The potential for micropollutants and microbial indicators as tracers of domestic sewage contamination along the Bow River

by

Jaime Danielle Hicks

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Environmental Engineering

Department of Civil and Environmental Engineering
University of Alberta

Abstract

Maintaining water quality is essential for overall watershed health. The **One Water** approach recognizes the interconnectedness of water resources, emphasizing the importance of integrated water resource management. A key threat to water quality is fecal contamination from human and non-human sources. Traditionally, fecal indicator bacteria (**FIB**) have been used for contamination analysis, but combining FIB with other microbial indicators (e.g., microbial source tracking [**MST**] markers) and a set of chemical indicators (e.g., micropollutants) offers multiple lines of evidence to assess fecal contamination sources in aquatic environments. The main objective of this thesis was to identify baseline conditions of microbial and chemical indicators at three sites near Calgary along the Bow River (**BR**): 1) an upstream source water site (**BR2**), 2) an urban site (**BR3**), and 3) a site downstream of the City's three wastewater treatment plants (**WWTPs**) (**BR4**). This thesis also developed site-specific fingerprints comprising a set of microbial and chemical indicators that differentiate samples from the three sites. Data processing and visualization of three FIB, 9 MST markers, and 56 chemical indicators collected from 2018 to 2023 were completed using the Python programming language. Parameters with less than seven numerical detections were excluded. Preliminary analysis indicated that the data was mostly non-normal (via the Shapiro-Wilk, D'Agostino's K^2, and Jarque Bera normality tests). Hence the $\log_{10}$ data transformation was used, with the exception of BR2-spring, which utilized the 1/x transformation. The Pearson correlation method was applied to the data in addition to the Ward clustering linkage method. Analysis of the microbial and chemical indicators revealed distinct patterns at the three sample sites. BR4 had the highest diversity of indicators, while BR2 showed minimal impacts. The presence of FIB, MST markers, and chemical indicators at all sites indicates varying levels of human and non-human fecal contamination. Six chemical indicators (caffeine, N,N-diethyl-meta-toluamide [DEET], metformin, O-Desmethyl-venlafaxine [ODV], sucralose, sulfamethoxazole [SMX]) were consistently detected. Total coliforms were positively correlated with the flow at all sites, likely due to increased inputs during high-flow events (i.e., flow enrichment). However, sucralose was found to be inversely correlated with the

flow, highlighting dilution effects. Overall, enterococci had higher median concentrations compared to the other FIB, and sucralose dominated the detected chemical indicators across all datasets. The presence of MST markers varied across the sites, with Rum2Bac (ruminant) dominating at BR2, CG01 (Canada goose) at BR3, and HF183 (human) at BR4. Additional relationships were revealed during the correlation and cluster analysis. For instance, at BR2 total coliforms and several chemical indicators (including metformin, ODV, and sucralose) were negatively correlated indicating different sources and/or pathways of entry into BR2. Additionally, the clustering analysis identified three primary clusters: 1) a chemical indicator cluster (ODV, metformin, sucralose), 2) a FIB cluster (*E. coli*, total coliforms), and 3) a combination of microbial and chemical indicators (caffeine, enterococci, Rum2Bac, DEET, SMX), suggesting potential co-occurrence of these indicators. Increased correlations between *E. coli* and enterococci as well as between total coliforms and caffeine from BR2 to BR3 were observed. Three of the four main clusters at BR3 contained microbial and chemical indicators, while one cluster was limited to chemical indicators exclusively. Downstream of the WWTPs at BR4, strong positive correlations were seen between *E. coli* and chemical indicators, while HF183, HumM2 (human), and total coliforms were negatively correlated with various chemical indicators, aligning with previous studies related to wastewater discharges. The diverse array of chemical indicators at BR4 also underscores the cumulative impacts of wastewater discharges at the site. The FIB (*E. coli*, total coliforms, enterococci), Rum2Bac, metformin, ODV, and sucralose are recommended indicators for site-specific fingerprinting due to their presence at all sites. Additionally, CG01, HF183, HumM2, caffeine, carbamazepine and diclofenac were added to the site-specific fingerprinting recommendations due to their unique presence at BR3 and BR4, indicating that their detection at elevated levels at BR2 can trigger additional source water protection initiatives.

Preface

This thesis is an original work by Jaime Danielle Hicks under the supervision of Dr. Maricor Arlos. Aspects of this work have been presented at the Society of Environmental Toxicology and Chemistry (SETAC) – Prairie Northern Chapter 2024 Annual General Meeting in Edmonton, AB. This work was funded by the NSERC Alliance grant (ALLRP 567652-21), the NSERC Discovery grant (RGPIN-2021-02412) and the Alberta Innovates - Water Innovation Project. No part of this thesis has been previously published.

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Dr. Maricor Arlos, for her invaluable guidance and support. Her expertise and insights have been instrumental to this work.

I am also deeply indebted to the following individuals for their contributions to this research. To Dr. Norma Ruecker and Eric Camm from the City of Calgary for their expertise, assistance, and encouragement. To Candis Scott, for sharing her knowledge of microbiology and Daniela Pulgarin Zapata for teaching me how to generate the site map in QGIS. Additional thanks to Bryanne Rempel from the City of Calgary for sharing the raw data used for this research, Joe Holstein for providing information about the wastewater treatment process in Crossfield, AB, and Willow from the Canadian Wildlife Service for assistance with Canadian migratory bird information.

I would like to extend a special thanks to my examiners, Dr. Norman Neumann and Dr. Jennifer Nafziger.

Finally, I would like to thank my family for their motivation and support.

Table of Contents

List of Abbreviations

AC/TC	Atypical colonies/total coliforms
AHTN	Tonalide
Amp^r	Ampicillin-resistant
Anthropog.	Anthropogenic DNA marker
BMPs	Best management practices
BPA	Bisphenol A
BPS	Bisphenol S
BR	Bow River
CCE	Calibrator cell equivalent
CSO	Combined sewer overflow
CST	Chemical source tracking
DEET	N,N-diethyl-meta-toluamide
DNA	Deoxyribonucleic acid
DO	Dissolved oxygen
ER	Elbow River
ESS	Sum of squares error
FIB	Fecal indicator bacteria
FWAs	Fluorescent whitening agents
GR	Glenmore Reservoir
HHCB	Galaxolide
Hmt	Human mitochondrial DNA
IQR	Interquartile range
KM	Kaplan-Meier
LABs	Linear albylbenzenes
LID	Low-impact development
MBA	Multi-barrier approach
MLE	Maximum likelihood
MPN	Most probable number
MST	Microbial source tracking
NaN	Not a Number
NC	Nose Creek
OBs	Optical brighteners
ODV	O-Desmethyl-venlafaxine
PAHs	Polycyclic aromatic hydrocarbons
PCBs	Polychlorinated biphenyls
PCR	Polymerase chain reaction
PF	Policeman's Flats
PFAS	Perfluoroalkyl and polyfluoroalkyl substances
PFCs	Perfluorinated compounds
PFGE	Pulsed-field gel electrophoresis
PFOA	Perfluorooctanoic acid
PFOS	Perfluorooctanesulfonic acid
PPCPs	Pharmaceuticals and personal care products
Q1	First quartile
Q3	Third quartile
qPCR	Quantitative PCR
R^2	Pearson correlation coefficient
RLs	Reporting limits
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
rep-PCR	Repetitive element sequence-based PCR
SMX	Sulfamethoxazole
SRCs	Sulphite-reducing clostridia
SWP	Source water protection
TCEP	Tris(2-carboxyethyl)phosphine
TCPP	Tris(chloropropyl) phosphate
TSS	Total suspended solids
UV	Ultraviolet
UWM	Urban water management
WWTP	Wastewater treatment plant
Zoog.	Zoogenic DNA marker

List of Tables

List of Figures

Chapter 1 – Introduction

1.1 Background

The multi-barrier approach (**MBA**) to drinking water is a comprehensive strategy for ensuring the safety and quality of drinking water from source to tap [1]. Water-borne disease outbreaks in Walkerton, ON and North Battleford, SK highlighted the critical need for this approach and brought it to the forefront of water management in Canada [2], [3], [4]. Designed to have integrated approaches and points of intervention, MBA consists of three primary components: 1) source water protection (**SWP**), 2) drinking water treatment, and 3) drinking water distribution. SWP is a critical first step in MBA as it safeguards the quality and quantity at its source (e.g., lakes, rivers, groundwater, reservoirs) to prevent potential contamination and depletion. At its core, SWP recognizes that protecting water at the source is more economically effective compared to relying solely on water treatment and purification [5], [6]. However, traditional approaches to water management by treating stormwater, wastewater, and drinking water as separate entities have been unable to address the challenges posed by rapid population growth, urbanization, and climate change. The **One Water** paradigm offers a more holistic approach by recognizing the interdependence of water resources while promoting integrated management strategies [7], [8], [9]. This approach aligns with the principles of MBA, which emphasizes the importance of protecting water at its source and integrating impacts from stormwater and (un)treated wastewater discharges.

Water supplies around the globe, including in many Canadian urban centres, continue to encounter significant challenges due to rapid population growth. For example, as cities continue to expand, source water quality may be compromised when the distance between wastewater discharge points and drinking water sources is reduced [10], [11]. Stormwater stemming from drainage surfaces (e.g., overland transport) and urban drainage systems (e.g., stormwater outfalls) may also impact source water quality by transporting high contaminant loads with limited treatment during wet-weather events [12], [13]. These conditions are further exacerbated by climate change-induced stressors, such as

droughts, heavy rainfall, rising temperatures, and increased wildfires, therefore, incorporating climate change adaptation measures within SWP plans remains a major challenge faced by municipalities worldwide [4], [14].

The City of Calgary in southern AB, Canada has taken a proactive approach to addressing risks to its source waters by implementing an SWP Plan (2015-2018) and adopting an SWP Policy (2020) which integrates watershed protection into future land use changes [5], [15]. Among the SWP Plan's goals is to manage stormwater effectively. As was common in many large cities, combined drainage pipes involving both sanitary (domestic sewage) and storm drainage were originally built in Calgary in 1890. However, the separation of these systems took place between the 1920s and 1960s [16]. Though Calgary utilizes separated sewer systems today, the literature suggests that storm drainage systems can be contaminated with untreated wastewater, particularly when they experience dry-weather flow conditions [17], [18], [19]. Cross-connections between the sewage systems as well as leakages from aging or deteriorating sewer systems may transport pollutants into the stormwater system or natural environment without treatment [20].

The potential for contamination from sewage systems into stormwater systems underscores the importance of accurate detections when identifying the origins of fecal pollution within water supplies and urban catchments. Fecal source tracking is important in the context of Calgary's SWP plan given the inputs of stormwater into its source waters (i.e., Glenmore [Elbow River (ER)] and Bearspaw [Bow River (BR)] reservoirs). Source water microbial quality is highly variable and influenced by climate and hydrometeorological events such as rainfall, snowmelt, bypasses, and sanitary sewer overflows during storm events [21], [22], [23]. For many years, fecal indicator bacteria (**FIB**) (e.g., *E. coli*, fecal coliforms, enterococci) have been used as surrogates for potentially harmful pathogens in environmental waters, however, they do not discriminate among the many possible sources [24], [25]. Hence, host-associated microbial source tracking (**MST**) methods have evolved to identify the originating host species such as

birds, humans, cattle or a combination by utilizing various human and non-human-associated microbial markers [26], [27], [28].

Other studies have also utilized chemical source tracking (**CST**) tools to assess fecal contamination [28], [29]. While CST can include a broad range of chemicals including human-specific fecal sterols and stanols and other human biomarkers, the utility of pairing MST data with wastewater micropollutants, such as pharmaceuticals (e.g., antidepressants, anti-inflammatories), personal care products (e.g., antibacterials), endocrine disrupting compounds (e.g., natural and synthetic hormones), and other household/industrial chemicals (e.g., flame retardants), can serve a dual purpose [20], [28], [24]. Since 2018, the City of Calgary has monitored micropollutants to determine baseline conditions and how they behave in the BR and ER watersheds [5]. Although micropollutant monitoring is not mandatory for water utilities in Canada, some jurisdictions have established human health (e.g., perfluoroalkylated substances [PFAS]) and environmental guidelines (e.g., birth control pill ingredients), which promote a more prudent approach towards water quality management. Most micropollutants are derived mainly from domestic uses and some remain persistent during wastewater treatment (e.g., metformin [antidiabetic], sucralose [artificial sweetener]) [31]. Therefore, pairing these datasets with MST results can provide multiple lines of evidence to detect sources and impacts of (un)treated sewage and support risk assessments. Most micropollutants are also frequently detected in surface waters and some chemicals exhibit greater source specificity and geographical stability in comparison to most established MST tools currently used [29], [32], [33]. However, there is no consensus on a single fecal source tracking tool that water utilities should employ. Each one has its advantages and limitations, but both tools remain complementary methods for identifying the source(s) of fecal contamination in water systems.

1.2 Problem Statement

The City of Calgary measured 56 micropollutants (see Table 3.5) as well as microbial indicators including general fecal indicators (3 FIB [culture-dependent, see Table 3.3] and 9 host-associated

microbial markers [i.e., MST - molecular techniques, see Table 3.4]) at three sites along the BR near Calgary from 2018 to 2023 (Figure 1.1). Given the highly urbanized land use and diverse pollution inputs into the BR, whether the microbial surface water quality indicators correlate with the micropollutant data is unclear. Each of these sites are influenced by distinct land use patterns, contributing to a unique combination of indicators (i.e., fingerprint) due to varying inputs resulting in differences in presence and concentrations. In this thesis, a composite **fingerprint** includes a unique combination of FIB, MST markers, and chemical indicators specific to the site that can be linked to a particular fecal contamination source.

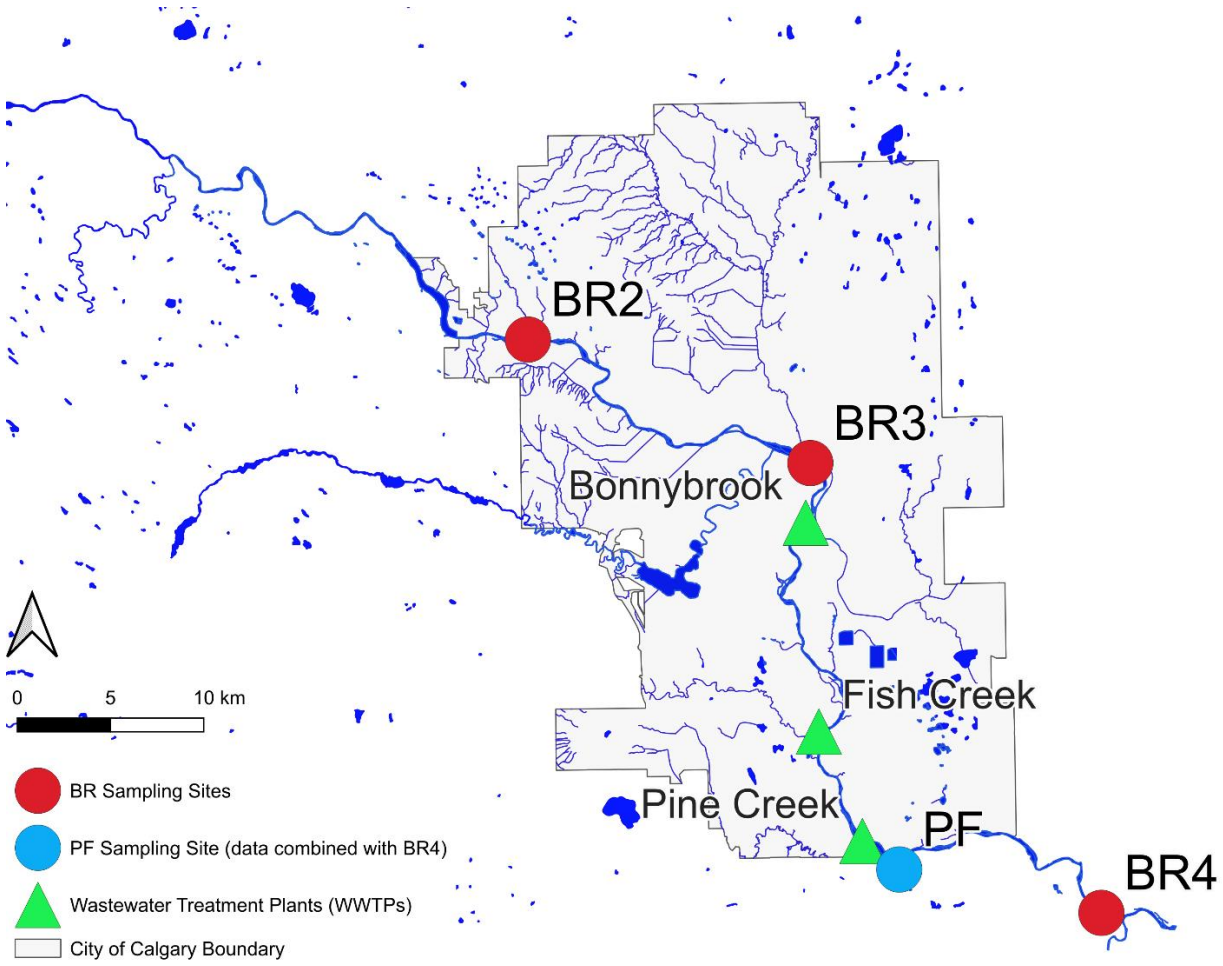


Figure 1.1 *Map of research sites along the Bow River (BR) near Calgary. PF = Policeman's Flats.*

1.2 Thesis Objectives

The main purpose of this thesis is to use microbial and chemical indicators to identify baseline conditions of anthropogenic fecal contamination at three sites along the BR near Calgary. To achieve this, the following objectives were developed:

i. investigate relationships between the microbial and chemical markers at three distinct (yet hydrologically connected) sites in and around Calgary consisting of 1) an upstream source water site (BR2), 2) an intermediate urban site (BR3), and 3) a cumulative downstream site that receives treated municipal wastewaters (BR4);
ii. develop a composite fingerprint comprising traditional fecal (FIB), microbial (MST) markers and chemical (CST) indicators to differentiate between samples taken from the three sites; and
iii. provide initial recommendations for monitoring approaches that can be utilized by the City of Calgary when assessing domestic sewage contamination along the BR.

1.3 Thesis Scope

Although microbial and micropollutant data have been measured at 11 sites in the ER and BR watersheds [31], this thesis focuses on only three sites (BR2, BR3, BR4) to test the composite fingerprint hypothesis, which suggests that each site has unique FIB, MST, and CST profiles that can be used to differentiate water samples. Site descriptions are provided below. Chapter 2 is a literature review on source water microbial quality and fecal source tracking tools. Chapter 3 explains the methodology including database management, data filtering, data visualization, hierarchical clustering, statistical, correlation and clustering analyses. Chapter 4 discusses the results and outcomes related to developing a composite fingerprint. Finally, Chapter 5 summarizes the key messages and outlines recommendations for future work.

1.4 Site Descriptions

As of 2022, the City of Calgary in southern AB, Canada had a population of 1.4 million [34]. Its drinking water is mainly supplied by the BR and the ER source watersheds. The ER supplies 40% of Calgary's drinking water and feeds into the Glenmore Reservoir (GR), a human-made lake that provides flood protection, serves as a non-motorized boating and recreational area, and is the intake source for the Glenmore Water Treatment Plant [5], [35], [36]. The remaining 60% of the City's drinking water is supplied by the BR, which is treated and distributed by the Bearspaw Water Treatment Plant [36].

Briefly, three monitoring sites (see Table 3.2) were selected among the 36 active watershed surface water quality monitoring locations within and around Calgary [37]. Site selection was curated to include sites with relatively different impacts, potentially allowing for clear identification of variations when identifying the distinct fingerprint for each site.

The most upstream site, BR2, is the intake source of the Bearspaw Water Treatment Plant, and it receives water from the BR which flows through smaller communities such as Canmore and Cochrane, AB [36], [38]. Since the nearest wastewater plant discharge point upstream of BR2 is ~66 km away from Calgary, potential sewage contamination containing high concentrations of persistent chemicals (i.e. pharmaceuticals, antidepressants) would likely originate from cross-connections with stormwater or septic fields from residential areas. As a result, this site is expected to have the least number of detections and magnitudes of microbial indicators and micropollutants.

BR3 is located within the City boundary and is impacted by the surrounding high-density residential and core commercial activities [39]. It also receives inputs from Nose Creek (NC), the BR, and the lower ER through 287 stormwater outfalls within Calgary. NC is known to be contaminated since it carries stormwater from Airdrie, AB, and is also impacted by the urban runoff from Calgary International Airport [40], [41]. In addition, the site is influenced by treated wastewater from Crossfield, AB (population: 3,599) [42], which is stored in a sewage lagoon and discharged once per year for 21 consecutive days within the timeframe of April 1st and November 30th [43], [44]. This site is expected to

reflect influences of high urbanization and both older (i.e., upstream wastewater treatment plant [WWTP] discharges) and untreated wastewater (i.e., NC inputs, stormwater cross-connections with sanitary sewers).

The third site, BR4, is located downstream of Calgary's three WWTPs: 1) Fish Creek, 2) Pine Creek, and 3) Bonnybrook. A prior study at these sites indicated a high frequency and magnitude of micropollutant detections along the BR and therefore likely provides a unique chemical (i.e., micropollutant) fingerprint [31]. The three WWTPs also utilize disinfection prior to discharge into the BR which is different from the type of sewage that BR3 might be exposed to. Based on this hypothesis, the MST fingerprint at BR4 may vary due to the limited micropollutant removal by the WWTPs.

Chapter 2 – Literature Review

2.1 Concept and Applications of the One Water Approach

Water quality management is critical for the provision of safe drinking water [45], [46] and the presence of contaminants in water has been shown to have negative impacts on human health with the potential to harm ecosystems [47]. However, effectively managing these substances can pose significant challenges. Traditional urban water management (**UWM**) strategies have relied on segregated systems required for drinking water treatment, stormwater drainage, and wastewater collection. Such systems were viewed as advantageous due to the cost-effectiveness of large-scale construction and the ability for water service trends to be monitored individually [7], [48]. Yet, these approaches have been experiencing stressors caused by various factors including population growth and climate change impacts (i.e., heatwaves, wildfires, floods) [8]. This confluence of factors has resulted in rising water demand in addition to declines in water availability and water quality in communities.

