Suffet, I. H., McGuire, M. J., & Kavanaugh, M. C. (2001). Implications of an MTBE odor study for setting drinking water standards. *Journal American Water Works Association, 93*(3).

Suffet, I. H.(1995). Advances in Taste and Odor Control. *American Water Works Research Foundation*

Suffet, I. H., Khiari, D., Bruchet, A. (1999). The drinking water taste and odor wheel for the millennium: Beyond geosmin and 2-methylisoborneol. *Water Science and Technology,* 40(6), 1-13. doi:10.1016/S0273-1223(99)00531-4.

Suffet, I. H. (2007). A re-evaluation of the taste and odour of methyl tertiary butyl ether (MTBE) in drinking water. *Water Science and Technology, 55*(5), 265-273. doi:10.2166/wst.2007.188

Tipping, M. E., & Bishop, C. M. (1999). Probabilistic Principal Component Analysis. *Journal of the Royal Statistical Society. Series B (Statistical Methodology), 61*(3), 611-622.

Tomboulian, P., Schweitzer, L., Mullin, K., Wilson, J., & Khiari, D. (2004). Materials; used in drinking water distribution systems: contribution to taste-and-odor. *Water Science and Technology, 49*(9), 219-226.

Vega, M., Pardo, R., Barrado, E., & Debán, L. (1998). Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research, 32*(12), 3581-3592. doi:https://doi.org/10.1016/S0043-1354(98)00138-9

Ventura, F., Quintana, J., Gomez, M., & Velo-Cid, M. (2010). Identification of alkyl-methoxypyrazines as the malodorous compounds in water supplies from Northwest Spain. *Bull Environ Contam Toxicol, 85*(2), 160-164. doi:10.1007/s00128-010-0053-6

Watson, S. B. (2003). Cyanobacterial and eukaryotic algal odour compounds: signals or by-products? A review of their biological activity. *Phycologia, 42*(4), 332-350. doi:10.2216/i0031-8884-42-4-332.1

Watson, S. B., Chariton, M., Rao, Y. R., Howell, T., Ridal, J., Brownlee, B., . . . Millard, S. (2007). Off flavours in large waterbodies: physics, chemistry and biology in synchrony. *Water Science and Technology, 55*(5), 1-8. doi:10.2166/wst.2007.155

Watson, S. B., Monis, P., Baker, P., & Giglio, S. (2016). Biochemistry and genetics of taste- and odor-producing cyanobacteria. *Harmful Algae, 54*, 112-127. doi:10.1016/j.hal.2015.11.008

Young, C. C., & Suffet, I. H. (1999). Development of a standard method — Analysis of compounds causing tastes and odors in drinking water. *Water Science and Technology, 40*(6), 279-285. doi:https://doi.org/10.1016/S0273-1223(99)00569-7

Zhang, K., Luo, Z., Zhang, T., Mao, M., & Fu, J. (2016). Study on formation of 2,4,6-trichloroanisole by microbial O-methylation of 2,4,6-trichlorophenol in lake water. *Environ Pollut, 219*, 228-234. doi:10.1016/j.envpol.2016.10.042

Zhang, L., Hu, R., & Yang, Z. (2005). Simultaneous picogram determination of "earthy-musty" odorous compounds in water using solid-phase microextraction and gas chromatography-mass spectrometry coupled with initial cool programmable temperature vaporizer inlet. *J Chromatogr A, 1098*(1-2), 7-13. doi:10.1016/j.chroma.2005.08.053