One innovative approach that has been proposed as a potential solution to address the limitations of traditional UWM systems is the concept of **One Water**. The One Water approach proposes that urban water (i.e., stormwater, groundwater, surface water, drinking water, wastewater) be viewed and treated as part of one interconnected system rather than in terms of separate management systems [7], [8], [9]. A more integrated approach to UWM would be to promote strategies to harvest stormwater, reduce water usage, and minimize waste generation through six guiding principles: 1) recognizing the value of all types of water, 2) utilizing strategies with multiple benefits, 3) investing in systems-based approaches, 4) implementing fit-for-purpose water use, 5) developing adaptive infrastructure, and 6) engaging the appropriate stakeholders [9]. By viewing all types of water and their by-products as resources, this approach can improve water security while promoting financial resilience and supporting the circular economy [7]. However, various barriers can hinder the adoption of the One Water approach. These include institutional barriers (i.e., fragmented government structures, outdated regulations), limited public

knowledge, political processes, inadequate resources, and a lack of necessary infrastructure. These obstacles, in addition to economic and policy considerations, underscore the need for collaborative efforts in order for the successful implementation of the One Water approach [49], [50].

The One Water guiding principles outline best practices for water policies and management strategies, reflecting shared values around water as a vital resource. These principles provide a framework for developing sustainable and equitable water management systems [51]. The first guiding principle recognizes the inherent values of all types of water. Such value can be categorized as economic, environmental, and social. Economic value is derived from the fact that water is key in the development and production processes within the natural resource sector, and underestimating the economic value of water has caused negative economic and environmental impacts (i.e., inefficient water use, resource depletion, release of untreated wastewater, etc.) [52]. The environmental value of water is seen in its regulation of water resources and facilitation of filtration, storage, and flow regulation. The worth of water is also connected with its role in sustaining life, as it is a fundamental part of organisms and is essential to life on Earth, while recreational, cultural, and spiritual attributes associated with water contribute to its social value [53]. Recognizing these values highlights the need for a holistic water management approach, treating all types of water (groundwater, surface water, stormwater, wastewater, and drinking water) as critical resources [9]. This approach includes integrating stormwater and wastewater treatment, reducing freshwater withdrawals, and treating and reusing certain waters (i.e., stormwater, greywater, wastewater). The implementation of this principle would require a shift from segregated water management systems to integrated strategies, understanding the connections between various water bodies and developing new, comprehensive management approaches [54].

The second principle emphasizes that multiple benefits are required within water management. Such an approach promotes environmental sustainability and can strengthen community resilience by optimizing available resources once water resources (i.e., wastewater, stormwater) have been integrated [9], [55]. One example is energy recovery from wastewater to produce heat and electricity while reducing

greenhouse gas emissions from the wastewater treatment process [8]. A circular economy approach to water values the resource for its various applications, unlike traditional methods which view water as a linear resource from source to end use [55]. The potential benefits from the implementation of this approach are dependent on the water resources available as well as the needs of the community (i.e. environmental, economic, or social) to address current challenges while ensuring water quality and security for future generations [8], [9], [54], [55].

System-based approaches, the third principle of the One Water framework, view water as a complex and interconnected system. Recognizing the interconnections from water resources to infrastructure and ecosystems enables the creation of effective and sustainable solutions [9], [55]. These solutions may include installing water-efficient appliances to reduce consumption, conducting water usage audits, optimizing infrastructure, and developing public awareness campaigns to promote water conservation and resource protection. From an industrial perspective, adopting water-efficient technologies, minimizing waste, and investing in watershed protection programs are key system-based approaches that can be used to implement the One Water concept [8], [54]. The successful implementation of system-based approaches at individual, community, and industrial levels along with collaboration among stakeholders, is essential for ensuring environmental health and sustainability of water resources [8], [9], [54], [55].

The fit-for-purpose principle tailors water treatment based on intended use, rather than uniformly treating all water to drinking water quality standards. This approach yields multiple advantages, including reducing treatment costs and energy consumption while also minimizing waste generation. Applications can include using reclaimed water for irrigation, industrial processes, and other non-potable purposes when high-quality water is not necessary [8], [54]. One example of this is using domestic reclaimed water for toilets and urinals in residential and commercial settings [56], [57], [58]. While successful water reuse projects have demonstrated the potential to address the challenges of increasing water scarcity and elevated water demands, high initial costs and delayed benefits have been barriers to widespread

implementation [58]. Therefore, a long-term perspective, strategic planning and stakeholder collaboration are essential for developing efficient water management systems [8], [9], [57], [58].

Developing adaptive or flexible infrastructure is the fifth principle in the One Water approach, essential for navigating the uncertainties caused by climate change. Traditional and aging infrastructure lacks the capacity to adjust to evolving conditions and demands, while adaptive systems are designed to meet new challenges [9]. Examples of this include: 1) incorporating green infrastructure (i.e., rain gardens, green roofs) and low-impact developments (LIDs) to manage and improve water quality in urban settings and 2) designing water and wastewater facilities with modular components that can be expanded or modified to accommodate changing water conditions [8], [55], [59]. Investing in adaptive infrastructure also promotes equity and resilience to climate change and other external impacts, while reducing long-term costs and enhancing water security [8], [9], [54], [60].

The sixth and final principle of the One Water approach is stakeholder engagement, which necessitates the collaboration of various groups, including community members, industry representatives, water professionals, and government agencies [8], [9]. This collaboration brings together diverse perspectives and expertise to enhance decision-making and develop appropriate solutions. Unlike the One Water approach, traditional segregated UWM systems do not prioritize stakeholder engagement collaboration, impacting resource allocation, the quality of water services, and overall sustainability [61]. Shared decision-making and emphasizing partnerships between the appropriate stakeholders allow for the development of resilient and sustainable water management systems [8], [9], [61].

Many locations around the world have developed sustainability and/or action plans, which include initiatives allowing them to actively transition to the One Water approach (e.g., Chicago, IL, USA; Copenhagen, Denmark; New York, NY, USA; Phoenix, AZ, USA; Rotterdam, Netherlands; Singapore; Sydney, Australia) [62], [63], [64], [65], [66], [67], [68], [69]. Table 2.1 outlines the One Water initiatives both proposed and implemented at the locations mentioned, as well as the targeted principles. It is clear that initiatives vary between locations due to their unique needs, resources, and

environmental conditions. The City of Calgary is one Canadian city that has begun to incorporate elements of the One Water approach into its water management strategy by identifying initiatives and priority actions required to maintain water security [70]. Climate change-related impacts (i.e. elevated temperatures, droughts, wildfires), regulatory requirements, population and economic growth have been identified as posing potential risks to the City's water security [70], [71]. The City articulates the need to develop a sustainable and resilient water supply system by planning for various demand scenarios, collaborating with stakeholders, and utilizing adaptive management approaches to maintain water security [70].

Table 2.1 *One Water case studies in North America, Europe, and Australasia along with the guiding principles associated with the specific initiatives.*

Location	One Water initiative	Guiding principle	Description	Reference
Chicago, IL, USA	Strengthen gray and green infrastructure to withstand climate change	2, 3, 5	Improve emergency response, adopt sustainable practices, and protect infrastructure from future impacts	[62]
	Improve water resource management and coordination	1, 3, 5, 6	Increase collaboration, data-driven decision-making, and planning across agencies	
	Improve water resource management and coordination	2, 3, 6	Utilize collaboration, data sharing, and planning across agencies and stakeholders to improve water resource management	
	Incorporate water resource management into local planning	1, 2, 3	Promote sustainable development, and protect water quality, quantity and ecosystems through local planning	
	Create and implement multi-objective watershed plans	2, 3, 6	Focus on watershed-based approaches, data-driven decision-making and stakeholder collaboration	
	Optimize water infrastructure investment	2, 3, 5, 6	Prioritize existing infrastructure, invest in resource recovery, promote stakeholder collaboration	
	Maintain and invest in gray and green infrastructure	2, 3, 5	Invest in and maintain green infrastructure to enhance flood resilience	
	Incorporate water supply and demand considerations into local and regional planning	2, 3, 5, 6	Protect water quality by aligning land use planning with water resource management	

Location	One Water initiative	Guiding principle	Description	Reference
Chicago, IL, USA	Strengthen regional water supply management	2, 3, 5, 6	Implement sustainable water management strategies through data sharing and collaborative planning	[62]
	Maintain drinking water infrastructure and manage demand	2, 4, 5, 6	Invest in and maintain drinking water infrastructure, promote water conservation, and explore alternative water sources	
Copenhagen, Denmark	Energy production – carbon neutral district heating	4	Reduce water consumption by using alternative cooling methods	[63]
	Energy production – carbon neutral utilities	2	Optimizing heating systems by using biogas from wastewater treatment, reducing carbon emissions	
	City administration initiatives – training and information	6	Facilitating a program which allows young individuals to learn about water-related issues	
	City administration initiatives – City of Copenhagen's woodlands	1, 3, 5, 6	Safeguard water resources by expanding tree coverage in water catchments and around drinking water wells	
New York, NY, USA	Protecting wastewater treatment facilities from storm surges	2, 3, 5	Safeguard public health, protect the environment, prevent damages caused by future storms	[64]
	Improving and expanding drainage infrastructure	2, 3, 5	Improve drainage infrastructure and incorporate green infrastructure to mitigate flooding risks	
	Promoting redundancy and flexibility to ensure a constant supply of high-quality water	2, 3, 5	Use the MBA, incorporate redundancy, and leverage advanced technologies to safeguard water supplies	
Phoenix, AZ, USA	Celebrate and protect rivers, washes, and waterways	1, 2, 3, 5, 6	Preserve natural waterways, improve public access, and collaborate with stakeholders to protect water resources	[65]
	Manage and plan for efficient delivery of safe and reliable water supplies	1, 2, 3, 4, 5, 6	Ensure reliable, high-quality water is available to the public by promoting efficient water use, investing in infrastructure, and collaborating with stakeholders	
	Managing stormwater efficiently and economically while minimizing stormwater pollution	2, 3, 5	Improve stormwater management through innovative practices, the development of new and green infrastructure, and public engagement	

Location	One Water initiative	Guiding principle	Description	Reference
Phoenix, AZ, USA	Treat, manage and use wastewater and related infrastructure efficiently and economically	1, 2, 3, 4, 5, 6	Promote sustainable development, infrastructure investment, and water reuse while minimizing environmental impacts	[65]
Rotterdam, Netherlands	Changing the water narrative	1, 6	Rebranding water from a nuisance to an opportunity by reshaping the narrative and engaging various stakeholders	[66]
	Cross-sectoral collaboration	2, 3, 6	Address urban challenges through cross-sectoral collaboration and generating innovative solutions	
	Co-production of knowledge	6	Engage diverse stakeholders to address gaps in the knowledge and develop new ways to address water management	
	Experiential evidence-based learning	2, 3, 6	Showcase projects highlighting novel water management solutions while engaging the public	
	Strategic use of trusted science	3, 6	Invest in large-scale research projects to accelerate water management transformation	
	Fostering and investing in networks	1, 3, 6	Foster knowledge sharing and collaboration from stakeholders across cities	
	Generating business from science-based innovation	1, 2, 3, 5, 6	Use water as an economic driver to develop sustainable water management	
Singapore	NEWater	1, 2, 3, 4, 5, 6	Investing in advanced technologies, strong governance, and public engagement to transform treated wastewater into a high-quality water resource	[67], [68]
Sydney, Australia	Decentralized water master plan	1, 2, 3	Reduce potable water consumption, improve water quality, and enhance urban resilience by implementing decentralized management strategies	[69]
	Stormwater harvesting and reuse	1, 2, 3	Capturing and utilizing stormwater for irrigation and other purposes	
	Water efficiency programs	1, 4	Encourage water-saving measures in residential, commercial, and industrial sectors	
	Foster relationships between the public and private sectors	6	Encourage public and private sectors to collaborate to develop and operate water recycling infrastructure	
	Green infrastructure	1, 2, 3	Integrate natural elements into urban environments to improve water quality and manage stormwater	

2.2 The Multibarrier Approach (MBA) in Water Quality Management

While the One Water approach adopts a holistic view by valuing all types of water as essential resources, the MBA ensures the safety of drinking water from the source to the consumers through five key elements: 1) source water protection, 2) effective water treatment, 3) reliable distribution systems, 4) continuous water quality monitoring, and 5) prompt responses to poor water quality [72], [73]. Employing multiple barriers minimizes the risk of waterborne contamination in the event that one barrier fails [74], [75]. Since public health and safety are paramount, the MBA is crucial for water utilities to maintain high water quality and prevent contamination [75]. While the MBA differs from the One Water approach, its strategies can be used to align with One Water principles. For instance, ensuring the safety of reclaimed or reused water for human consumption involves a combination of more advanced treatment processes, rigorous monitoring, and strict regulatory standards [76].

The MBA has been brought to the forefront of water quality management in Canada in light of the water-borne outbreaks in Walkerton, ON and North Battleford, SK [77]. The water crisis in Walkerton was caused by *E. coli* contamination from local farmland as a result of inadequate infrastructure and reporting, while the North Battleford water crisis was caused by insufficient wastewater treatment allowing *Cryptosporidium* to be released into the drinking water [2], [3]. As a result of these outbreaks, many individuals became ill, and seven individuals tragically died in Walkerton. The significant impact of these events prompted the development of the "Guidance on the Multi-Barrier Approach to Safe Drinking Water" (2004) as well as strong recommendations to rigorously implement MBA [73], [74], [76]. Such outbreaks in Canada and around the world highlight the need for an MBA in water quality management to ensure robust systems are in place to prevent contamination and safeguard public health [78].

2.3 Source Water Protection (SWP) Strategies

SWP is a key component of a successful MBA, aiming to safeguard drinking water sources from potential pollution [4], [79]. Watershed protection strategies vary depending on the type of source water available (i.e., surface water or groundwater) for municipal, industrial, and agricultural purposes. SWP strategies can be separated into two main categories: 1) LID which focuses on stormwater impacts on source waters, and 2) best management practices (BMPs) that manage impacts of industrial, municipal, and agricultural activities [79]. LID promotes the decentralization of on-site stormwater management through a wide range of technologies such as open bioswales, stormwater retention and detention ponds, green roofs, and erosion control. These technologies increase infiltration potential and allow for the removal of pollutants including, nutrients, heavy metals, and pathogens. Due to the diverse technologies available, LID can be implemented in both microscale and macroscale environments [80]. By contrast, BMPs are a broader category of SWP strategies that address water quality and hydrology challenges in industrial, urban, and agricultural areas. Examples of BMPs for SWP include constructed wetlands, bioretention systems, and nutrient management [79], [81].

SWP is an effective drinking water quality management strategy that reduces the entry of pathogens and chemicals into source waters, improving source water quality and reducing treatment costs and public health risks [79], [82], [83], [84]. Hence, efforts related to MST and CST can significantly enhance SWP management by identifying the origins of contamination. These techniques provide precise information about the sources and pathways of pollutants, enabling targeted and effective management strategies.

2.4 Microbial Source Tracking (MST) Methods to Indicate Fecal Contamination

As public health is a main driver for maintaining water quality, one strategy to identify fecal contamination and inform microbial risk assessment is to utilize a set of methodologies in environmental water quality microbiology to determine the origin of fecal pollution [85], [86], [87]. Early MST

methodologies were developed as a result of social and legal pressures [87] and attempted to discriminate between human and non-human fecal sources in environmental waters using microbial indicators such as fecal streptococci (i.e., *Enterococcus*), *E. coli*, and DNA genetic markers (i.e. *Bacteroides*, MST markers) [88], [89], [90], [91]. While FIB are valuable microbial indicators, they can originate from various human and non-human sources. Therefore the combination of FIB and more host-specific (i.e., cow, ruminant, seagull, dog) microbial (i.e., MST) markers are required for source attribution [24], [87].

There are two main categories of MST methodologies: 1) culture-dependent and 2) culture-independent. Culture-dependent methods require the cultivation of the bacteria in a laboratory prior to analysis. One example of a culture-dependent method is standard FIB analysis, first established in the late 1800s, which is a technique where the indicator bacteria (i.e., *E. coli*, total/fecal coliforms, *Enterococcus*) are enumerated [92], [93]. The three classical culturing methods used to enumerate and analyze FIB are 1) presence-absence tests, 2) the most probable number (MPN) method, and 3) membrane filtration. Presence-absence tests rely on fluorescent changes within the sample after it has been exposed to an enzyme substrate. A fluorogenic or chromogenic signal indicates the presence of the target FIB and no change signifies that the target FIB is not present in the sample. Prior to the development of the membrane filtration method, the MPN approach was used, which reports a statistical estimate of the number of FIB present in the sample. For the membrane filtration method, the water sample is filtered through a membrane with a diameter of 47 mm, retaining the FIB on its surface [93]. These culture-dependent methods are inexpensive, sensitive, and effectively amplify the target indicator. However, the lack of uniformly distributed bacteria, inadequate growth conditions and the varying growth rates of the indicators can pose a challenge [93], [94]. Culture-independent methods have evolved to incorporate deoxyribonucleic acid (DNA)/ribonucleic acid (RNA) extraction from the samples directly [26], [86], [87], [95]. Unlike culture-dependent methods, culture-independent methods have the ability to identify larger portions of the target microbial community, however, they can be less accurate when determining the number of viable microbes present in the sample [96].

Culture-dependent and independent MST methodologies can be further classified as library-dependent and library-independent. Library-dependent methods require the development of a reference library of fecal profiles (i.e. FIB and fecal strains) from target hosts present in and around the watershed which may serve as a potential source of fecal contamination. Phenotypic (i.e., biochemical) or genotypic profiles of the fecal bacteria collected are stored in such libraries, allowing researchers to use the distinguishing characteristics of the profiles within the library to identify contaminant sources in water samples collected from the watershed [24], [86], [87], [97], [98]. Here, considerations regarding the geographic and temporal stability of the library must be accounted for.

Molecular species diversities have been shown to result in variations in library profiles in instances of enterococci antibiotic resistance and certain strains of *E. coli*. Since different strains of indicator organisms have varying temporal stabilities, indicators must be selected based on their relevance within the context of the target host(s) of the watershed as well as the long-term stability of the indicators in the library, limiting temporal variations within profiled populations [87]. While not all phenotypic and genotypic MST methods are quantitative, they still allow for the identification of contamination from host groups. Additional limitations include the fact that reference strains are to be cultivated, and a geographic and temporal-specific reference library must be developed [87], [99].

Library-independent MST methods often rely on initial reference libraries for the characterization of host-specific markers from samples. Once the markers have been validated, they can be used as microbial indicators in similar environments, eliminating the need for the continual development of libraries for unique watersheds [100]. For example, this method allows researchers to target specific regions of the 16S ribosomal RNA (rRNA) gene in MST markers since they originate from human and non-human intestines (e.g., HF183 [human], CG01 [Canada goose], Rum2Bac [ruminant]). Unlike library-dependent methods, library-independent methods target specific markers associated with host groups [24], [86], [100]. Regardless of the approach, culture-dependent and culture-independent methods can be applied to library-independent MST. Table 2.2 outlines the common microbial methods for fecal

source tracking. Culture-dependent methods for FIB (i.e., *E, coli,* enterococci), various polymerase chain reaction (PCR) methods and automated enzymatic methods have been listed in the Guideline for Canadian Recreational Water Quality microbiological sampling and analysis technical document by Health Canada [101]. Various culture-based (i.e., cell culture assays, membrane filtration, presence-absence tests) and molecular-based (i.e. PCR) methods are proposed in the Canadian Drinking Water Quality bacteriological and microbiological parameter guide documents for *Giardia, Cryptosporidium,* enteric viruses, enterococci*, E. coli*, and total coliforms [102].

Table 2.2 *Common library-dependent and independent molecular microbial source tracking (MST) methods (adapted from* [26]*). PCR = Polymerase chain reaction.*

Culture dependency	Library dependency	Category	Approach	Description	Reference
Culture-dependent	Library-dependent	Phenotypic	Antibiotic resistance	Gut bacteria from various host animals develop unique antibiotic resistance patterns in response to varying antibiotic exposures	[26], [87], [95], [103]
			Carbon source utilization	Comparison of the unique patterns created when the target organism consumes carbon and nitrogen substrates	[26], [95], [103]
		Genotypic	Repetitive element sequence-based PCR (rep-PCR)	Uses PCR to amplify repetitive DNA sequences, resulting in strain-specific banding patterns	[87], [95], [103]
			Pulsed-field gel electrophoresis (PFGE)	Separates DNA molecules by using agarose gel and alternating electric fields to analyze the microbial genome directly	[26], [87], [95], [103]
			Ribotyping	Digestion of restriction enzymes is used to analyze variations in rRNA genes when differentiating between strain-specific branding patterns	[26], [87], [95], [103]
Culture-dependent	Library-independent	Phenotypic or genotypic	Bacteriophage	Distinguish between human and non-human fecal contamination based on the presence of the *E. coli* bacteriophage (F+RNA coliphage)	[95]
			Bacterial culture	Isolating and growing fecal indicator bacteria cultures	[26], [95]
Culture-independent	Library-independent	Genotypic	Host-specific bacterial PCR	Uses PCR to extract, amplify, and identify fecal sources based on target sequences in bacterial DNA (e.g. *Bacteroides, Enterococcus*)	[95]
			Host-specific viral PCR	Analyzes signature genes (e.g., 16S rRNA, antibiotic resistance) in fecal bacteria to differentiate between human and non-human sources	[26], [95]
			Host-specific quantitative PCR (qPCR)	Combines conventional PCR techniques of amplifying host-specific markers in fecal bacteria with quantification of the bacteria	[95], [103]

2.5 Chemical Source Tracking (CST) Methods

CST has been employed to identify the presence of anthropogenic fecal pollution in environmental waters [28], [87]. Many of these substances have anthropogenic origins and can provide unique signals depending on the land use within the area. Though using CST as a fecal source tracking method has emerged as an area of research in the last few decades, less focus has been placed on refining CST for such applications compared to FIB and MST methods [87], [104], [105].

Since there is no singular chemical that can serve as an indicator of anthropogenic pollution in all aquatic environments, four categories of chemical compounds have shown to have potential when identifying the presence of anthropogenic pollution: 1) pharmaceuticals and personal care products (PPCPs), 2) fecal sterols and stanols, 3) optical brighteners (OBs)/fluorescent whitening agents (FWAs), and 4) common wastewater tracers (i.e. caffeine, sucralose) [28], [87], [106].

PPCPs have been used as chemical indicators due to their unique presence in treated and untreated wastewater, though their concentrations may vary. While methodologies for PPCP measurement are considered to be reliable, standardized extraction procedures are still under development. In addition, analytical equipment (i.e., liquid/gas chromatography, mass spectrometry) and training costs may contribute to higher analytical costs compared to MST approaches [87]. Analgesics and anti-inflammatories (i.e., acetaminophen, ibuprofen, naproxen), antibacterials (i.e., triclocarban, triclosan), antidiabetics (i.e., metformin), antiepileptics (i.e., carbamazepine), beta-blockers (i.e., propranolol), and insect repellants (i.e., N,N-diethyl-meta-toluamide [DEET]) are a few of the PPCPs that have been detected in wastewater effluents and stormwater outfalls [28], [87], [107]. Research has shown that PPCPs can impact aquatic ecosystems and affect the growth and reproduction of aquatic populations. However, the persistence and degradation of these micropollutants as well as their public health impacts remains an ongoing area of research, further demonstrating the need to monitor for PPCPs from human wastewater pollution in the environment [87], [108].

Fecal sterols and stanols are another set of chemicals directly associated with sewage. Originating from the metabolism of dietary plant and animal sterols by enteric bacteria, sterols are part of the lipids family and are key components of cell membranes within eukaryotic organisms. Once consumed, dietary sterols are converted into stanols by bacteria within the intestine, while some sterols remain unchanged [104], [109], [110]. Both sterols and stanols are then excreted in animal and human waste, where they bind to particulate matter. Though fecal sterols have been shown to degrade in aerobic conditions, they have been detected in sediments [87], [111]. Since the metabolism of sterols produces end products that vary based on the diet and gut bacteria of the host, fecal sterol and stanol ratios can be analyzed to differentiate between human and non-human waste. For instance, approximately 60% of human stanols are comprised of sterol cholesterol which has been metabolized into 5 β-stanol, while non-human waste contains lower amounts of 5 β-stanol and contains wider ranges of sterol metabolic byproducts [112], [113]. Fecal sterol and stanol ratios can suggest that human sewage may be present and complement common water quality indicators (i.e., *E. coli*) in water quality analysis [114]. However, specialized equipment, a lack of standardized methods, and dilution concerns in large water bodies limit the widespread application of fecal sterols/stanols in CST [87].

OBs, also known as FWAs, are organic compounds found in laundry detergents which absorb ultraviolet (UV) light and re-emit light between the blue and indigo portions of the visible spectrum, enhancing the appearance of whites and brighter colours by counteracting yellow colours within the fabrics [87], [115], [116]. In addition to laundry detergents, OBs have been applied to textiles, plastics, synthetic fibres, and paper in medical, chemical, and petroleum-related applications [115]. Even though OBs have been known to photo-degrade once exposed to UV light and be removed by other means of disinfection (i.e. chlorination), they have been detected in wastewater effluents. Since OBs are not naturally occurring within the environment, they are clear indicators of anthropogenic pollution [87], [115], [116], therefore, they are a good candidate when identifying the presence of wastewater contamination in water due to their presence in household and industrial products. Additional analysis

may be required to determine whether the contamination contains human fecal matter. Detection methods currently available are simple and cost-effective, and results can be obtained quickly compared to the tests used to determine the presence of other micropollutants. However, OB dilution and in-stream fate/transport in large water bodies can lead to interference from other chemicals, highlighting the need for additional research to improve the accuracy and reliability of OBs as CST indicators [87], [105].

Caffeine (1,3,7-trimethylxanthine or $C_8H_{10}N_4O_2$) is a naturally occurring alkaloid stimulant found in various beverages (i.e., coffee, tea), food products, and medications, representing another potential chemical marker for anthropogenic contamination [87], [117]. It has been considered as a potential indicator of pollution due to its high consumption rates and subsequent releases into the wastewater system [117], [118], [119], [120]. Although caffeine has a half-life of 12 days when exposed to natural light [87], it has a half-life between 100 and 240 days when in water, demonstrating its potential persistence within aquatic environments [117], [120], [121]. Current detection methods for caffeine are highly sensitive, allowing for detection at low concentrations, and the presence of caffeine in water allows for the detection of recent contamination events, unlike certain microbial indicators (e.g., fecal coliforms). However, the natural degradation of caffeine in the environment as well as the presence of sources containing caffeine (i.e., coffee and tea vegetation) may affect detection, highlighting the need for additional indicators to be used in conjunction with caffeine [87], [117], [122].

In addition to the substances discussed above, many other chemical markers have been explored as potential source trackers for domestic sewage. These markers include but are not limited to, bile acids, polychlorinated biphenyls (PCBs), artificial sweeteners (i.e., acesulfame, sucralose), flame retardants, flavourants, fragrances, perfluorinated compounds (PFCs), pesticides and herbicides, petroleum hydrocarbons, plasticizers (i.e., bisphenol A [BPA], bisphenol S [BPS]), polycyclic aromatic hydrocarbons (PAHs), preservatives, solvents, and surfactants [87], [123], [124], [125], [126]. Persistence and detection of these compounds are variable within the environment, and their presence may not always be indicative of domestic sewage contamination. The broad range of chemical indicators that are being

analyzed for their potential as anthropogenic indicators demonstrates the need for multiple markers when determining the presence or absence of human fecal contamination. Hence, no single marker will be suitable for all cases; CST is dependent on site-specific factors, such as location, land use, and naturally occurring substances in the environment.

2.6 Integration of traditional (FIB), MST and CST Methods for Fecal Source Indication

A combination of traditional FIB, MST and CST methods has been proposed as a more reliable approach when determining the presence of anthropogenic contamination, as it utilizes multiple water quality indicators with varying persistence and fates within aquatic environments [28], [87], [127], [128]. Traditional methods have relied on *E. coli* or fecal coliforms; however, these do not differentiate between human and non-human sources. New microbial methods (i.e., MST) utilizing host-specific markers to identify fecal contamination sources were developed to address this limitation. CST complements MST by focusing on chemical indicators (i.e., PPCPs) associated with human waste [129], [130], [131]. Table 2.3 lists the advantages and limitations of MST and CST methods.

Table 2.3 *Advantages and limitations of microbial source tracking (MST) and chemical source tracking (CST) methods when employed separately.*

Approach			Advantages	Limitations	Reference
MST	Traditional		Simple and cost-effective Standardized protocols Easy interpretability of results	Inability to differentiate between human and non-human sources Potential underestimation of contamination due to non-culturable bacteria	[85], [86], [87]
	Molecular	Culture-dependent	Increased ability to differentiate between human and nonhuman sources Ability to detect non-culturable bacteria	The process of culturing the bacteria may be time-consuming Not as accurate for all sources	[85], [86], [87], [88], [89], [90], [91], [132], [133], [134]
	Molecular	Culture-independent	Increased specificity and sensitivity Rapid analysis time	Requires specialized equipment and expertise Increased cost compared to traditional MST methods	[85], [86], [87], [95], [96], [100], [101], [102], [103]
CST			Not affected by non-culturability or regrowth Can provide information regarding contaminant sources (i.e., pharmaceuticals, artificial sweeteners)	Site-specific-effectiveness is dependent on proper indicator selection based on location, land use, and naturally occurring substances Persistence and fate in the environment are variable Increased cost compared to traditional MST	[28], [87], [104], [105], [106], [135]

MST and CST encompass a wide range of indicators, and their respective limitations highlight the need for a combined approach to identify domestic sewage contamination. Leveraging the strengths of both methods can provide a comprehensive understanding of the presence and source of fecal contamination in aquatic environments. The combined approach can further improve the accuracy and effectiveness of fecal source identification [129], [130]. Figures 2.1 and 2.2 present the geographical

distribution of 28 case studies that employ a combined MST and CST approach to identify domestic sewage contamination in aquatic environments.

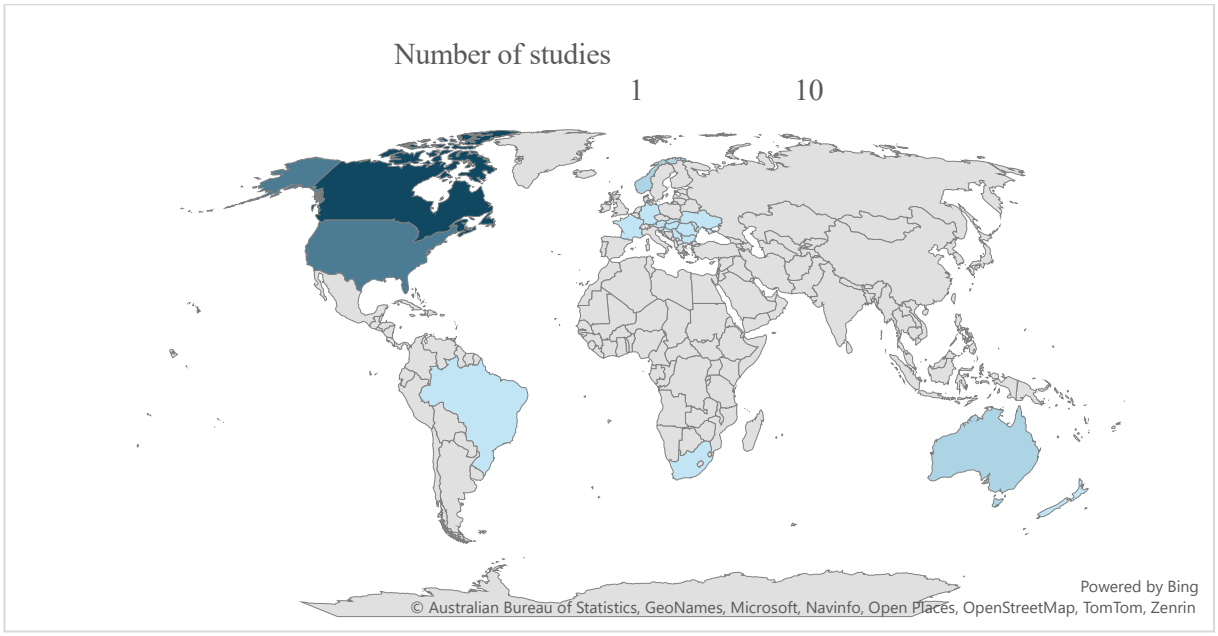


Figure 2.1 *Geographic distribution of studies which utilize a combined microbial source (MST) and chemical source tracking (CST) approach.*

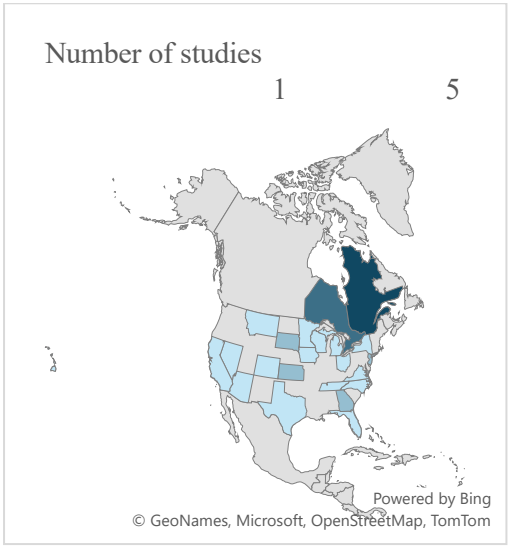


Figure 2.2 *Geographic distribution of studies in Canada and the USA which utilize a combined microbial source (MST) and chemical source tracking (CST) approach.*

A review of the 28 case studies which employed microbial and CST methods identified *E. coli* as the most prevalent microbial indicator (27 studies). Enterococci/*Enterococcus* and HF183 were also commonly employed (10 and 11 studies, respectively). Among the chemical indicators, caffeine was the most frequently analyzed (21 studies), followed by carbamazepine (16 studies) and acetaminophen (11 studies). The frequency of indicator usage in the case studies is shown in Figure 2.3.

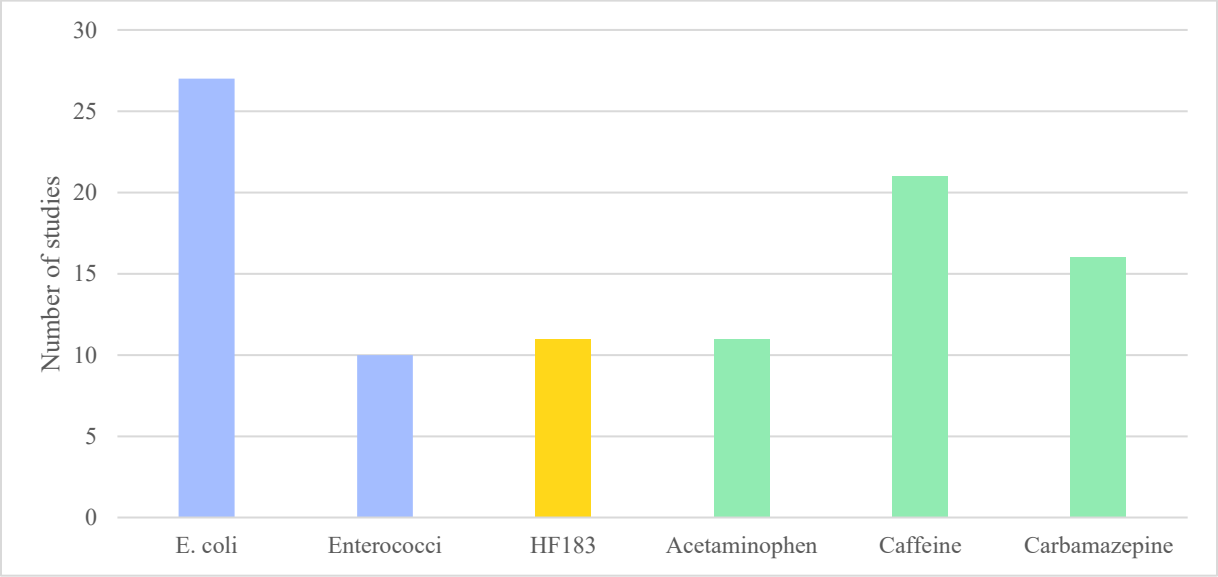


Figure 2.3 *Most commonly used microbial and chemical indicators across analyzed case studies.*

E. coli and enterococci are well-established fecal contamination indicators that have been used to determine the microbiological quality of water [106], [136] and inform public health decisions due to their strong relationship with gastrointestinal illnesses [131]. However, their inability to differentiate between human and non-human fecal sources limits their effectiveness in identifying contamination origins [137], [138]. To address this limitation, HF183 was often used as a microbial marker indicative of human fecal pollution [139], [140]. This marker, which targets bacteria found in the human gut, has shown improved specificity compared to traditional FIB. However, its short half-life of 0.4-8 days can

limit its applicability [140]. Therefore, it is recommended that HF183 be used in conjunction with other microbial indicators to obtain a more comprehensive assessment of the fecal pollution profile.

The case studies explored various chemical indicators to identify their potential as fecal pollution markers. Acetaminophen, a widely used pharmaceutical, has emerged as a promising fecal marker due to its frequent occurrence in wastewater effluents [33], [139], [141]. Caffeine, the most commonly analyzed chemical indicator in the case studies, has been used as an indication of recent untreated wastewater discharges due to its frequent consumption within the population and rapid biodegradation in the environment [106], [131]. Carbamazepine has also been identified as a potential indicator of human fecal contamination in aquatic environments since its persistence can indicate the presence of older untreated sewage [140], [142].

To overcome the limitations of individual indicators, a combined approach has been recommended in the literature [106], [129], [143]. The integration of microbial and CST methods provides a more comprehensive understanding of the fecal contamination sources, timings, and extent of pollution present at study sites, enhancing the accuracy and reliability of the analysis. While the ideal set of indicators varies based on site-specific conditions, the consensus within the literature is that an integrated approach offers enhanced source identification, and differentiation between recent and historical contamination, resulting in improved decision-making for water quality management [28], [106], [136]. Table 2.4 summarizes the 28 case studies that have employed a combined MST and CST approach to identify human fecal contamination in diverse aquatic environments.

Table 2.4 *Studies which have employed a combined microbial and chemical source tracking (CST) approach. CSO: Combined sewer overflow. FIB: Fecal indicator bacteria. PPCPs: Pharmaceuticals and personal care products. WWTP: Wastewater treatment plant.*

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
Toronto, ON, Canada	River	*B. dorei* *B. thetaiotaomicron* Bac32 *Bif. Adolescentis* CF128 *Catellicoccus marimammalium* *Clostridium perfringens* CowM2 DG37 *E. coli* (culturable and Amp^r [Ampicillin-resistant]) *Eu. rectale* *F. prausnitzii* GenBactF3 Gull2/tc HF183 *P. acnes* *R. bromii* *S. epidermidis* *Turicibacter sanguinis* qGull4	Caffeine Carbamazepine Codeine Cotinine Acetaminophen Acesulfame	Sewage contamination is present. Potential sources include WWTPs, sewage cross-connections in stormwater systems, or leaking septic systems. Caution should be used when using Amp^r *E. coli* as an indicator. The presence of caffeine, cotinine, and acetaminophen indicates recent raw sewage contamination, while carbamazepine is a possible indicator of older sewage contamination. CST markers did not always correlate with human microbial markers.	[28]
AZ, CO, GA, IA, KS, MN, NV, NJ, NY, SD, USA	WWTP discharges	*E. coli* Enterococci	110 chemical indicators tested Most common chemical indicators included: 1,4-dichlorobenzene 1,7-dimethylxanthine 3,4-dichlorophenyl isocyanate 4-nonylphenol diethoxylate 4-nonylphenol monoethoxylate 4-octylphenol diethoxylate 5-methyl-1H-benzotriazle Acetaminophen Benzophenone Bisphenol A (BPA) Caffeine Carbamazepine Cholesterol Codeine Coprostanol	The concentrations of most chemical indicators and the number of chemical indicators detected near the WWTPs decreased downstream, suggesting that in-stream degradation (dilution, biodegradation, sorption) occurs as they are transported downstream. Fire retardants and fecal and plant sterols were detected the most frequently, while the frequency of detection varied for the chemical groups (i.e., cotinine vs ibuprofen within the non-prescription drug group). The ratio of coprostanol-to-cholesterol in WWTP effluents and downstream indicates human waste contamination.	[131]

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
AZ, CO, GA, IA, KS, MN, NV, NJ, NY, SD, USA	WWTP discharges	Listed on previous page	Cotinine Dehydronifedipine Diazinon Diltiazem Diphenhydramine Ethanol,2-butoxy-,phosphate Ethyl citrate Galaxolide (HHCB) N,N-diethyl-m-toluamide (DEET) Pentachlorophenol Phenol Sitosterol Sulfamethoxazole Tonalide (AHTN) Tri(2-chloroethyl) phosphate (TCEP) Tri(dichlorisopropyl) phosphate Tributyl phosphate Triclosan Trimethoprim Triphenyl phosphate	Since *E. coli* and enterococci are not human-specific microbial indicators, they are not reliable when attempting to identify the presence of human fecal contamination. Of the 35 chemical indicators most frequently detected, five emerged as potentially effective tracers: two pharmaceuticals (carbamazepine, diphenhydramine), coprostanol, OBs, and caffeine.	[131]
GA, KS, MI, NC, NJ, OH, SD, TN, TX, VA, USA	Rivers, lakes, springs	*E. coli* Enterococci Fecal coliforms	1,7-dimethylxanthine 4-tert-octylphenol 5-methyl-1H-benzotriazole Benzophenone Beta-sitosterol Bromacil Caffeine Carbamazepine Chlorpyrifos Cholesterol Ciprofloxacin Codeine Coprostanol Cotinine Dehydronifedipine Diazinon Diethoxynonylphenol-total Diethoxyoctylphenol-total Diltiazem Enrofloxacin Erythromycin H_2O Ethanol 2-butoxy-phosphate Ethyl citrate	Fecal pollution was indicated by gene-based and/or chemical markers rather than FIB. Strong suggestion to utilize multiple fecal source indicators with variable persistence/fate to strategize decisions specific to the site. Environmental persistence and conditions during the sampling period may influence FIB concentration and the occurrence of certain indicators, resulting in inconsistencies within the data. Select chemical and gene-based indicators were present when the FIB standards were met.	[129]

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
GA, KS, MI, NC, NJ, OH, SD, TN, TX, VA, USA	Rivers, lakes, springs	Listed on previous page	Fluoxetine Galaxolide (HHCB) Indole Metalaxyl Methyl salicylate Metolachlor N,N,-diethyl-toluamide (DEET) Para-cresol Prometon Sarafloxacin Skatol Tetrachloroethylene Tonalide (AHTN) Tri(2-chloroethyl) phosphate (TCEP) Tri(dichlorisopropyl) phosphate Tributyl phosphate Triclosan Trimethoprim	Listed on previous page	[129]
Justiçou, Pen an Traon and La Fresnaye catchments, France	Headwater and coastal catchments	*E. coli* Enterococci HF183 Pig2Bac Rum2Bac	Bstanol Hstanol Pstanol	The source(s) of fecal contamination were identified for 83% of all samples when the combination of MST and CST markers was used. Concentrations of *E. coli* were impacted by rainfall and catchment size. The small urban catchment showed no modification of *E. coli* concentrations due to a constant source, while the larger catchment showed an increase in *E. coli* concentrations when rain intensity increased. Rainfall events led to higher detection rates of bovine markers across all catchments.	[136]
Brisbane, Sydney, and Melbourne, Australia	Catchments	Adenovirus *E. coli* *Enterococcus* HF183 *nif*H Polyomavirus	Acesulfame Caffeine Paracetamol Salicylic acid	Sewage-associated markers were found at all sites. HF183 and acesulfame were detected in 96% of stormwater samples, while caffeine was detected in 91% of samples. The results from the combined MST and CST approach demonstrated that human contamination was the main source of contamination in the urban catchments. The presence of chemical indicators at trace levels (μg/L) is an indication of a recent raw sewage release.	[106]

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
Kleinmond, South Africa	Rainwater harvesting tanks and gutter systems	Adenovirus HF183	Acetaminophen Caffeine Salicylic acid	HF183 and the human adenovirus were detected in samples from rainwater harvesting tanks and gutter systems, suggesting fecal contamination of the rainwater. Caffeine, salicylic acid and acetaminophen were detected in the samples, indicating human contamination. HF183 had a concurrence of 57.5% and caffeine had a concurrence of 82.5%.	[139]
Avon River, Christchurch, New Zealand	River	Atypical colonies/total coliforms (AC/TC) B. adol *Campylobacter* *Clostridium* *Cryptosporidium* Dog *E. coli* GenBac3 *Giardia* HumBac HumM3 Phage Wildfowl	Avian-associated steroid ratios: P1, Av1, Av2 FWAs General steroid ratios: F1, F2 Herbivore: R1 Human-associated steroid ratios: H1, H2, H3, H4, H5, H6	A combination of microbial and steroid tests indicated the presence of human sewage contamination. While PCR tests are rapid and less labour-intensive, steroid markers provide long-term storage stability as well as the ability to distinguish human from non-human sources. Steroid markers can detect historical human sewage contamination in sediments, including when the discharge events have ceased.	[137]
Ås and Ski, Norway	Rural creek and urban stream	Anthropogenic DNA marker (Anthropog.) *E. coli* Zoogenic DNA marker (Zoog.)	2-hydroxyibuprofen 4-hydroxydiclofenac Atenolol Bisphenol A (BPA) Caffeine Carbamazepine Carboxyibuprofen Chloramphenicol Diclofenac Erythromycin Furosemide Gabapentin Hydrochlorothiazide Ibuprofen Iohexol Ketoprofen Metoprolol	The urban stream contained human fecal contamination while the rural stream was affected by animal and environmental (i.e., non-human) sources. PPCPs were more prevalent in the urban stream and correlated strongly with human markers. Eutrophication-causing nutrients (i.e., total nitrogen and total phosphorus) were linked to fecal pollution sources (human sources for the urban stream and animal/environmental sources for the rural stream).	[138]

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
Ås and Ski, Norway	Rural creek and urban stream	Listed on previous page	Naproxen O-desmethylnaproxen Paracetamol Saccharin Tramadol Venlafaxine	Listed on previous page	[138]
Greater Montréal area, QC, Canada	Creek watersheds	*E. coli* Fecal coliforms HF183 Human mitochondrial DNA (Hmt)	Acetaminophen Carbamazepine Caffeine Theophylline	Human fecal contamination was detected in all samples, indicating the presence of sanitary sewer cross-connections within the watersheds sampled. Fecal coliforms and alternative markers (i.e., HF183, Hmt, caffeine, theophylline, and acetaminophen) showed significant correlations. There were significant correlations between *E. coli,* caffeine and theophylline. Due to the limitations of the *E. coli* threshold, an index that pinpoints cross-connections and ranks drainage basins for remediation is a more reliable method to determine sites that require rehabilitation.	[142]
Canada	Engineered urban canal	*E. coli* Fecal coliforms	Acetaminophen Carbamazepine Caffeine Theophylline	Due to its low temporal variability, *E. coli* is the best indicator of the exfiltration (release of groundwater from the channel) of wastewater and fecal discharges from sewers. Carbamazepine, caffeine, and theophylline degrade and can be diluted in surface water, therefore, they are not ideal indicators of exfiltration. High concentrations of acetaminophen corresponded with concentrations of *E. coli* in water samples, and acetaminophen can be an indicator of past discharges in sediments.	[33]

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
Lee County, FL, USA	Drainage ditches, canals, creeks	*E. coli* Enterococci HF183	Acetaminophen Carbamazepine Ibuprofen Naproxen Sucralose	Septic systems located adjacent to waterbodies can contaminate surface waters. The majority of the fecal bacteria originated from anthropogenic waste since HF183 was positively correlated with *E. coli* and enterococci. Sucralose detections indicated that contamination from human waste was present at the sites. Detections of pharmaceutical pain relievers (i.e., acetaminophen, ibuprofen, naproxen) further indicated the presence of human contamination. Increased infections of *V. vulnificus,* which has been associated with fecal bacteria, can become elevated when human wastewater is present in warm coastal waters, particularly following coastal storms and heavy runoff.	[140]
Greater Montréal area, QC, Canada	Combined sewer overflow (CSO) outfalls, receiving river, sewer system and WWTP effluent	*E. coli*	Acetaminophen Carbamazepine Caffeine Theophylline	Differences in *E. coli* and chemical indicator concentrations varied between sampling sites due to land use and population. Concentrations of carbamazepine, caffeine, and theophylline at the WWTP inlet fluctuated daily as a result of human activity. The concentration of acetaminophen remained relatively constant, indicating that it may be a suitable WWTP tracer for raw sewage in CSOs. Seasonal variabilities were observed; however, the study did not observe high dilution as a result of snowmelt. Dilution of stormwater caused by CSOs was observed.	[144]
Greater Montréal area, QC, Canada	Urban catchment	*E. coli*	Carbamazepine	The simulation model showed that sewage is the main source of *E. coli* contamination in urban water supplies. Carbamazepine is a stable tracer of sewage contamination since it was not present in stormwater.	[145]

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
Paranaguá and Guaratuba bays, Brazil	Subtropical estuaries	*E. coli* Enterococci	Coprostanol Linear alkylbenzenes (LABs)	No relationships were found between the microbial and chemical indicators when linear models (statistical) were applied. However, relationships between microbial indicators and coprostanol were observed when logistic regression was applied. It was hypothesized that temperature can affect the relationships between microbial indicators and coprostanol due to varying threshold values in different climates. Increased concentrations of enterococci were observed in the winter, and low concentrations were observed in the summer.	[146]
Milwaukee, WI, USA	Streams	Adenovirus group A Adenovirus groups A-F Adenovirus groups C, D, F *Campylobacter jejuni* *Cryptosporidium spp.* *Enteropathogenic E. coli* (eae gene) Enterovirus Hepatitis A virus Human *Bacteroides* Human *Lachnospiraceae* 3 Human Polyomavirus Norovirus genogroup I Norovirus genogroup II Pepper mild mottle virus Rotavirus group A (NSP3 gene) Rotavirus group A (VP1 gene) Rotavirus group C Salmonella (invA gene) *Salmonella* (ttr gene) Shiga toxin 1-producing bacteria (stx1 gene) Shiga toxin 2-producing bacteria (stx2 gene) sHSV	106 chemical indicators tested within the following categories: Analgesic-Anti-inflammatory: 15 Anesthetic: 1 Antacid: 4 Anthelmintic: 1 Anti-parkinson: 1 Antiallergen: 1 Antianxiety: 1 Antibiotic: 5 Anticonvulsant: 2 Antidepressant-Neurochemical Modulation: 17 Antidepressant: 3 Antidiarrheal: 1 Antifungal: 2 Antihistamine: 3 Antihistamine: 1 Antihyperglycemic: 1 Antihypertensive: 1 Antimalarial: 1 Antitussive: 1 Antiviral: 7 Anxiolytic, sedative: 1 Asthma Relief: 4 Cardiovascular Care: 15 Chemotherapeutic: 3 Chronic Condition: 1 Corticosteroid: 2 Degradate: 2 Estrogen inhibitor: 1	Many of the markers were not adequate as sewage contamination markers due to their lack of sensitivity. Chemical indicators had increased temporal variability compared to microbial indicators. Chemical indicators correlated with each other and were more consistent when estimating sewage pollution compared to microbial indicators. Multiple lines of evidence in monitoring and research are recommended when using indicators to determine the presence of human fecal contamination in water systems. Five human sewage indicators were detected consistently: pepper mild mottle virus, human *Bacteroides*, human *Lachnospiracaea*, acetaminophen, and metformin.	[147]

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
Milwaukee, WI, USA	Streams	Listed on previous page	Estrogen: 1 Opioid: 2 Stimulant: 5	Listed on previous page	[147]
Germany, Austria, Slovakia, Hungary, Serbia, Croatia, Romania, Bulgaria, Republic of Moldova, Ukraine	River	AllBac BacHum BacR *Cl. perfringens* *E. coli* Enterococci HF183II Pig2Bac	49 chemical indicators tested including: 4-acetylaminoantipyrine 4-formylaminoantipyrine Artificial sweeteners Benzotriazoles Betablockers Caffeine Carbamazepine Clofibric acid Cytostatic drugs and other pharmaceuticals Iodinated X-ray contrast media Lipid-lowering drugs Nonsteroidal anti-inflammatory drugs Salicyclic acid	High concentrations of *E. coli* were observed due to the influence of WWTPs. Certain *E. coli* trends cannot be explained (i.e., lower concentrations after the merge of tributaries, massive increases). Underestimation and processing errors may have impacted the enterococci data. HF183II, BacHum and *E. coli* were strongly correlated, indicating that the fecal pollution originated from human hosts. No strong correlations were observed between traditional fecal indicators, ruminant, and pig-associated markers. Regression analysis shows no detectable relationships between caffeine with genetic or bacterial markers, while a relationship between HF183II and carbamazepine was identified.	[148]
Helena Valley, MT, USA	Groundwater	Coliphage, male-specific Coliphage, somatic *E. coli* Enterococci Total coliforms	17-alpha-estradiol 17-alpha-ethynyl-estradiol 17-beta-estradiol Androstenedione Atrazine Bisphenol A (BPA) Caffeine Carbamazepine DEET Diazepam Diclofenac Diethylstilbestrol Dilantin Estriol Estrone Fluoxetine Gemfibrozil	Most frequently detected PPCPs include sulfamethoxazole, atrazine, carbamazepine, dilantin and diclofenac. It is hypothesized that atrazine is present in domestic wastewater since it correlates strongly with chloride and total dissolved solids (which are domestic wastewater discharge indicators). Poor correlations were observed between microbial indicators and PPCPs. Coliphages are not suitable indicators of fecal contamination in groundwater due to their lack of detection alongside PPCPs.	[149]

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
Helena Valley, MT, USA	Groundwater	Listed on previous page	Hydrocodone Ibuprofen Meprobamate Naproxen Oxybenzone Pentoxifylline Progesterone Sulfamethoxazole Testosterone Triclosan Trimethoprim	Listed on previous page	[149]
St. Lucia, Caribbean	River watershed	BacBovine BacGeneral BacHuman *E. coli* Total coliforms	Acetaminophen Caffeine Fluconazole Ibuprofen Metformin Sucralose	Increased total coliforms and *E. coli* concentrations from river grab samples did not coincide with concentrations detected, indicating additional pollution sources are contributing to the bay. The majority of microbial contamination in the river originated from non-human sources (i.e., ruminants and other warm-blooded animals). Areas of increased contamination were located near areas where pit latrines are used and open defecation is practiced, likely causing the increased detection of chemical indicators. Correlations between sucralose, caffeine, and BacHuman indicate human fecal contamination.	[150]
Gwelup and Jandakot borefields, Perth, Australia	Production bores, monitoring bores, surface water	*Clostridium perfringens* *E. coli* Enterococci F-RNA coliphages Faecal streptococci Presumptive sulphite-reducing clostridia (SRCs) Somatic coliphages	8 faecal sterols 4 hormones Caffeine	Microbial indicator concentrations were highest at the basins and lowest at the production bores. Aquifer filtration resulted in bacterial indicator concentrations decreasing by 4-5 orders of magnitude while cholesterol concentrations decreased by 3 orders of magnitude. Enterococci is a more sensitive indicator compared to *E. coli.* It was observed that the physicochemical indicators (i.e., pH, dissolved oxygen [DO], etc.) were not adequate surrogates for microbial indicators.	[151]

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
Mille Île's river, greater Montréal area, QC, Canada	River	BomtDNA *E. coli* HF183 HumtDNA PomtDNA Total coliforms	Acetaminophen Caffeine Carbamazepine Dihydro-carbamazepine Theophylline	One snowmelt/rainfall event resulted in elevated *E. coli* and HF183 concentrations, indicating that CSOs and/or the municipal water resource water recovery facility may have contributed to the elevated concentrations of raw sewage. It was observed that HF183 was the most sensitive marker to wet-weather and dry-weather overflows compared to all other indicators. *E. coli* and HF183 concentrations peaked during a contamination event, demonstrating that human sewage was the primary source of contamination. Caffeine, theophylline and carbamazepine peaked earlier than *E. coli* and other human markers, highlighting their potential as early CSO discharge indicators. Concentrations of caffeine and theophylline decreased during wet weather events while carbamazepine concentrations remained high, showing that carbamazepine is an indicator suitable for detecting human sewage contamination in wet weather conditions.	[141]
Grand River Watershed, ON, Canada	River watershed	BacBov BacGen BacHum *E. coli*	Acesulfame-K Caffeine Carbamazepine Gemfibrozil Ibuprofen Naproxen Sulfamethoxazole	Multiple anthropogenic tracers were found at the downstream sites throughout the sampling period, strongly indicating the presence of domestic wastewater impacts. Elevated concentrations of BacHum and BacBov were found close to areas used for agricultural purposes. The highest concentrations of pharmaceuticals were found at sites where there was an elevated abundance of FIB and human markers. It was hypothesized that low concentrations of organic compounds were due to groundwater removal or low consumption rates in communities near the sample sites.	[152]

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
Santa Barbara, CA, USA	Storm drains and creeks	*E. coli* Enterococci *Enterococcus spp.* HF183 *Methanobrevibacter smithii* Total coliforms	Caffeine Cotinine	The first study to directly compare the accuracy of canine responses to laboratory-based methods for detecting human waste in creeks and storm drains. The dogs accurately detected the presence of human waste contamination (70% for Dog 1 and 100% for Dog 2), demonstrating that canine scent detection is a viable method for anthropogenic waste detection. No false negatives were observed; however, both dogs indicated that human waste was present in four samples where human waste was not present. Canine responses were associated with certain assays (HF183 for Dog 1, HF183 and caffeine for Dog 2).	[153]
Gryteland stream, Norway	Stream (also known as a catchment)	Coliforms *E. coli* Human *Bacteroidales* Non-human *Bacteroidales*	46 chemical indicators tested Most common chemical indicators included: 2-hydroxy-ibuprofen Bisphenol A (BPA) Caffeine Carboxy-ibuprofen Gabapentin Ibuprofen Paracetamol Saccharin	The first Norwegian study to use MST to track pollution sources in water runoff collected in a natural wetland system. *Bacteroidales* DNA profiling methods were effective when distinguishing between human and non-human pollution sources. A strong positive correlation was seen between PPCPs and human *Bacteroidales,* while PPCPs were not correlated with high concentrations of *E. coli.* The most frequently detected PPCP was gabapentin. The nature-based treatment system was seen to remove *E. coli,* while PPCPs were unaffected.	[154]
Toronto, ON, Canada	River watershed	*E. coli* GooseMt Gull4 HF183 HuMt	Acesulfame K Acetaminophen Caffeine Carbamazepine Codeine Cotinine Paraxanthine Sucralose Theophylline	Microbial indicators were observed to be useful tools when interpreting elevated concentrations of *E. coli* in samples. HF183 was detected in all samples from the Don River and associated outfalls. The presence of sewage contamination was indicated by the detection of HF183 with chemical markers (i.e., caffeine, acesulfame).	[155]

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
Toronto, ON, Canada	River watershed	Listed on previous page	Listed on previous page	Increases in concentrations of *E. coli* and human DNA markers were observed during wet weather events. During dry weather flows, HF183, caffeine, acesulfame, and carbamazepine were detected at harbour CSO outfalls, which may be the result of CSO cross-connections.	[155]
Greater Montréal area, QC, Canada	Storm outfalls and storm sewer sub-catchments	*E. coli* GenBac HF183 Total coliforms	Acetaminophen Caffeine Carbamazepine Theophylline	High detections of HF183, caffeine and theophylline across all samples suggest the infiltration of sanitary sewer contamination into the storm pipe system. The utilization of traditional indicators to determine fecal pollution sources was found to be insufficient, suggesting that the combination of traditional indicators with human markers is necessary. It was determined that HF183, caffeine, and theophylline are indicative of human waste contamination from cross-connections, and carbamazepine indicates sewer exfiltration contamination. Positive correlations between HF183 and *E. coli* suggested that human fecal pollution was present.	[156]
Kauai, HI, USA	WWTP effluent and perennial streams	*E. coli* Enterococci Total coliforms	Sucralose	Average sucralose concentrations in the WWTP effluents were 39,167 ng/L (~9 ng/L/person), demonstrating its potential use as an indicator of water contaminated with wastewater. Enterococci concentrations exceeded the standards in all streams analyzed. It was observed that concentrations of FIB were not dependent on the presence of sucralose, however, using sucralose and enterococci can indicate the risk of human-related pathogens in recreational waters and drinking water.	[157]

Location	Aquatic environment	Microbial indicators	Chemical indicators	Key findings	Reference
Lake Ontario, Toronto, ON, Canada	Lake	*C. perfringens* *Cryptosporidium* *E. coli* *Giardia*	2-OH-carbamazepine Acesulfame Aspartame Caffeine Carbamazepine Sucralose	Significant correlations were observed between caffeine, carbamazepine, 2-OH-carbamazepine, *Giardia, E. coli,* and *C. perfringens* in raw sewage samples collected during wet weather events, indicating that these indicators are indicative of raw sewage discharges. Weak, yet significant, correlations were seen between flowrate, *Giardia, E. coli,* and *C. perfringens.* Carbamazepine and 2-OH-carbamazepine were positively correlated with flowrate, while total suspended solids and acesulfame were negatively correlated with flowrate. This indicates that acesulfame is diluted by high flows. It was determined that carbamazepine and 2-OH-carbamazepine are the most suitable indicators when determining the presence of human fecal contamination during wet weather conditions.	[158]
Ythan and Don rivers, Scotland	Catchments	*E. coli* Fecal coliforms Total coliforms	Acesulfame Caffeine Saccharin Sucralose Tryptophan	Positive correlations were seen between fecal coliforms, *E. coli*, total phosphorous, caffeine, and saccharine, demonstrating the potential for these indicators to be used as fecal discharge tracers. Positive correlations were also seen between caffeine, artificial sweeteners, coliforms, *E. coli*, and total phosphorous, suggesting that the chemical indicators can be used as reliable effluent discharge tracers. It was determined that a combination of physical and chemical tracing approaches is needed when evaluating the presence of septic tank effluents.	[143]

Chapter 3 – Methodology

3.1 Data Collection

The data used for this study was provided by the City of Calgary and spanned from 2018 to 2023 (see Tables 3.2 – 3.5). The dataset contained three main categories of parameters: 1) FIB (i.e., *E. coli*, total coliforms, enterococci) (see Table 3.3), 2) MST markers (see Table 3.4), and 3) micropollutants (i.e., chemical indicators) (see Table 3.5). The frequency of data collection as well as the number of data points provided by the City of Calgary for BR2, BR3, BR4, and PF are listed in Table 3.2.

Table 3.3 presents the detection limits for the FIB analyzed by the City of Calgary. Note that the detection limit for enterococci cannot be determined due to the nature of the enterococci qPCR assay (US EPA Method 1611). This method uses amplification to detect the DNA of enterococci bacteria (lsrRNA, 23S rRNA) in water samples, which is reported as calibrator cell equivalents (CCEs) rather than a quantitative measurement based on a standard curve [159]. The qPCR assay detects both viable and non-viable bacterial cells and is preferred over traditional culture methods due to its rapidity and specificity [159]. The usage of CCEs has been shown to accurately predict the relationship between enterococci and gastrointestinal illnesses, which indicate risks to human health, making it the preferred metric for assessing enterococci levels [160]. MST markers are shed by warm-blooded hosts, such as humans, ruminants, and birds. Table 3.4 lists the target genes and assay names used to identify fecal sources from the microbial markers provided in the dataset from the City of Calgary. Micropollutants monitored fell into the following broad classifications: 1) hormones/contraceptives, 2) industrial compounds, 3) pharmaceuticals, and 4) WWTP tracers (see Table 3.5).

3.2 Data Visualization

The Python programming language, which was executed within the Jupyter Notebook environment accessed through the Anaconda distribution, was used to process and generate visualizations of the dataset. Python provides flexibility and a broader spectrum of analytical capabilities, making it

ideal for data-driven research [161]. Jupyter Notebooks are interactive collaborative platforms that allow for the execution of code in modular cells and provide immediate visualization of results. Preprocessing of the dataset via Python included removing non-numerical values, omitting unnecessary columns, and converting the data to numerical values for analysis. The main Python libraries and modules utilized as well as their functions are summarized in Table 3.1. Sample Python codes used to analyze and visualize the data can be found in Appendix A.

Table 3.1 *Functions of Python libraries and modules used for data analysis.*

Python library/module	Version	Function	Reference
Pandas	1.5.3	Data analysis and manipulation	[162]
Seaborn	0.12.2	Generating statistical visualizations	[163]
Matplotlib	3.7.1	Plotting static and interactive graphics	[164]
OS	3.10.9	Accessing files within the local operating system	[165]

Table 3.2 *Site information and parameters monitored where microbial and chemical indicators were sampled. Data points considered are limited to sample dates where both microbial and micropollutant sampling took place. a) CowM3 and Pig2Bac sampling ended in 2021/2022. The remaining MST markers were sampled from 2018 – 2023. b) Sampling of certain micropollutants ended in 2021/2022 (see section 4.1). c) Micropollutant sampling did not take place during 2023 (see section 4.1). BR: Bow River. CCE: Calibrator cell equivalents. FIB: Fecal indicator bacteria. MPN: Most probable number. MST: Microbial source tracking. PF: Policeman's Flats.*

Site Designation	Site Name	Latitude	Longitude	Parameters	Unit	Number of data points from 2018 to 2023
BR2	Bearspaw Source Water	51.09971	-114.228	Enterococci, FIB	CCE/100 mL	49
				E. coli and total coliforms, FIB	MPN/100 mL	49
				MST markers (9 genetic markers, see Table 3.4)	Copies/100 mL	29-49[a]
				Chemical indicators (56 substances, see Table 3.5)	ng/L	28-49[b]
BR3	Bow River Cushing Bridge	51.03857	-114.013	Enterococci, FIB	CCE/100 mL	48
				E. coli and total coliforms, FIB	MPN/100 mL	48
				MST markers (9 genetic markers, see Table 3.4)	Copies/100 mL	39-48[a]
				Chemical indicators (56 substances, see Table 3.5)	ng/L	37-48[c]
BR4	Bow River Upstream of Highwood River	50.81972	-113.796	Enterococci, FIB	CCE/100 mL	46
				E. coli and total coliforms, FIB	MPN/100 mL	46
				MST markers (9 genetic markers, see Table 3.4)	Copies/100 mL	31-46[a]
				Chemical indicators (56 substances, see Table 3.5)	ng/L	31-46[b]
BR4 - PF	Policeman's Flats	50.841969	- 113.9511340	Enterococci, FIB	CCE/100 mL	8
				E. coli and total coliforms, FIB	MPN/100 mL	8
				MST markers (9 genetic markers, see Table 3.4)	Copies/100 mL	4-8[a]
				Chemical indicators (56 substances, see Table 3.5)	ng/L	2-8[b]

Table 3.3 *Detection limits of the fecal indicator bacteria (FIB) monitored by the City of Calgary. CCE: Calibrator cell equivalents. MPN: Most probable number. a) There is no reportable limit for Enterococcus CCE/100 mL. All values are reported.*

Component	Detection Limit	Units
E. coli	<1	MPN/100 mL
Enterococci[a]	-	CCE/100 mL
Total coliforms	<1	MPN/100 mL

Table 3.4 *Culture-dependent and library-independent microbial source tracking (MST) markers monitored by the City of Calgary. Quantitative polymerase chain reaction (qPCR) was used to conduct the assays listed. rRNA: Ribosomal RNA. Reporting limit: 1200 Copies/100 mL.*

Target host	Target gene	Assay	References
Canada goose	16S rRNA	CGO1 (or CG01)	[166], [167], [168]
Cow	Sialic acid-specific 9-*O*-acetylesterase secretory protein homolog	CowM3	[87], [169], [170]
Dog	Long chain fatty acid -CoA ligase	Dog3	[168], [171]
Human	16S rRNA	HF183	[87], [168]
	Cell surface-associated genes	HumM2	[87]
Muskrat	16S rRNA	MuBac	[85], [168]
Pig	16S rRNA	Pig2Bac (or Pig-2-Bac)	[87]
Ruminant	16S rRNA	Rum2Bac (or Rum-2-Bac)	[85], [168]
Seagull	16S rRNA	LeeSG (or LeeSg)	[168], [172]

Table 3.5 *List of micropollutants monitored by the City of Calgary (adapted from [31]). BPA: Bisphenol A. BPS: Bisphenol S. DEET: N,N-diethyl-meta-toluamide. PFOA: Perfluorooctanoic acid. PFOS: Perfluorooctanesulfonic acid. TCEP: Tris(2-carboxyethyl) phosphine. TCPP: Tris(chloropropyl) phosphate. WWTP: Wastewater treatment plant.*

Compound	Broad Classification	Subcategory	Reporting Limit (ng/L)
17α-Estradiol	Hormone/Contraceptive	Hormone	10
17α-Ethynylestradiol	Hormone/Contraceptive	Contraceptive	0.8
17β-Estradiol	Hormone/Contraceptive	Hormone	10
Androstenedione	Hormone/Contraceptive	Hormone	2
Equilenin	Hormone/Contraceptive	Hormone	5
Equilin	Hormone/Contraceptive	Hormone	5
Estriol	Hormone/Contraceptive	Hormone	5
Estrone	Hormone/Contraceptive	Hormone	5
Norethindrone	Hormone/Contraceptive	Contraceptive	5
Norgestimate	Hormone/Contraceptive	Contraceptive	5

Compound	Broad Classification	Subcategory	Reporting Limit (ng/L)
Progesterone	Hormone/Contraceptive	Hormone	2
Testosterone	Hormone/Contraceptive	Hormone	2
4-n-Nonylphenol	Industrial Compounds	Industrial Surfactant	10
4-t-Octylphenol	Industrial Compounds	Industrial Surfactant	100
Benzyl butyl phthalate	Industrial Compounds	Plasticizer	20
BPA	Industrial Compounds	Plasticizer	50
BPS	Industrial Compounds	Plasticizer	5
Cotinine	Industrial Compounds	Cotinine	1
DEET	Industrial Compounds	Insect Repellent	5
Di-n-octyl phthalate	Industrial Compounds	Plasticizer	100
Di(2-ethylhexyl) phthalate	Industrial Compounds	Plasticizer	1500
Dibutyl phthalate	Industrial Compounds	Plasticizer	300
Diethyl phthalate	Industrial Compounds	Plasticizer	50
Dimethyl phthalate	Industrial Compounds	Insect Repellent	100
PFOA	Industrial Compounds	Industrial Surfactant	10
PFOS	Industrial Compounds	Industrial Surfactant	20
TCEP	Industrial Compounds	Flame Retardant	5
TCPP	Industrial Compounds	Flame Retardant	50
Triclosan	Industrial Compounds	Antibacterial Agent	10
Acetaminophen	Pharmaceutical	Analgesic	5
Atenolol	Pharmaceutical	Cardiovascular Drugs	5
Carbamazepine	Pharmaceutical	Antiepileptic	1
Citalopram	Pharmaceutical	Antidepressant	20
Clarithromycin	Pharmaceutical	Antibiotic	5
Codeine	Pharmaceutical	Opiate	10
Diclofenac	Pharmaceutical	Analgesic	5
Erythromycin	Pharmaceutical	Antibiotic	2
Fluoxetine	Pharmaceutical	Antidepressant	5
Gemfibrozil	Pharmaceutical	Cardiovascular Drugs	5
Ibuprofen	Pharmaceutical	Analgesic	10
Metformin	Pharmaceutical	Antidiabetic	5
Naproxen	Pharmaceutical	Analgesic	5
Nifedipine	Pharmaceutical	Cardiovascular Drugs	5
Norfluoxetine	Pharmaceutical	Antidepressant (metabolite)	5

Compound	Broad Classification	Subcategory	Reporting Limit (ng/L)
O-Desmethyl-venlafaxine	Pharmaceutical	Antidepressant (metabolite)	1
Pantoprazole	Pharmaceutical	Acid Reducer	5
Primidone	Pharmaceutical	Antiepileptic	1
Propranolol	Pharmaceutical	Cardiovascular Drugs	5
Salbutamol	Pharmaceutical	Others	5
Sulfamethoxazole	Pharmaceutical	Antibiotic	1
Trimethoprim	Pharmaceutical	Antibiotic	1
Venlafaxine	Pharmaceutical	Antidepressant	1
Zopiclone	Pharmaceutical	Others	10
Acesulfame	WWTP Tracer	Artificial Sweetener	50
Caffeine	WWTP Tracer	Stimulant	5
Sucralose	WWTP Tracer	Artificial Sweetener	10

3.3. General Analysis Workflow

Data filtering and clean-up were completed prior to data processing in Python. First, values that were below the reporting limits (RLs) were excluded from the analysis. Although substitution of censored data (i.e., data < RLs) with its RL or ½ RL is common practice, this approach produces poor estimates [173]. Maximum likelihood (MLE) and Kaplan-Meier (KM) techniques are example alternatives employed for substitution [174], [175], but this is outside the scope of this thesis. Hence, it was deemed more appropriate to exclude the data rather than to substitute as information from the original dataset may be modified through substitution.

Parameters with fewer than seven numerical detections were excluded from the analysis, and values below the RLs or detection limits were converted to “Not a Number” (NaN) values in Python. Although there is no consensus related to the sample size required for correlation analysis in water quality, a minimum of $n = 5$ for Pearson correlation analysis and $n \geq 7$ for regression analysis (straight line model only) have been recommended [176]. Sample sizes for the three sites across all datasets are listed in Table B.1 of Appendix B.

Pearson and Spearman correlation coefficients are the most frequently employed in water quality correlation analysis (see section 3.3.2), with Pearson requiring data to be normal for it to be a valid approach. Furthermore, preliminary data processing suggested that the datasets were mostly non-normal and subsequent data transformation was completed and checked for normality: 1) $\log_{10}$, 2) $1/x$, 3) $\sqrt[2]{x}$, and 4) x^2. The normality of the dataset was analyzed following each transformation method using the Shapiro-Wilk, D'Agostino's K^2, and Jarque Bera normality tests (see section 3.3.1). The data transformation method which yielded the highest number of normalities was selected. The same approach was applied when analyzing seasonal data for each site. Once the data transformation methods were selected, the appropriate correlation methods and clustering linkage methods were identified. A visualization of the steps involved in this workflow is shown in Figure 3.1. Further information regarding normality, correlation and clustering linkage method selection are described in the following sections.

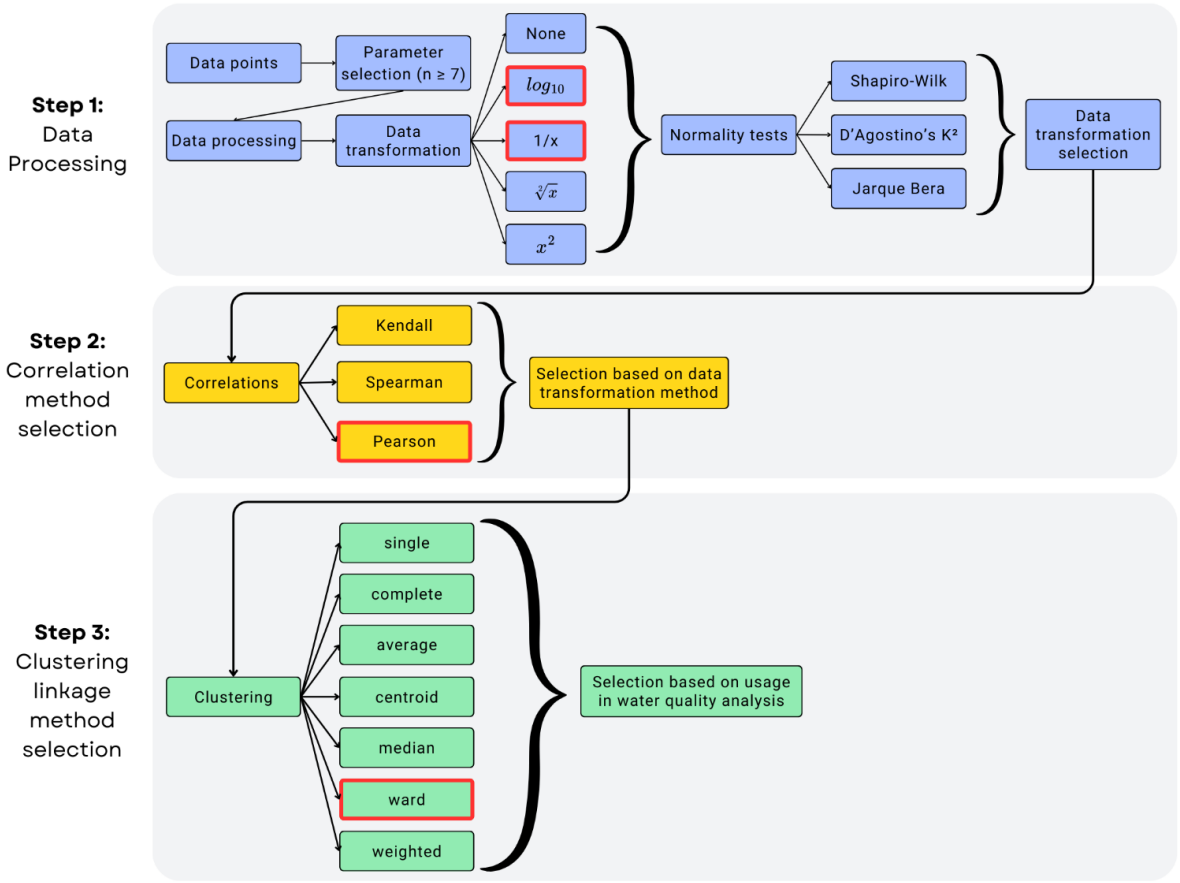


Figure 3.1 *Data analysis flow chart. Boxes highlighted in red indicate the final approaches used.*

Preliminary visualizations were used to gain an initial understanding of the dataset provided. Presence/absence plots were used to identify the detection of parameters over the sampling period. Temporal variations of representative parameters from the FIB, MST markers, and chemical indicators were observed through time series plots and their relationship with the flow of the BR was analyzed. Box plots were utilized to visualize data distributions at each site. Individual box plots, representing the interquartile range (IQR), median, and minimum and maximum data points (shown within the whiskers), were used to assess data normality and skewness. The IQR is illustrated by the rectangles in each plot where smaller IQRs indicate consistent data and larger IQRs show that the data points have a larger spread [177], [178], [179]. For normally distributed data, the median is located in the centre of the box plot and the whiskers are equal in length. Right-skewed distributions exhibit a median in the lower part of the box plot with a lower whisker shorter than the upper whisker. A left-skewed distribution is characterized by a median in the upper part of the box plot with a longer lower whisker and a shorter upper whisker [180], [181]. Analysis of the dataset as a whole and seasonally was conducted for each site to determine overall and seasonal trends and accurately interpret site-specific and seasonal variations. Using the start dates of the seasons as defined by the National Research Council of Canada (i.e., Spring: March 20, Summer: June 20/21, Fall: September 22/23, Winter: December 21/22) [182], the overall datasets for each site were manually sorted into the seasons using Microsoft Excel.

3.3.1 Normality Tests

The equations used to calculate normality for the Shapiro Wilk, D'Agostino's K^2 and Jarque Bera tests are outlined in Table 3.6, however, these calculations were performed using Python (see Appendix A). Using multiple normality tests is recommended in the literature, as relying on the results of a single test may not provide a comprehensive assessment of the data [183], [184], [185]. The Shapiro-Wilk test is one of the most common normality tests and is ideal for small to moderate sample sizes ($n < 50$) [184], [186]. Using this approach, a dataset's normality is determined by comparing the ordered sample values to

the expected values from a normal distribution. A W statistic value of 1 indicates that the data is likely normal, while a value closer to 0 rejects the normality null hypothesis [193]. The D'Agostino's K^2 and Jarque-Bera tests determine departures from normality based on the skewness and kurtosis of a dataset [183], [188], [189], [190], [191]. D'Agostino's K^2 test can be used for small to moderate sample sizes ($n < 20$), and the Jarque-Bera test is more accurate with larger sample sizes [189], [190]. Skewness refers to the extent to which the distribution is symmetrical, while kurtosis measures the dispersion of the data from the normal distribution [179].

Table 3.6 *Shapiro-Wilk, D'Agostino's K^2 and Jarque-Bera normality test equations [192], [193].*

Normality test	Equation	Description of variables
Shapiro-Wilk	$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$ where $(a_1, \dots, a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}}$	x_i = ith order statistic $\bar{x} = (x_1 + \cdots + x_n/n$ = sample mean $m = (m_1, \dots, m_n)^T$ and $m_1, \dots, m_n$ = the expected order statistics V= covariance matrix of the order statistics sampled from the normal population
D'Agostino's K^2	$Z_1(g_1) = \delta ln\left(\frac{g_1}{\alpha\sqrt{\mu_2}} + \sqrt{\frac{{g_1}^2}{\alpha^2\mu_2} + 1}\right)$ where $W^2 = \sqrt{2\gamma_2 + 4} - 1$, $\delta = 1/\sqrt{lnW}$, $\alpha^2 = 2/(W^2 - 1)$	g_1 = transformation for sample skewness α and δ = constants $\mu_2 = \mu_2(g_1)$ = variance of g_1 $\gamma_2 = \gamma_2(g_1)$ = kurtosis of g_1
Jarque-Bera	$JB = n[(b_1)/6 + (b_2 - 3)^2/24]$	n = sample size b_1 = skewness b_2 = kurtosis

The normality of the datasets for each site (BR2, BR3, BR4), both as a whole and seasonally, was assessed using the Shapiro-Wilk, D'Agostino's K^2 and Jarque-Bera tests following various data transformations. The SciPy "**stats**" library submodule (version 0.13.5) was utilized to access the necessary statistical functions. Each parameter was analyzed using the three normality tests, providing a statistic and corresponding *p*-value. Parameters with a *p*-value greater than 0.05 were deemed significant

[194], [195]. This process was repeated for each data transformation considered. Transformation methods resulting in the highest portion of parameters classified as “Normal” were selected. Sample Python code for the Shapiro-Wilk normality test can be found in Appendix A.

3.3.2 Correlation Method Selection

Correlation analysis identifies the relationship between variables [196], with positive correlations (i.e., +1) indicating that as one variable increases, the second variable tends to increase as well, while negative correlations (i.e., -1) show that as one variable increases, the second variable decreases [197]. The three correlation methods available when using the “**.corr”** function (version 1.5.3) in Python are 1) Pearson, 2) Spearman, and 3) Kendall. Pearson is employed when the data is normally distributed while Spearman and Kendall are used when the data is non-parametric, or not normally distributed [198].

A brief literature review suggested that Spearman and Pearson are commonly used correlation methods in water quality analysis [199], [200], [201], [202], [203]. The Pearson correlation method was selected for datasets for each site (overall and seasonally) that were considered to be normal following the appropriate data transformation. This tailored approach to the dataset was employed to enhance the reliability of the results generated by the correlation matrices.

While correlation strength criteria have been developed within the scientific community, there is no general consensus on a universally accepted standard. As illustrated in Table 3.7, a wide range of correlation strength interpretations for correlation strength exist [204], [205], [206], [207]. For the purposes of this thesis, the correlation coefficients are categorized as weak (0 to 0.3), low (0.3 to 0.5), moderate (0.5 to 0.7), strong (0.7 to 0.99), or perfect (1) based on the absolute value of the Pearson correlation coefficient (R^2). This will be used for any discussions related to correlation.

Table 3.7 *Correlation strength interpretations across various studies and what was employed for this thesis. R^2: Pearson correlation coefficient.*

Correlation strength interpretation	**$\|R^2\|$ coefficient**	**Reference**	**This study** Correlation strength interpretation	$\|R^2\|$ coefficient
Little or (very) weak	0 to 0.29	[204]	Weak	0 to 0.3
	0 to 0.3	[205]		
	0 to 0.3	[206]		
	0 to 0.2	[207]		
Low or weak	0.3 to 0.49	[204]	Low	0.3 to 0.5
	0.3 to 0.5	[206]		
	0.2 to 0.4	[207]		
Moderate	0.5 to 0.69	[204]	Moderate	0.5 to 0.7
	0.3 to 0.7	[205]		
	0.5 to 0.7	[206]		
	0.4 to 0.6	[207]		
High or strong	0.7 to 0.89	[204]	Strong	0.7 to 0.99
	0.7 to 1	[205]		
	0.7 to 1	[206]		
	0.6 to 0.8	[207]		
Very high/strong or perfect	0.9 to 1	[204]	Perfect	1
	1	[205]		
	0.8 to 1	[207]		

3.3.3 Clustering Linkage Method Selection

Hierarchical clustering, a type of unsupervised machine learning, was used to group similar data into clusters based on a specified similarity or distance metric [208]. This technique is unsupervised since it does not follow explicit instructions. Instead, it identifies patterns and relationships within the data based on a specified method (i.e., single, average, Ward) [209]. Dendrograms are one type of visualization generated using clustering analysis. Shorter cluster distances indicate that there are similarities between parameters, while larger cluster distances suggest dissimilarities between parameters [210]. Figure 3.2 demonstrates this concept, where parameters f and g are similar due to their shorter cluster distance, while parameters h and i are dissimilar due to their larger cluster distance [208], [211].

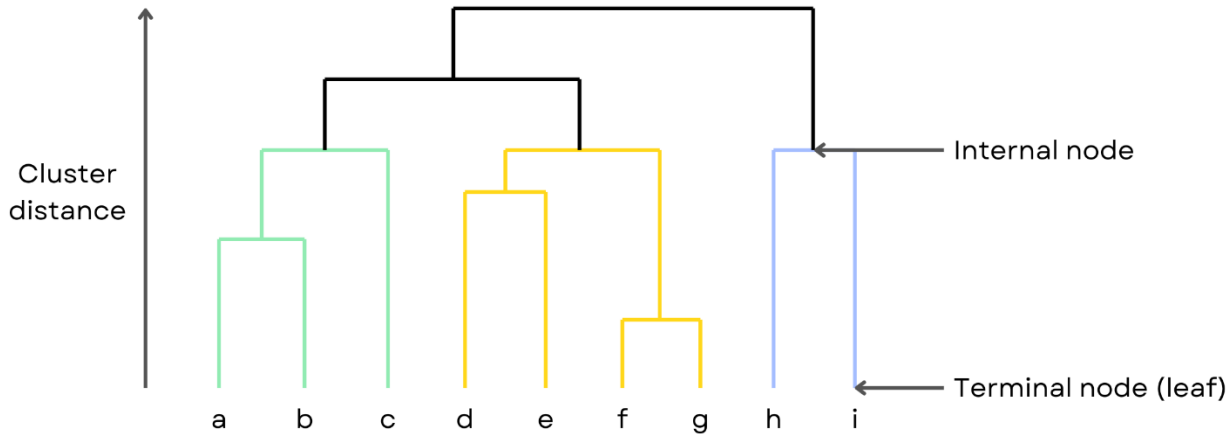


Figure 3.2 *Visual representation of hierarchical clustering of 9 parameters (a-i) forming three clusters (adapted from [212]).*

The determination of "short" and "large" cluster distances is data-dependent and requires the identification of a threshold distance in order to identify main clusters [213]. All clusters below the threshold distance will be considered similar, while those above the threshold are dissimilar. Such thresholds are determined through the subjective analysis of the dendrogram and selecting a point that reflects the desired level of granularity, aiding with the identification of patterns and relationships within the dataset.

There are two main categories of hierarchical clustering: 1) agglomerative and 2) divisive. Agglomerative hierarchical clustering considers each data point to be an individual cluster and groups the data into larger clusters, with the entire data set being the last cluster [214], [215]. This is opposite to divisive hierarchical clustering, which first assumes that the entire dataset is the first cluster, then proceeds to group the data into smaller clusters until each data point is considered to be individual clusters [215], [216]. This concept is illustrated in Figure 3.3 [215], [217], [218].

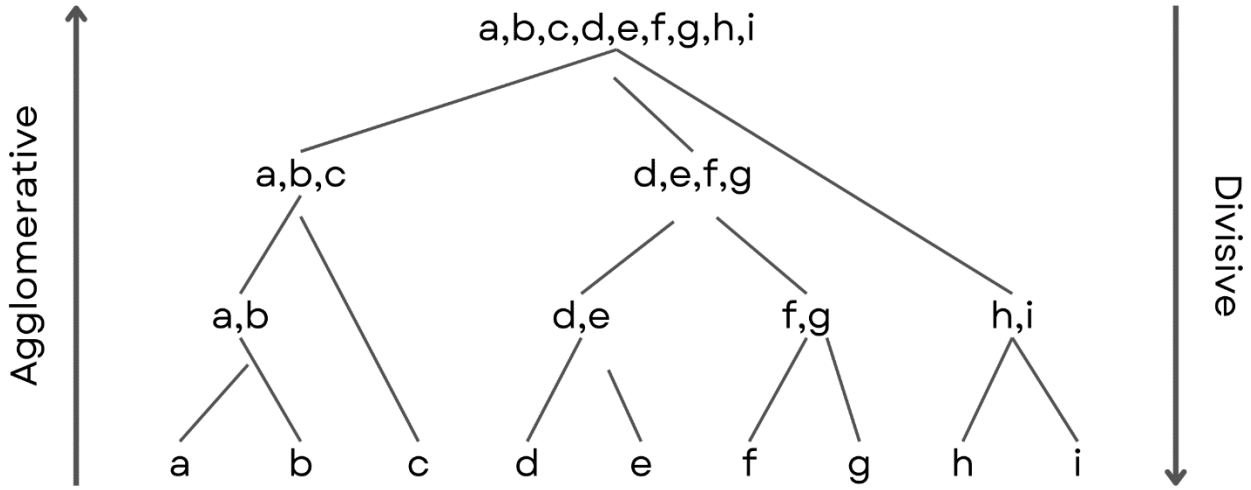
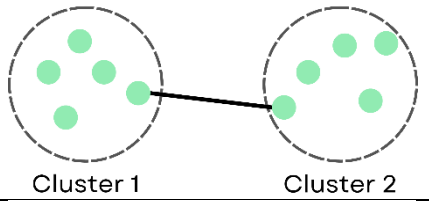
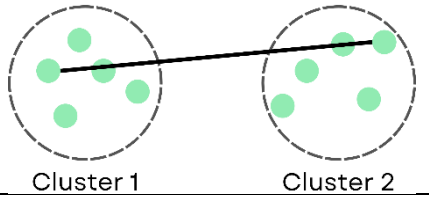
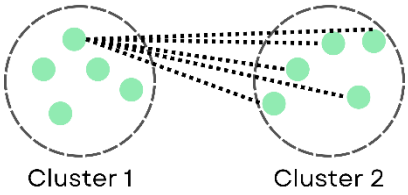
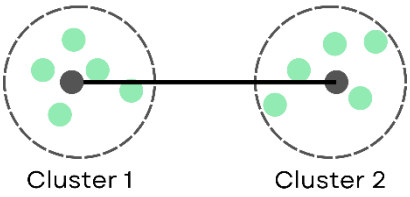
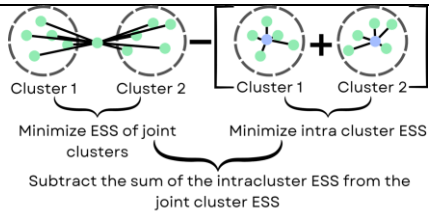


Figure 3.3 *Visual representation of agglomerative and divisive hierarchical clustering of 9 parameters (a-i) (adapted from [215], [217], [218]).*

Within the Python code, the "**clustermap**" function (version 0.12.2) was used to visualize the hierarchical sorting of a dataset by utilizing a specific linkage method. There are seven linkage methods available when using the "**clustermap**" function: 1) single, 2) complete, 3) average, 4) centroid, 5) median, 6) Ward, and 7) weighted. These 7 methods are the most commonly used agglomerative hierarchical methods and incorporate both graphical and geometric methodologies [215]. Descriptions, advantages, disadvantages and visual representations of the linkage methods are listed in Table 3.8 [215], [217], [219], [220].

Table 3.8 *Summary of the agglomerative hierarchical clustering linkage methods (adapted from [215], [217], [219], [220]).*

Linkage method	Description	Advantages	Disadvantages	Visual representation
Single	Measures the similarity between clusters by prioritizing regions where the clusters are closest	Can effectively handle nonelliptical and elongated clusters	Sensitive to noise and outliers	
Complete	Defined by the distance between the farthest members of neighbour clusters	Can produce compact clusters	Sensitive to noise and outliers	
Average (weighted and unweighted)	Average distance between clusters is taken as the average distances between all data point pairs	Performs well with ball-shaped clusters	Computationally expensive for large datasets	
Centroid	Measures similarity between clusters by comparing centroids and merging clusters based on centroid similarities	Performs well with compact data structures	Assumes Euclidean distance; sensitive to differences in cluster sizes	
Median	Uses Euclidean distance between weighted centroids of the clusters to measure similarities	Alleviates disadvantages of the centroid linkage method; performs well with elongated data structures	Not suitable for correlation coefficients	-
Ward	Computes the sum of squares error (ESS) when merging clusters	Effectively identifies dense spherical clusters amidst background noise	Determination of the minimum ESS is not guaranteed	

Cluster maps for all sites were generated using the linkage methods listed in Table 3.8. Preliminary analysis showed minimal differences among these methods, therefore, Ward linkage was selected due to its frequent use and effectiveness for water quality data, as observed in the literature. Studies by Trábert et al., Tiri et al., Ukpatu et al., and Guzman et al. have applied the Ward method when analyzing the water quality of the Danube River in Budapest, the Koudiat Medouar Watershed in Algeria, the Okoro River Estuary in Nigeria, and across the United States, respectively [221], [222], [223], [224]. Tiri et al. and Ukpatu et al. explicitly indicated Ward's ability to produce distinctive groups and detect multivariate patterns and relationships across seasons for various water quality parameters [222], [223].

Since hierarchical clustering is also used for exploratory data analysis, it is ideal for investigating the utility of microbial indicators and micropollutants of interest [212]. The outcomes of the correlation matrices and cluster maps will inform the decision as to which parameters should be included in the composite fingerprint for each site.

Chapter 4 – Results and Discussions

This chapter explores the qualitative and quantitative relationships between FIB, MST markers and chemical indicators at the BR2, BR3, and BR4 sites along the BR. These relationships were analyzed through the generation of presence/absence plots, time series plots, box plots, correlation matrices, and cluster maps. The chapter then discusses the applicability and effectiveness of combining microbial and chemical indicators to identify the presence of human sewage pollution in source waters and aquatic environments. The chapter concludes with a proposed composite fingerprint specific to the three sites.

4.1. Frequency of Detection – Presence/Absence

Presence/absence plots (Figures 4.1 – 4.3) summarize the number of detections (for parameters with $n \geq 7$) and non-detections for all target indicators during the 2018 – 2023 sampling period. The sequence of sites with the lowest to highest number of detections as per the criteria developed in this thesis is BR2 (10 parameters) < BR3 (16) < BR4 (33), which is consistent with the hypothesis that sites near treated effluent discharge points (which also services larger populations) are likely to have more frequent detections of microbial and chemical indicators. For instance, BR2 (source water) is downstream of small communities with a combined population of only ~35,650 [42], whereas BR4 is located approximately ~3.22 – 14.20 km downstream of Calgary's three WWTPs servicing over one million residents. Additionally, since these sites are located sequentially downstream, the most downstream site (BR4) is expected to accumulate more impacts from anthropogenic (urban) activities.

All three FIB (*E. coli,* enterococci, and total coliforms) were consistently detected at the three sites (i.e., 100% at all sites), with the exception of *E. coli* at BR2 which was only detected 44.90% over the sampling period. Overall, frequent FIB detections were expected as they are ubiquitous in the environment since they originate from fecal and non-fecal sources. However, BR2 is a source water site which is considered to be the least impacted than the other sites.

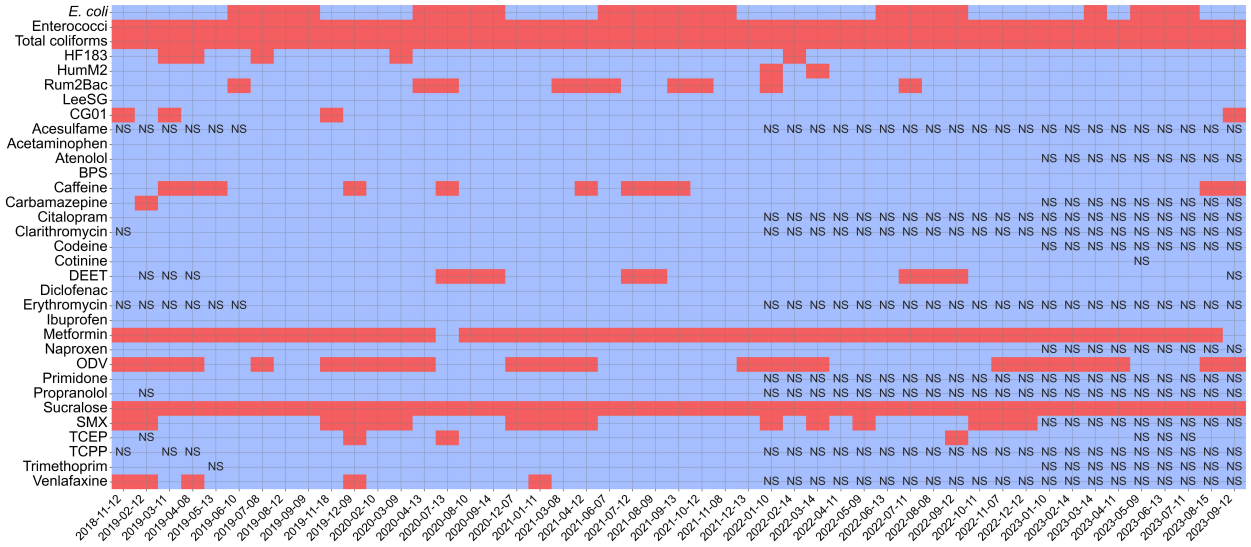


Figure 4.1 *Presence/absence plot for* ***BR2****. Red indicates presence and blue indicates that the parameter was not detected or was below the reporting limit. Sampling that includes all indicators began in October 2018. Indicators which met the numerical detection limit (n ≥ 7) for BR4 are shown here for comparison. Indicator sample sizes for BR2 can be found in Table B.1 of Appendix B. BPS: Bisphenol S. DEET: N,N-diethyl-meta-toluamide. NS: Not sampled. ODV: O-Desmethyl venlafaxine. SMX: Sulfamethoxazole. TCEP: Tris(2-carboxyethyl)phosphine. TCPP: Tris(chloropropyl) phosphate.*

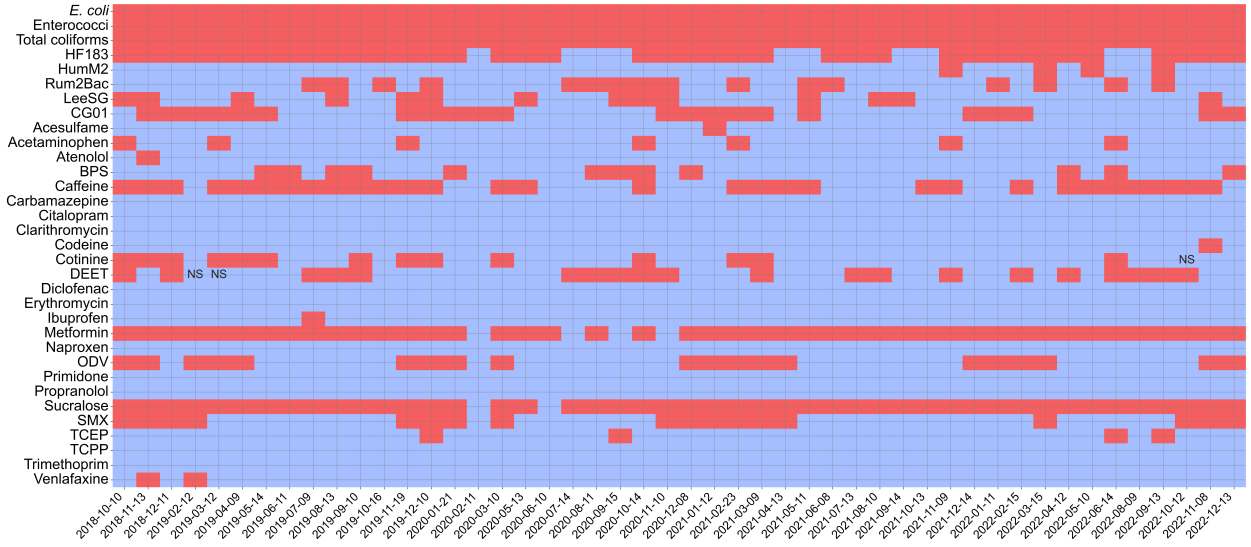


Figure 4.2 *Presence/absence* plot *for* ***BR3****. Red indicates presence and blue indicates that the parameter was not detected or was below the reporting limit. Sampling that includes all indicators began in October 2018. Indicators which met the numerical detection limit (n ≥ 7) for BR4 are shown here for comparison. Indicator sample sizes for BR3 can be found in Table B.1 of Appendix B. Sampling of chemical indicators was discontinued at BR3 in 2023. BPS: Bisphenol S. DEET: N,N-diethyl-meta-toluamide. NS: Not sampled. ODV: O-Desmethyl venlafaxine. SMX: Sulfamethoxazole. TCEP: Tris(2-carboxyethyl)phosphine. TCPP: Tris(chloropropyl) phosphate.*

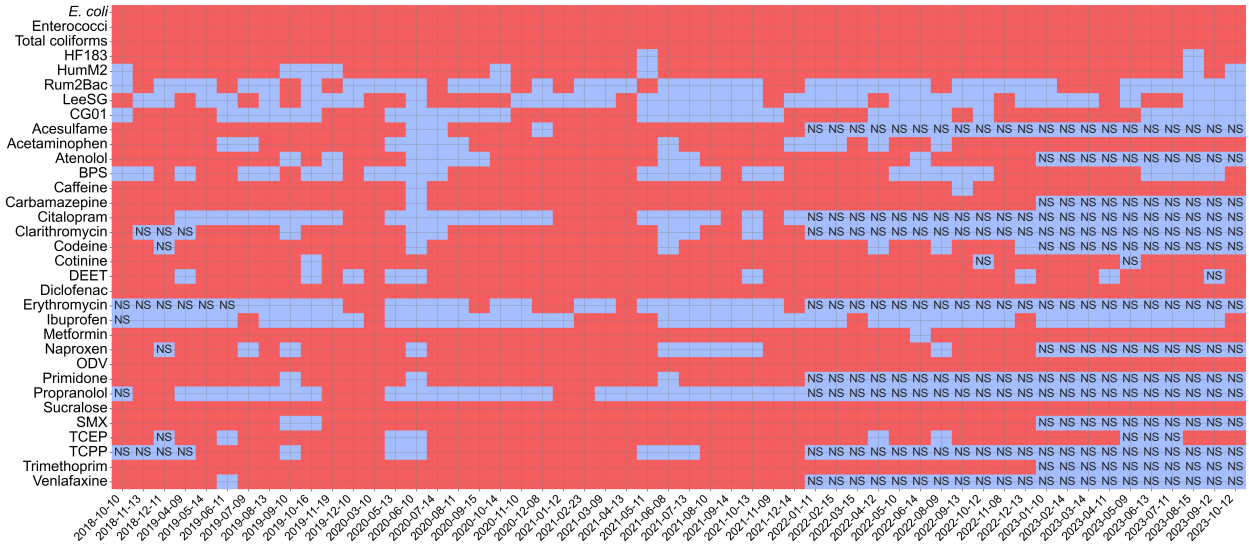


Figure 4.3 *Presence/absence plot for* ***BR4****. Red indicates presence and blue indicates that the parameter was not detected or was below the reporting limit. Sampling that includes all indicators began in October 2018. Indicator sample sizes for BR4 can be found in Table B.1 of Appendix B. Only a select set of chemical indicators were considered from 2022 to 2023 (acetaminophen, BPS, BPA, caffeine, cotinine, DEET, diclofenac, metformin, ODV, sucralose, TCEP). BPS: Bisphenol S. DEET: N,N-diethyl-meta-toluamide. NS: Not sampled. ODV: O-Desmethyl venlafaxine. SMX: Sulfamethoxazole. TCEP: Tris(2-carboxyethyl)phosphine. TCPP: Tris(chloropropyl) phosphate.*

As for monitoring of the MST markers, Rum2Bac (ruminant) was detected more than seven times at **all** three sites. Given the frequent sightings of white-tailed deer, mule deer, and occasional moose in Calgary, the widespread detection of Rum2Bac at all three sites is likely attributed to the presence of these ruminant species. A 2017–2018 report noted that the majority of wildlife activity within the city is dominated by deer species (78% of events), further supporting the frequent occurrence of Rum2Bac at the three sites [225].

CG01 (Canada goose) and LeeSG (seagull) were detected more than seven times at BR3 and BR4, but only ≤ four times at BR2. Both gulls and geese have been known to gather along the BR [226]. Their presence may be influenced by warmer waters from storm sewers, food availability (i.e., spring insect hatchings, fish remaining from snowmelt), and nesting sites [226], [227], [228]. The higher frequency of detection of CG01 and LeeSG at BR3 and BR4 compared to BR2 suggests a more localized distribution, potentially influenced by land use patterns and the location of wildlife habitats. For example,

the Inglewood Bird Sanctuary, located approximately 1 km downstream of BR3, is known to host the Canada goose and a variety of gull species, with some of these species passing through the area during the migration period [229].

At BR2, six chemical indicators (caffeine, DEET, metformin, O-Desmethyl-venlafaxine [ODV], sucralose, sulfamethoxazole [SMX]) were measured at least seven times throughout the monitoring period (Figure 4.1). Since caffeine degrades quickly in aquatic environments, its detection at low concentrations (as in the case of BR2) could suggest that it is a remnant of an older release or that a site is located in the far-field zone of a WWTP discharge [141], [155]. Regardless of its source, the sporadic detection of caffeine at BR2 indicates that the site is minimally impacted by anthropogenic activities, reinforcing BR2 as a high-quality source water for Calgary [230]. Additionally, the detection of other chemical indicators, including insect repellent (DEET), pharmaceuticals (metformin, ODV, SMX), and an artificial sweetener (sucralose) suggests that BR2 may be influenced by persistent anthropogenic activities. Hence, BR2 and areas further upstream will require regular monitoring as populations outside of Calgary continue to grow and additional discharges (i.e., treated effluent) into the BR upstream may be approved.

Although there is no effluent discharge near BR3, it receives inputs from NC (~1.3 km above BR3), a tributary that is impacted by seasonal sewage lagoon discharge (Crossfield, population: 3,599) [42] and runoff from the Calgary International Airport [231]. As a result, higher detection frequencies of microbial and chemical indicators were observed at BR3 (Figure 4.2). There are also 287 stormwater outfalls upstream of BR3 and cross-connections with the sanitary sewer network can lead to transfer of untreated sewage via leaks or infiltration. Furthermore, in addition to the FIB and other non-human markers (i.e., ruminants, geese, gull), HF183 (human marker) was detected at least seven times (Figure 4.2). The detection of HF183 at BR3 further suggests a potential source of untreated sewage contamination. More specifically, HF183 has been identified as a sensitive human fecal marker [57] and can be used in tandem with FIB to trace human fecal signals in the watershed network [232].

The detection of all six chemical indicators previously identified at BR2 were also present at BR3, with the addition of a pharmaceutical (acetaminophen), and industrial compounds (bisphenol S [BPS], cotinine). The increasing presence of pharmaceuticals and industrial compounds, often released into the environment through wastewater effluents and other anthropogenic sources, can pose risks to aquatic ecosystems; however, their human health impacts are not fully understood [233], [234], [235], [236]. While chemical indicators such as caffeine, carbamazepine and cotinine have been identified as potential indicators of human sewage (i.e., caffeine as an indication of recent untreated sewage and carbamazepine being an indication of older untreated sewage releases), the presence of BPS, DEET, metformin, ODV, SMX, and their association with human sewage is less established (see Table 2.4). The number of numerical detections increased for caffeine (11/49 to 30/48), DEET (8/49 to 20/48), and SMX (16/49 to 18/48), while numerical detections decreased for metformin (47/49 to 44/48), ODV (26/49 to 20/48), and sucralose (49/49 to 46/48) between BR2 and BR3. This result suggests that BR3 may experience a higher frequency of **more recent untreated** sewage releases, as indicated by the increased detections of caffeine, and the presence of chemical indicators which were absent from BR2. The increase in chemical indicators at BR3 is consistent with the site's exposure to high-density residential and commercial activities. The unique combination of human and non-human MST markers as well as the presence of various chemical indicators highlights the complex nature of potential contamination at BR3.

At BR4, a wide variety of microbial and chemical indicators were detected, including HF183, HumM2 (human marker), Rum2Bac, LeeSG, CG01, and 25 chemical indicators (Figure 4.3). Similar to BR3, all six chemical indicators previously identified at BR2 were also present at BR4, demonstrating the persistence of these contaminants in the downstream environment. Furthermore, a more diverse range of chemical indicators was observed at BR4, including a higher abundance of antidepressants (n=3), artificial sweeteners (2), antibiotics (4), analgesics (4), cardiovascular drugs (2), antiepileptics (2), opiates (1), and flame retardants (2). This result suggests that the downstream site is exposed to a broader spectrum of contaminants, likely due to the cumulative impact of multiple WWTP discharges as well as stormwater outfalls. However, the cumulative stormwater discharges are likely masked by the substantial

impact of the wastewater discharges. Previous research has identified high levels of chemical indicators (i.e., micropollutants) at this site, which aligns with the characteristics of the sample site location, being downstream of multiple WWTPs [31]. While the presence of these chemical indicators does not necessarily indicate direct human sewage contamination, the limited removal capabilities of the WWTPs in addition to stormwater are likely contributing to the detection of a wider range of chemical indicators at BR4. The presence of a diverse array of MST markers, including two human-specific markers, further supports the hypothesis that the downstream environment is impacted by treated sewage discharges.

Overall, the presence/absence plots demonstrate the distinct characteristics of the three sites, as evidenced by several unique microbial and chemical indicators at each location. While these plots provide the frequency of detection for various microbial and chemical indicators at the sites, they do not offer information regarding the magnitude of these detections. This limitation necessitates further exploration of spatio-temporal trends to develop a better understanding of the water quality along the BR.

4.2. Parameter Relationships with Flow

The temporal trends of the water quality parameters for the three study sites are presented in Figures 4.4 – 4.6. For this analysis, only representative indicators frequently detected at all sites were used to determine preliminary temporal patterns. For instance, enterococci and total coliforms were selected since they were the most frequently detected FIB (100% at all three sites). HF183 and CG01 were the predominant human- and non-human-associated MST markers, with detection frequencies increasing from BR2 (10.20% and 8.16% respectively) to BR4 (96.30% and 46.30% respectively). Sucralose was the most widely detected chemical indicator (100% at BR2 and BR4, 95.84% at BR3) at median concentrations above 26 ng/L (see Tables B.2 – B.4 in Appendix B).

Furthermore, given that BR flow shows a temporal pattern, the five parameters were compared with the flow at the time of sampling (Water Survey of Canada, Bow River at Calgary, Gauge ID: 05BH004) to identify seasonality within the dataset via a correlation analysis (R^2). The BR in particular

experiences low flow during the fall and winter, gradually rising in early spring (April - May) and peaking from mid to late June due to rainfall and glacial snowmelt. Hence, this distinct pattern can inform whether the presence of both microbial and chemical indicators is flow-enriched (i.e., higher concentrations at high flow conditions) or dominated by dilution patterns (i.e., lower concentrations at low flow).

Enterococci exhibited weak/low negative correlations while total coliforms showed low/moderate positive correlations with the BR flow at all sites (see Table 4.1). HF183 demonstrated weak/low negative correlations with the flow at BR3 (R^2: -0.35) and BR4 (R^2: -0.11). In contrast, sucralose displayed strong negative correlations with the flow at all three sample sites (BR2 R^2: -0.81, BR3 R^2: -0.70, BR4 R^2: -0.88). These findings suggest that the temporal variations of these indicators, particularly sucralose and total coliforms, may be influenced by hydrological conditions. For example, high flow conditions (e.g., spring snowmelt, rainfall events) can lead to increased levels of total coliforms in the BR, indicating its non-specificity to a particular source (i.e., can be from human sewage [point] or from other animals [diffused]). The total coliform trends also align with historical trends observed along the BR from 1951 – 1994, where total coliforms were found to have a positive dependency on the flow [237]. This result also indicates that the relationship between total coliforms and flow in the BR has remained relatively consistent over time, potentially as a result of increased bacterial inputs from runoff during higher flow events [237].

Table 4.1 *Pearson correlation coefficients (R^2) between Bow River (BR) flow and five representative parameters.*

	BR2	**BR3**	**BR4**
Enterococci	-0.13	-0.32	-0.17
Total coliforms	0.60	0.39	0.51
HF183	-	-0.35	-0.11
CG01	-	-0.42	0.18
Sucralose	-0.81	-0.70	-0.88

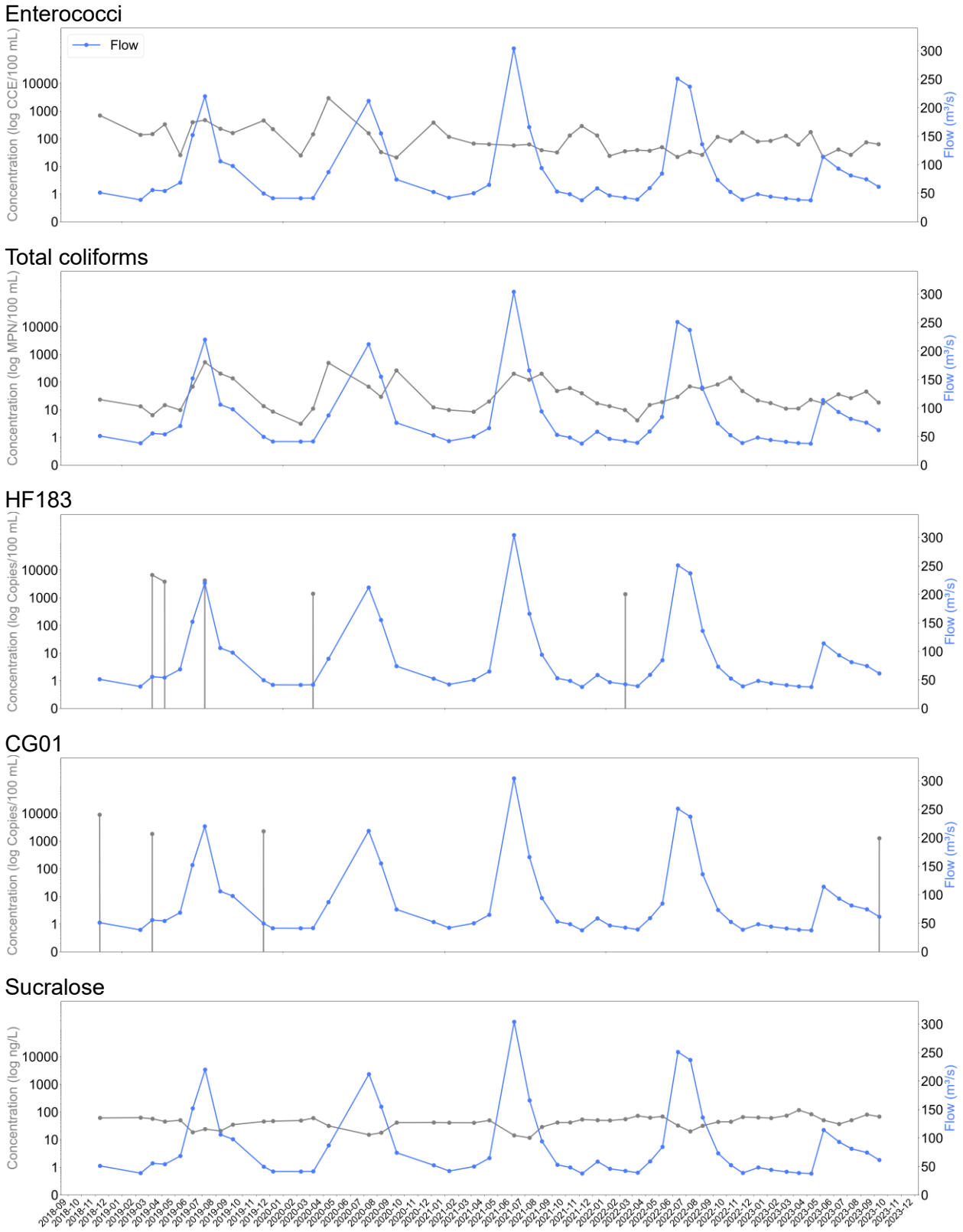


Figure 4.4 *Time series for enterococci, total coliforms, HF183, CG01 and sucralose at* ***BR2****. Vertical lines indicate periods of non-detection. Data points between vertical lines represent numerical detections.*

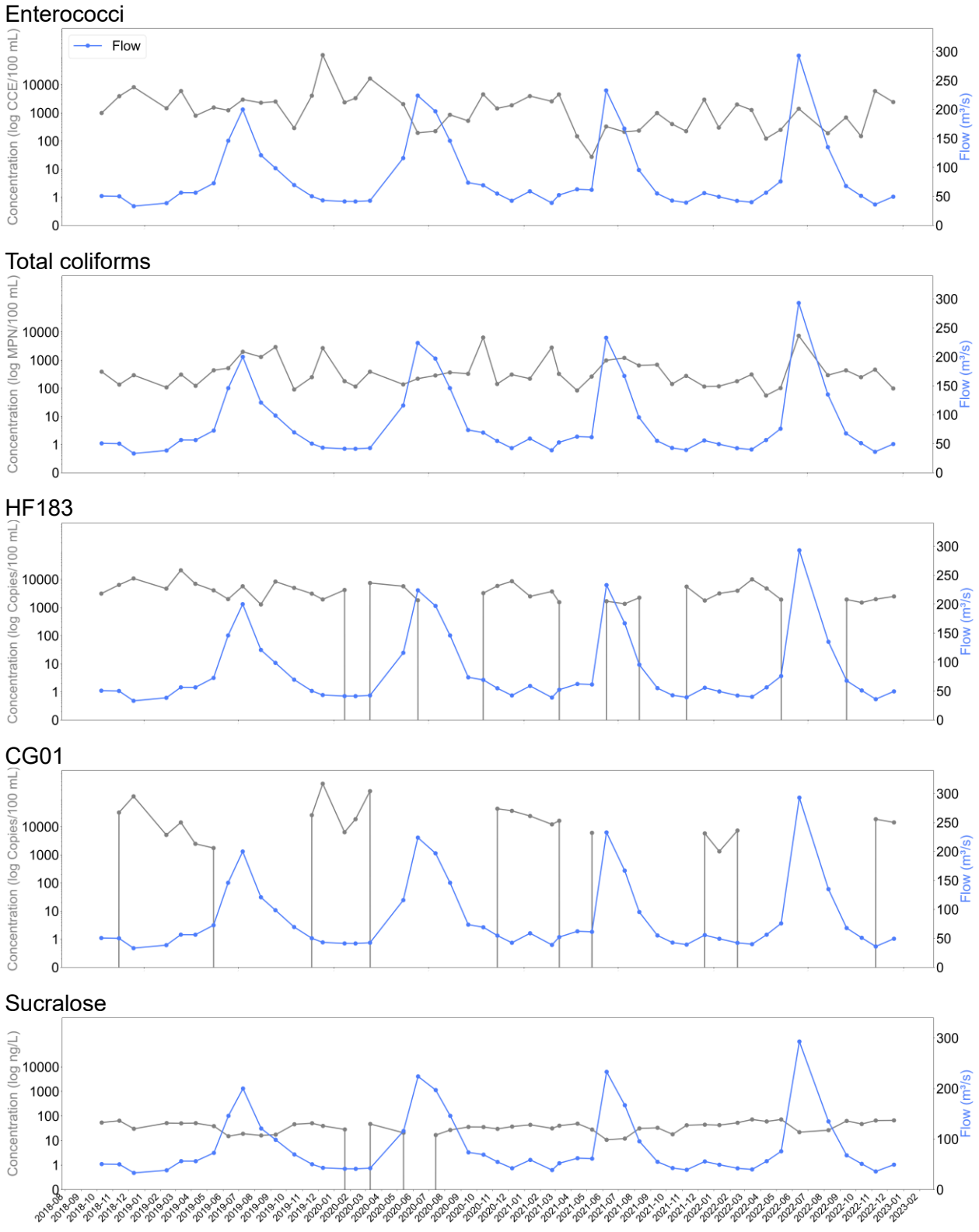


Figure 4.5 *Time series for enterococci, total coliforms, HF183, CG01 and sucralose at* ***BR3****. Vertical lines indicate periods of non-detection. Data points between vertical lines represent numerical detections.*

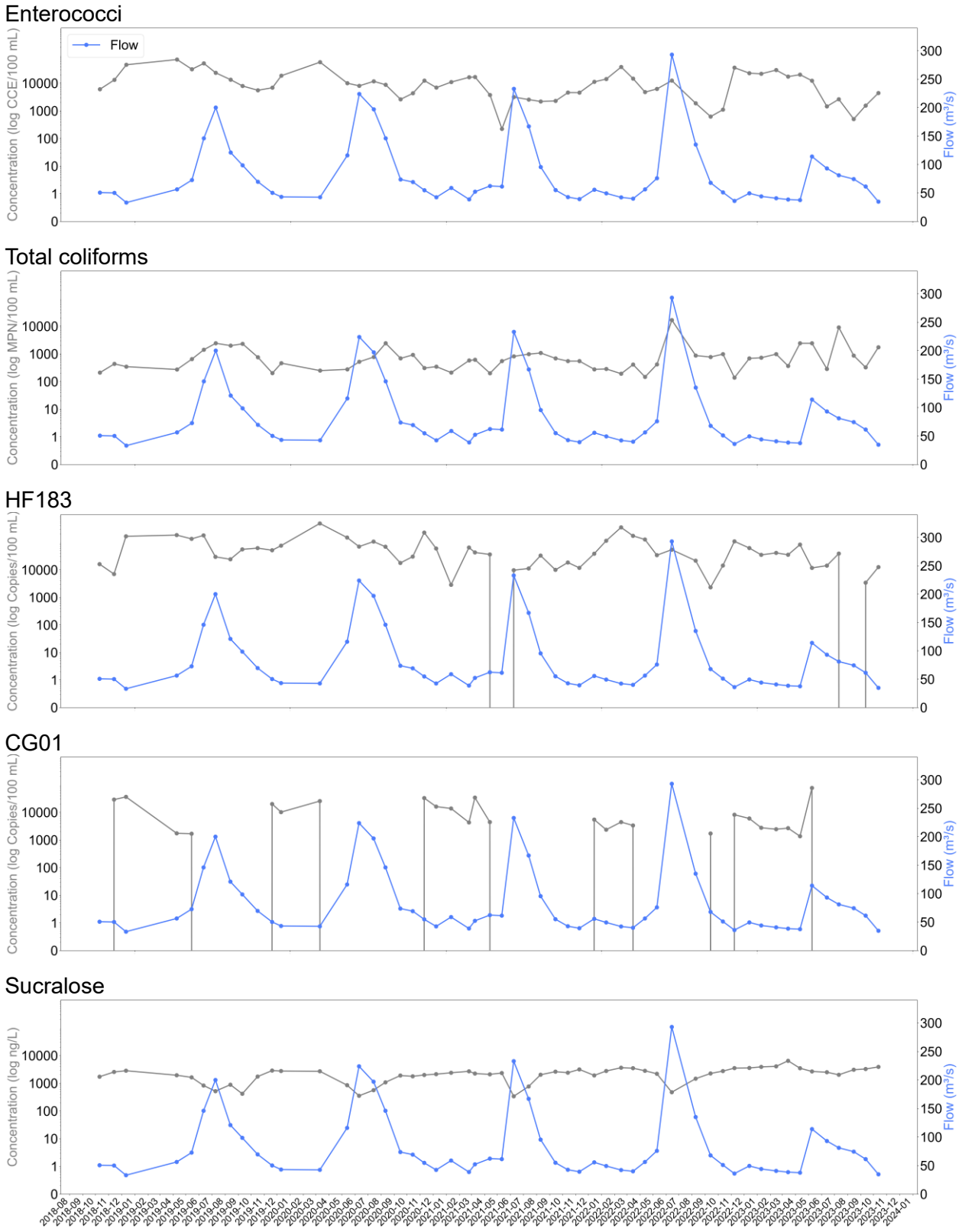


Figure 4.6 *Time series for enterococci, total coliforms, HF183, CG01 and sucralose at* ***BR4****. Vertical lines indicate periods of non-detection. Data points between vertical lines represent numerical detections.*

While the relationship of river flow with total coliforms is straightforward, it appears to be more complex for other FIB (e.g., enterococci, Figures 4.4 – 4.6). Previous studies often associate higher FIB concentrations with high-flow events and lower FIB concentrations with low-flow events [238], [239], [240], [241], but the findings suggest a more nuanced relationship at BR2. For instance, enterococci displayed a weak to moderate negative correlation with the flow at BR2, BR3, and BR4 (R^2: -0.13, -0.32, -0.17, respectively), indicating that other factors (i.e. dilution) may influence enterococci concentrations at the site.

The increase in negative correlation between enterococci from BR2 to BR3 also points to a potential point-source contamination in the area, including the stormwater outfalls within the sampling locations. At BR4, enterococci concentrations occasionally decreased prior to flow peaks, while total coliforms exhibited greater consistency (Figure 4.6). However, a weak negative correlation between enterococci and flow (R^2: -0.17) indicates that factors other than flow may be more influential to enterococci at this location.

Sucralose exhibited a clear inverse relationship with the flow at all three sites (BR2 R^2: -0.81, BR3 R^2: -0.70, BR4 R^2: -0.88), where higher concentrations were observed during low-flow events and lower concentrations were observed during high-flow events (Figures 4.4 – 4.6). This relationship is consistent with the findings of previous studies which indicate that dilution is a primary contributor to contaminant fate/transport in large riverine systems [31], [242], [243]. Since sucralose is widely used and is known to be a persistent chemical within aquatic environments, its detection at all three sites highlights the widespread distribution of this anthropogenic indicator along the BR [140], [157], [244], [245]. Although not presented in Figure 4.6, the BR flow was found to have a moderate positive correlation with LeeSG (R^2: 0.55) and total coliforms (R^2: 0.51) at BR4. However, the BR flow was negatively correlated with all chemical indicators except for BPS (R^2: 0.52), citalopram (R^2: 0.30), and DEET (R^2: 0.30). Of these five parameters, total coliforms and DEET had the highest numerical detections (100% and 83.34%, respectively). The positive correlation between total coliforms and flow suggests that increased flow rates may be associated with higher levels of total coliforms in the wastewater effluent due to wastewater

discharges or overflows. The negative correlations between flow and most chemical indicators align with findings from other studies that suggest that dilution and river processes can impact the levels of chemical indicators in aquatic environments [31], [242], [243].

4.3. Seasonal Differences in Magnitude

Box plots were used as a visual exploratory tool to compare concentration magnitudes (including uncensored data) among all parameters that had ≥ 7 detections at each site (Figures 4.7 – 4.9). The median concentration, first quartile (Q1), third quartile (Q3), IQR, and skewness for the three sites overall and seasonally are listed in Tables B.2 – B.4 in Appendix B. Among the FIB, enterococci consistently exhibited the highest median concentrations overall (BR2: 79.00 CCE/100 mL, BR3: 1,411.16 CCE/100 mL, BR4: 9,385.00 CCE/100 mL) and the highest IQRs (BR2: 123.20, BR3: 2,646.93, BR4: 13,294.47) at all three sampling sites, indicating that the enterococci data had a wider spread (i.e., increased variation) compared to the other FIB. As for MST markers, the indicators with the highest median concentrations at each site overall were Rum2Bac at BR2 (3,715.00 Copies/100 mL), CG01 at BR3 (15,189.00 Copies/100 mL), and HF183 at BR4 (40,107.00 Copies/100 mL). Sucralose had the highest median concentrations of all chemical indicators at all sites (BR2: 46.60 ng/L, BR3 overall: 39.02 ng/L, BR4: 2,300.00 ng/L). Chemical indicators at BR4 generally had higher median concentrations compared to BR2 and BR3, due to the influence of incomplete removal through the WWTPs upstream from the site.

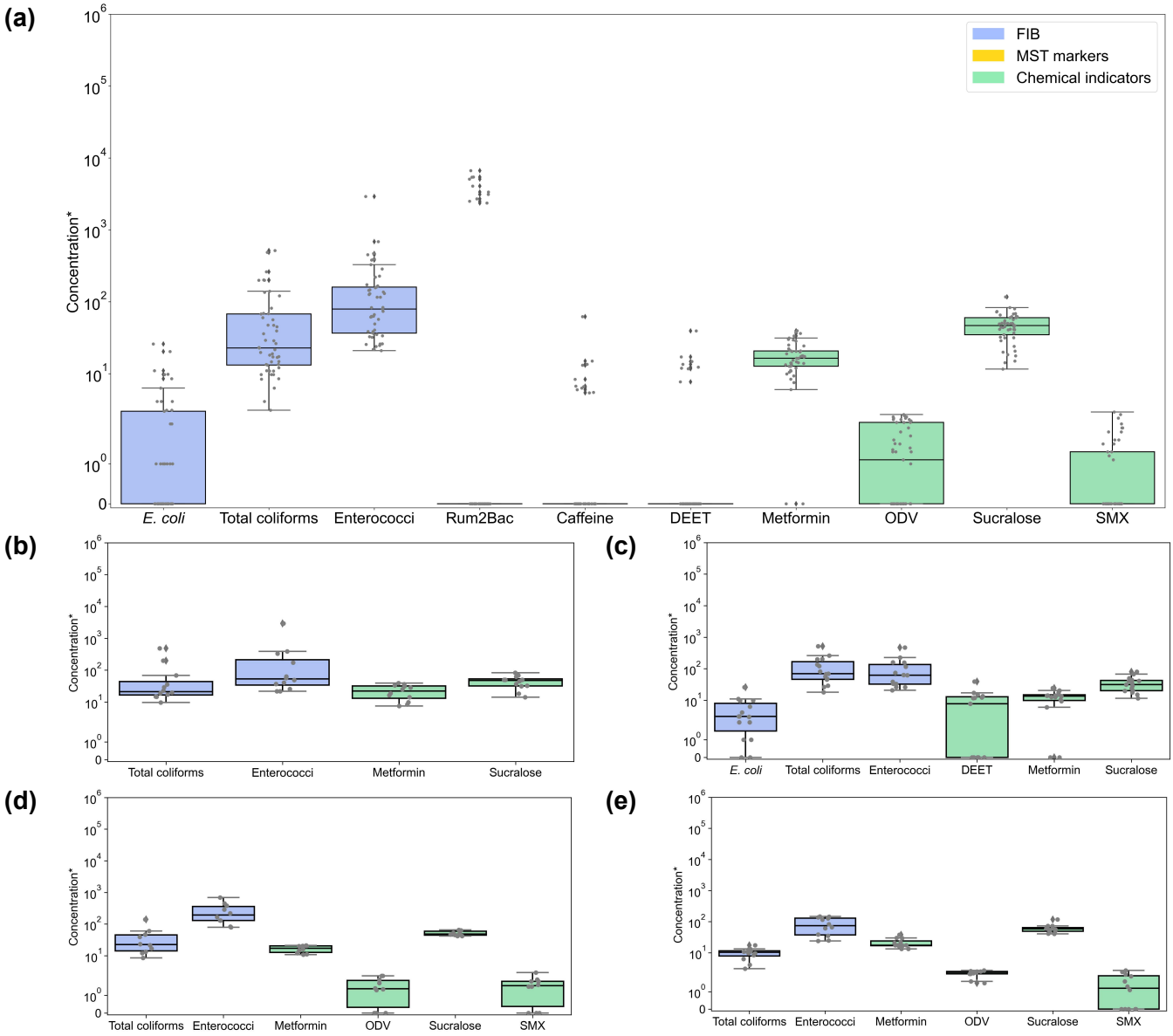


Figure 4.7 *Box plots of parameters with n ≥ 7 numerical detections for the (a) overall, (b) spring, (c) summer, (d) fall, and (e) winter data at* ***BR2****. *MPN/100 mL: E. coli, total coliforms; CCE/100 mL: Enterococci; Copies/100 mL: Microbial source tracking (MST) markers; ng/L: chemical indicators. MPN: Most probable number. CCE: Calibrator cell equivalent. DEET: N,N-diethyl-meta-toluamide. ODV: O-Desmethyl-venlafaxine. SMX: Sulfamethoxazole.*

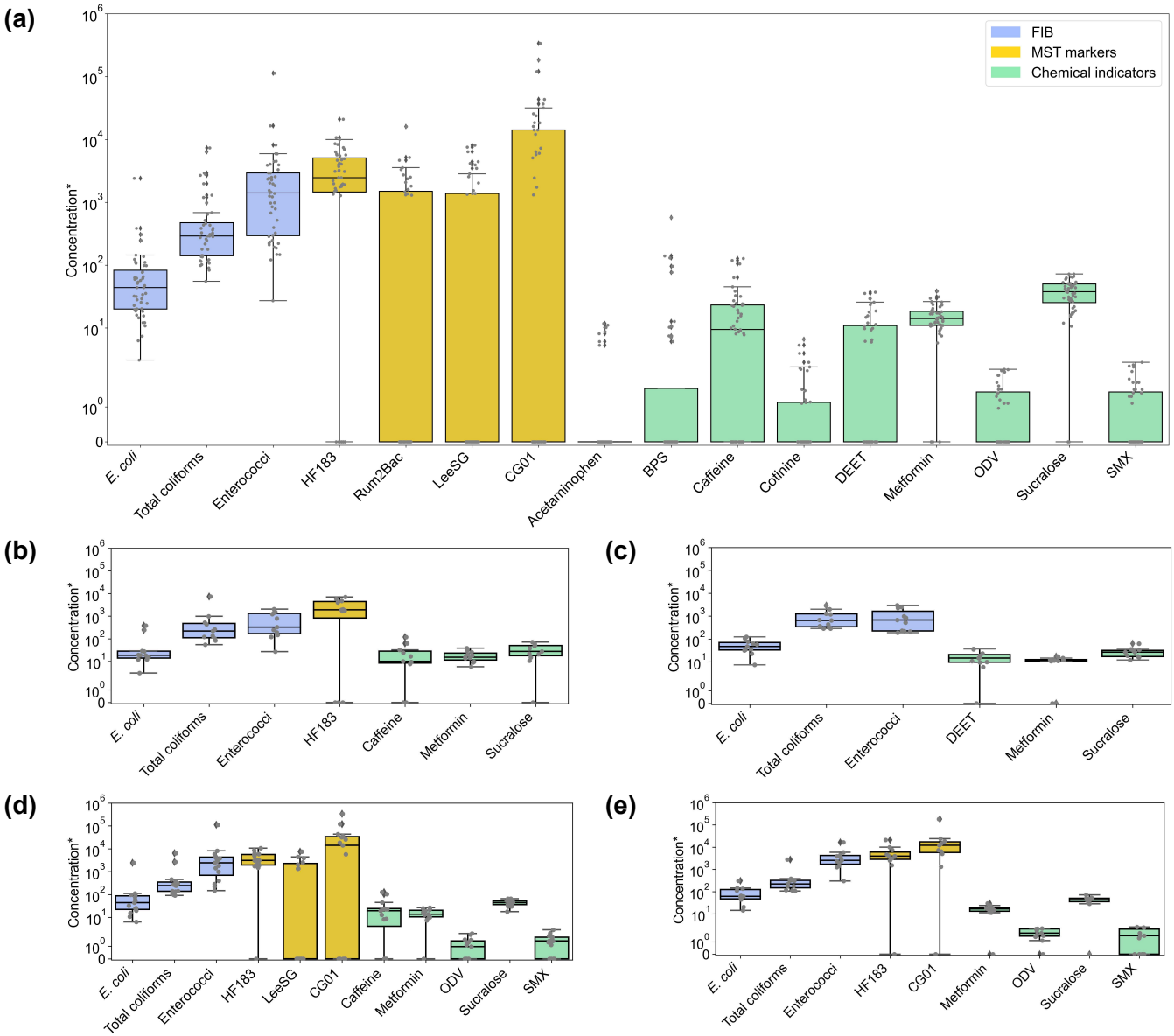


Figure 4.8 *Box plots of parameters with n ≥ 7 numerical detections for the (a) overall, (b) spring, (c) summer, (d) fall, and (e) winter data at* ***BR3****. *MPN/100 mL: E. coli, total coliforms; CCE/100 mL: Enterococci; Copies/100 mL: Microbial source tracking (MST) markers; ng/L: chemical indicators. MPN: Most probable number. CCE: Calibrator cell equivalent. BPS: Bisphenol S. DEET: N,N-diethyl-meta-toluamide. ODV: O-Desmethyl-venlafaxine. SMX: Sulfamethoxazole.*

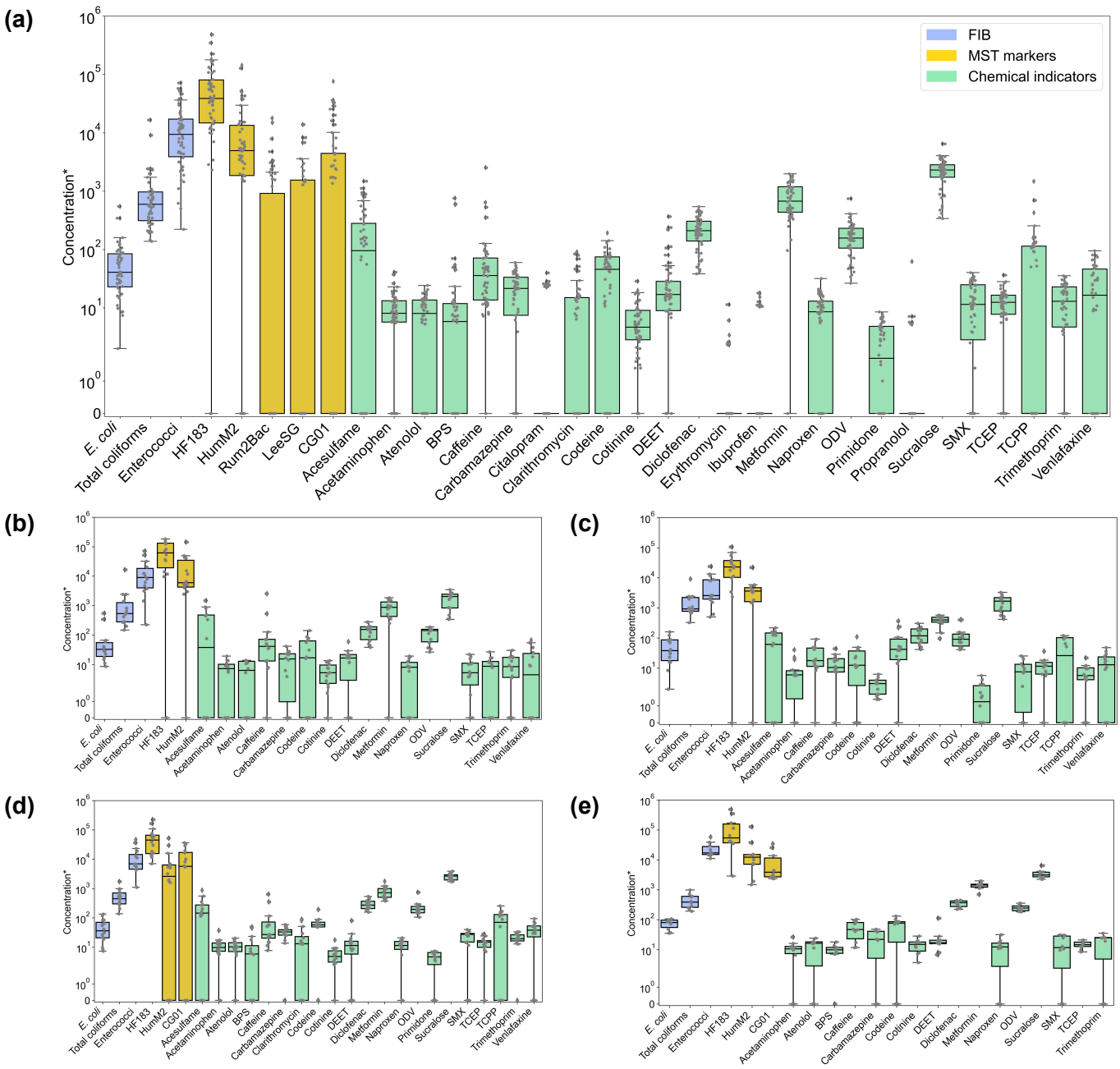


Figure 4.9 *Box plots of parameters with n ≥ 7 numerical detections for the (a) overall, (b) spring, (c) summer, (d) fall, and (e) winter data at* ***BR4****. *MPN/100 mL: E. coli, total coliforms; CCE/100 mL: Enterococci; Copies/100 mL: Microbial source tracking (MST) markers; ng/L: chemical indicators. MPN: Most probable number. CCE: Calibrator cell equivalent. BPS: Bisphenol S. DEET: N,N-diethyl-meta-toluamide. ODV: O-Desmethyl-venlafaxine. SMX: Sulfamethoxazole. TCEP: Tris(2-carboxyethyl)phosphine. TCPP: Tris(chloropropyl) phosphate.*

Furthermore, the boxplots for spring, summer, fall, and winter data for BR2 (Figures 4.7b – 4.7e) showed distinct patterns in the distribution of microbial and chemical indicators. For instance, *E. coli* was limited to the overall and summer datasets, with a consistent median concentration of 3.10 MPN/100 mL. The dataset also exhibited a tendency towards right-skewed distributions, with the exception of total coliforms (winter: -0.063), metformin (fall: -0.303), and ODV (overall: -0.145, winter: -1.036). For *E. coli*, a slightly lower IQR (6.95) overall compared to the summer (7.70) was observed, suggesting greater variability in the summer. The widest concentration range for enterococci was observed in the fall (IQR: 226.58), followed by the spring (IQR: 178.22). The winter dataset exhibited the least skewness (0.095), while the fall dataset had the highest median concentration (193.37 MPN/100 mL) and exhibited a more balanced distribution (skewness: 1.284). The highest median concentration for total coliforms was observed in the summer (69.10 MPN/100 mL) where the largest IQR (121.00) for the parameter was also seen. Rum2Bac, with analysis unique only to the overall dataset ($n \geq 7$), had a wider concentration range (IQR: 2,530.00) compared to the other microbial indicators. Note that filtering of data to only include detections that are ≥ 7 may be biased when looking at seasonal changes. Nonetheless, the presence of a non-human marker aligns with the understanding that the site has minimal human impacts, therefore the main source of fecal contamination likely originates from wild animals, particularly in surface waters with strong SWP [246].

As for the chemical indicators, ODV showed a narrow IQR (0.88) and low skewness (-0.146) overall but had distinct skewness characteristics across the fall and winter at BR2 (0.404 and -1.036, respectively). The higher detection rates in the fall (70%) and winter (100%) suggest that several factors influence its presence in the aquatic environment. Since ODV is a human metabolite of venlafaxine (an antidepressant), its elevated appearance in the winter months may be due to the fact that the half-life of venlafaxine is longer in the winter as a result of reduced sunlight exposure, leading to a longer degradation process [247]. Increased antidepressant prescription and usage during the fall and winter months, as observed in various Canadian studies, may also contribute to the seasonal detection of ODV [248].

Sucralose and metformin maintained relatively higher concentrations and more normal distributions compared to the other chemical indicators at BR2, with sucralose showing a wider concentration range in the overall dataset (IQR: 25.40) and varying median concentrations across the seasons. The highest median concentration was observed in the winter at 58.38 ng/L. SMX had lower concentrations compared to the other chemical indicators, with an overall median of 1.70 ng/L and a small IQR (0.48), however, elevated positive skewness was seen in the fall (2.109). The presence of metformin, ODV, sucralose, and SMX highlights their persistence in aquatic environments even in source water [244], [249], [250], [251]. The lower concentrations of these chemical indicators at BR2 compared to BR3 and BR4 suggest that their primary source may be located upstream (i.e. WWTP, sewage lagoon).

The seasonal boxplots for BR3 (Figure 4.8b – 4.8e) illustrate the influence of seasonality on the presence of microbial and chemical indicators. This is highlighted by the absence of MST markers during the summer, and the presence of migratory bird markers in the fall and winter. Metformin and sucralose were consistently detected throughout all seasons, with caffeine being limited to the spring and fall, DEET being limited to the summer, and ODV and SMX being present in the fall and winter. Like BR2, the BR3 dataset demonstrated right-skewed distributions for the majority of indicators, with the exception of acetaminophen (overall: -0.262) and ODV (winter: -0.476).

Overall at BR3 (Figure 4.8a), *E. coli* exhibited the lowest concentration among the FIB with a few outliers. The largest median concentration was seen in the summer (47.30 MPN/100 mL), with the lowest and highest IQRs being observed in the spring (14.55) and winter (76.35), respectively, suggesting greater variability in the winter data. Total coliforms displayed a wider concentration range overall (IQR: 334.20) compared to *E. coli* (IQR: 62.83), with a right-skewed distribution (3.413), indicating a tendency towards higher levels. Enterococci demonstrated the highest median concentration (1,411.16 CCE/100 mL) and IQR (2,646.93) of the FIB overall, indicating significant distribution variability. Based on the 2024 Health Canada Guidelines for Canadian recreational water quality, FIB concentrations that exceed ≤ 235 *E. coli* cfu/100 mL and ≤ 1,000 enterococci CCE/100 mL can present elevated health risks to the

public [101]. The concentrations of *E. coli* and enterococci exceeded these guidelines four (8.34%) and 27 (56.25%) times, respectively, within the overall dataset. Seasonally, *E. coli* and enterococci exceedances occurred 18.18% and 36.36% in the spring, 0% and 27.27% in the summer, 6.67% and 66.67% in the fall, and 9.09% and 90.90% in the winter, respectively. While *E. coli* levels exceeded the guideline only a few times, the frequent exceedance of the enterococci guideline suggests a higher likelihood of fecal contamination. Neither FIB are indicative of human contamination exclusively, however, other studies have noted that when FIB exceed such guidelines, this is often caused by sewage contamination [252]. In contrast, the presence of FIB was observed to be below these guidelines at BR2. Since the presence of fecal indicators does not always directly correlate with sewage contamination, this highlights the need to use both microbial and chemical indicators when determining the presence of contamination [253]. The increase in FIB concentrations from BR2 to BR3 aligns with the observation that stormwater is a potentially significant source of FIB, specifically *E. coli*, in urban sites from sewage and stormwater inputs [254].

HF183, Rum2Bac, LeeSG, and CG01 were detected at BR3, however, Rum2Bac did not meet the numerical detection criteria ($n \geq 7$) in the seasonal datasets. Rum2Bac had the lowest median concentration (2,235.00 Copies/100 mL) and IQR (1,836.00) compared to the other MST markers, while also displaying the largest skewness (3.318). CG01 had the largest median concentration overall (15,189.00 Copies/100 mL), followed by HF183 (3,480.00 Copies/100 mL) and LeeSG (3,165.00 Copies/100 mL). The presence of HF183 at BR3 is another indication of human sewage contamination at the site [255]. This pattern of elevated levels of human and non-human markers at the site aligns with previous findings, which suggest widespread sewage contamination in urban stormwater systems are associated with factors such as aging infrastructure, the presence of cross-connections, illicit discharges, and CSOs [147], [156], [168].

Of the chemical indicators detected at BR3, caffeine and sucralose had the highest concentrations overall (caffeine median: 20.75 ng/L, sucralose median: 39.02 ng/L) and seasonally (spring: 16.60 ng/L,

33.32 ng/L; summer: N/A, 26.40 ng/L; fall: 23.00 ng/L, 44.30 ng/L; winter: N/A, 45.45 ng/L respectively). While both were right-skewed, sucralose was closest to a normal distribution overall (skewness: 0.194). Similar to BR2, metformin and sucralose were detected across all seasons, with median sucralose concentrations exceeding those of metformin. In addition, overall median concentrations of caffeine (6.70 to 20.75 ng/L) and DEET (12.90 to 13.44 ng/L) increased, while the concentrations of metformin (17.00 to 14.55 ng/L), ODV (2.00 to 1.51 ng/L), sucralose (46.60 to 39.02 ng/L), and SMX (1.70 to 1.60 ng/L) decreased from BR2 to BR3. The increase in concentration and detection of acetaminophen (8.70 ng/L, 14.58%), BPS (11.65 ng/L, 25.00%), caffeine (20.75 ng/L, 62.50%), and cotinine (2.15 ng/L, 29.17%) are indicative of anthropogenic influences, demonstrating the potential for human contamination impacts at the site.

At BR4, *E. coli* consistently exhibited the lowest concentrations among the FIB across all datasets (Figures 4.9a – 4.9e). Similar median *E. coli* concentrations were seen at BR3 (44.10 MPN/100 mL) and BR4 (40.70 MPN/100 mL), and while the overall dataset at BR4 suggested a right-skewed distribution (4.036), seasonal variations were apparent. For instance, a left-skewed distribution during the winter (-0.554) suggests higher concentrations during the sampling period. Total coliforms showed a broader distribution compared to *E. coli*, with a larger IQR (*E.* coli IQR: 61.48, total coliforms IQR: 653.43) and a more pronounced right-skewness overall (5.179), indicating greater variability within the data. The highest median concentration for total coliforms was observed in the summer (925.40 MPN/100 mL) and the lowest median concentration in winter (387.70 MPN/100 mL). Similar to BR3, enterococci had the highest magnitude of concentrations of the FIB across the overall and seasonal datasets. The highest overall median concentration for enterococci was observed at BR4 (BR2: 79.00 CCE/100 mL, BR3: 1,411.16 CCE/100 mL, BR4: 9,385.00 CCE/100 mL). This indicates a higher prevalence of enterococci at BR4 may be due to wastewater treatment efficiency and upstream levels in the BR. Many studies have noted that enterococci are more resistant to environmental impacts (i.e., sunlight, salinity conditions) and wastewater treatment processes, resulting in elevated concentrations in wastewater effluents [256]. The

winter dataset revealed the largest median concentration (17,022.98 CCE/100 mL), followed by the spring (9,015.00 CCE/100 mL), fall (6,901.95 CCE/100 mL), and summer (2,580.00 CCE/100 mL) datasets, indicating seasonal variations. The skewness between datasets for enterococci remained relatively consistent, indicating a prevalence towards higher values.

Human MST markers were a major component of the microbial indicators at BR4 (Figures 4.9a – 4.9e), with HF183 demonstrating higher overall median concentrations (40,107.00 Copies/100 mL) compared to BR3 (3,480 Copies/100 mL). The distribution of HF183 was right-skewed, with the highest skewness value being seen for the overall dataset (2.712), suggesting a concentration of data points toward lower values, likely influenced by the conversion of non-detects to zeroes. Other markers, including HumM2 (85.19%), Rum2Bac (25.93%), LeeSG (31.48%), and CG01 (46.30%) were seen at BR4 and were less abundant compared to HF183 (96.30%). These markers also exhibited right-skewed distributions, with skewness ranging from 0.381 to 3.338. Seasonal variances were evident within the MST markers. The presence of both HF183 and HumM2 at this site reflects human fecal contamination stemming from either WWTPs or stormwater outfalls. While HF183 remained the dominant marker throughout the datasets, CG01 was the only non-human MST marker to meet the detection criteria in the fall and winter, suggesting environmental factors may influence its distribution within the aquatic environment. The increase in median MST marker concentrations during the winter, particularly for HumM2 (12,134.00 Copies/100 mL), may be attributed to reduced dilution effects caused by ice formation in the BR, a decline in the decay of MST markers in response to reduced sunlight, and lower rates of biological activity due to cooler temperatures [257]. However, it has also been noted that increased concentrations of human-associated markers in wastewater effluents during the winter can indicate poor WWTP performance [257].

Sucralose emerged as the predominant chemical indicator across all BR4 datasets with consistently high median concentrations ranging from 1,680 ng/L (summer) to 3,135 ng/L (winter) (Figures 4.9b – 4.9e). These concentrations were elevated compared to the other sample sites and reflect

the characteristics of sucralose as a domestic wastewater tracer [258]. The overall distribution of sucralose was right-skewed with the exception of the spring dataset which exhibited left-skewness (-0.108). Additional chemical indicators were observed at BR4 compared to BR2 and BR3 as a result of BR4 being downstream of three WWTPs. This aligns with our understanding of the sample site and the trends seen in the literature, as WWTPs are inefficient at removing micropollutants which cause WWTP effluents to be the main sources of micropollutants in aquatic environments [259], [260]. The occurrence of the chemical indicators typically varies between studies due to site characteristics, land use, population, and water consumption trends [261]. Similar to the FIB and MST markers, seasonal variances were observed regarding the chemical indicators at BR4. The highest number of chemical indicators that met the detection criteria was observed in the fall, followed by equal numbers of chemical indicators in the spring and summer, with the winter showing the fewest chemical indicators. Peak concentrations of the chemical indicators varied across seasons, suggesting that a multitude of factors influence their presence and distribution within the aquatic environment (i.e., effluent concentration variabilities, seasonal fluctuations, dilution, degradation, consumption rates) [262], [263], [264].

Overall, it is clear that although the sites are hydrologically connected (upstream to downstream), the presence of microbial and chemical indicators varies. To identify ideal indicators for site-specific fingerprints, correlation and cluster analysis were employed to examine additional relationships among the parameters.

4.4. Correlation Analyses – Data Transformation and Normality Test Results

A preliminary assessment of normality (see section 3.3) indicated that the datasets were non-normal, suggesting that data transformations are needed for subsequent statistical analyses (e.g., use of Pearson correlation coefficient for normal datasets only). The final decisions on data transformations, correlation, and clustering linkage methods are summarized in Table 4.2. Logarithm transformations were primarily applied to improve data normality, except for BR2-spring where the 1/x transformation was

used. The results of the normality tests are presented as a ratio of the number of parameters that were normally distributed after data transformation to the total number of parameters tested for the dataset. As can be observed, the data transformation improved the normality of datasets, with 50 – 100 % of the datasets conforming to normality. While the Shapiro-Wilk test returned the lowest normality results (BR2 all: 5/10, BR3 all: 11/16, BR4 all: 14/33), the use of two additional normality tests (D'Agostino's K^2, Jarque-Bera) provided multiple lines of evidence to support data transformation approaches. Based on the results, the Pearson correlation method on transformed datasets was used as the more appropriate method, followed by the Ward agglomerative clustering linkage method (see section 3.3.2).

Table 4.2 *Summary of data transformation, correlation, and clustering linkage methods selected for each site. a) The log_{10}, 1/x, and √x data transformations generated the same normality test results. The log_{10} transformation was chosen since it is used most often across all sites and seasons. b) The log_{10} and √x data transformations generated the same normality test results. The log_{10} transformation was chosen since it is used most often across all sites and seasons. c) The log_{10} and √x data transformations generated the same normality test results. The log_{10} transformation was chosen since it is used most often across all sites and seasons. N/A: D'Agostino's K^2 normality test requires all parameters to have a minimum of 8 numerical values. Any parameter with 7 numerical detections caused the test to fail.*

		Data transformation	**Normality test** Shapiro-Wilk	D'Agostino's K^2	Jarque-Bera	**Correlation method**	**Clustering linkage method**
BR2	All	log_{10}	5/10	7/10	8/10	Pearson	Ward
	Spring	1/x	2/4	¾	4/4	Pearson	Ward
	Summer	log_{10}	6/6	5/6	6/6	Pearson	Ward
	Fall	log_{10}[a]	5/6	N/A	6/6	Pearson	Ward
	Winter	log_{10}[b]	6/6	N/A	6/6	Pearson	Ward
BR3	All	log_{10}	11/16	N/A	13/16	Pearson	Ward
	Spring	log_{10}	6/7	6/7	7/7	Pearson	Ward
	Summer	log_{10}[c]	6/6	6/6	6/6	Pearson	Ward
	Fall	log_{10}	9/11	N/A	9/11	Pearson	Ward
	Winter	log_{10}	8/9	N/A	8/9	Pearson	Ward
BR4	All	log_{10}	14/33	N/A	26/33	Pearson	Ward
	Spring	log_{10}	19/22	N/A	22/22	Pearson	Ward
	Summer	log_{10}	20/22	N/A	21/22	Pearson	Ward
	Fall	log_{10}	23/27	24/27	25/27	Pearson	Ward
	Winter	log_{10}	20/22	N/A	22/22	Pearson	Ward

4.5. Correlation and Cluster Analysis for Site-Specific Fingerprinting of Fecal Indicators

4.5.1 BR2 – Source Water

When considering all datasets from 2018 – 2023, the correlation analysis of the combined BR2 dataset revealed perfect positive correlations between *E. coli* and caffeine, Rum2Bac and DEET, and a strong positive correlation between Rum2Bac and SMX (0.99) (Figure 4.10a). Correlations between parameters are pairwise, therefore the correlations are only examined when the two parameters in question are present. While the correlation analysis shows strong and perfect relationships between certain parameters, the low detection rates of these parameters at the site (*E. coli*: 44.90%, Rum2Bac: 20.41%, caffeine: 22.45%, DEET: 16.33%, SMX: 32.65%) suggest that these correlations may be an artifact of the analysis. Hence, caution must be exercised when interpreting the results.

E. coli and caffeine demonstrated a perfect positive correlation. Daneshvar et al. also observed a positive relationship between fecal coliform and caffeine (Spearman correlation: 0.45) when they examined the Greater Montréal area source waters [265]. This finding aligns with the growing body of evidence supporting caffeine as a potential indicator of fecal contamination [28], [141], [143], [155]. Nevertheless, the degradation of caffeine in aquatic environments, as discussed earlier (see section 2.5), necessitates caution, i.e., only elevated caffeine levels at BR2 (e.g., ~3x higher than the median concentrations at BR2) may suggest substantial fecal contamination, and confirmation should be sought through the analysis of additional fecal indicators (e.g., FIB, MST markers).

Although there have been many studies which analyze the correlations between microbial markers, studies which use a combined MST and CST approach rarely explore correlations between parameters quantitatively (see Table 2.4) or consider temporal factors when assessing contaminant levels (as most are completed as one-time sampling campaigns) [141]. Correlation analyses performed indicate moderate correlations between *E. coli* and total coliforms (0.62) as well as between enterococci and DEET (0.51) (Figure 4.10a). In contrast, strong negative correlations were shown between *E. coli* and both ODV (-0.83) and sucralose (-0.75). Similarly, moderate negative correlations were observed between

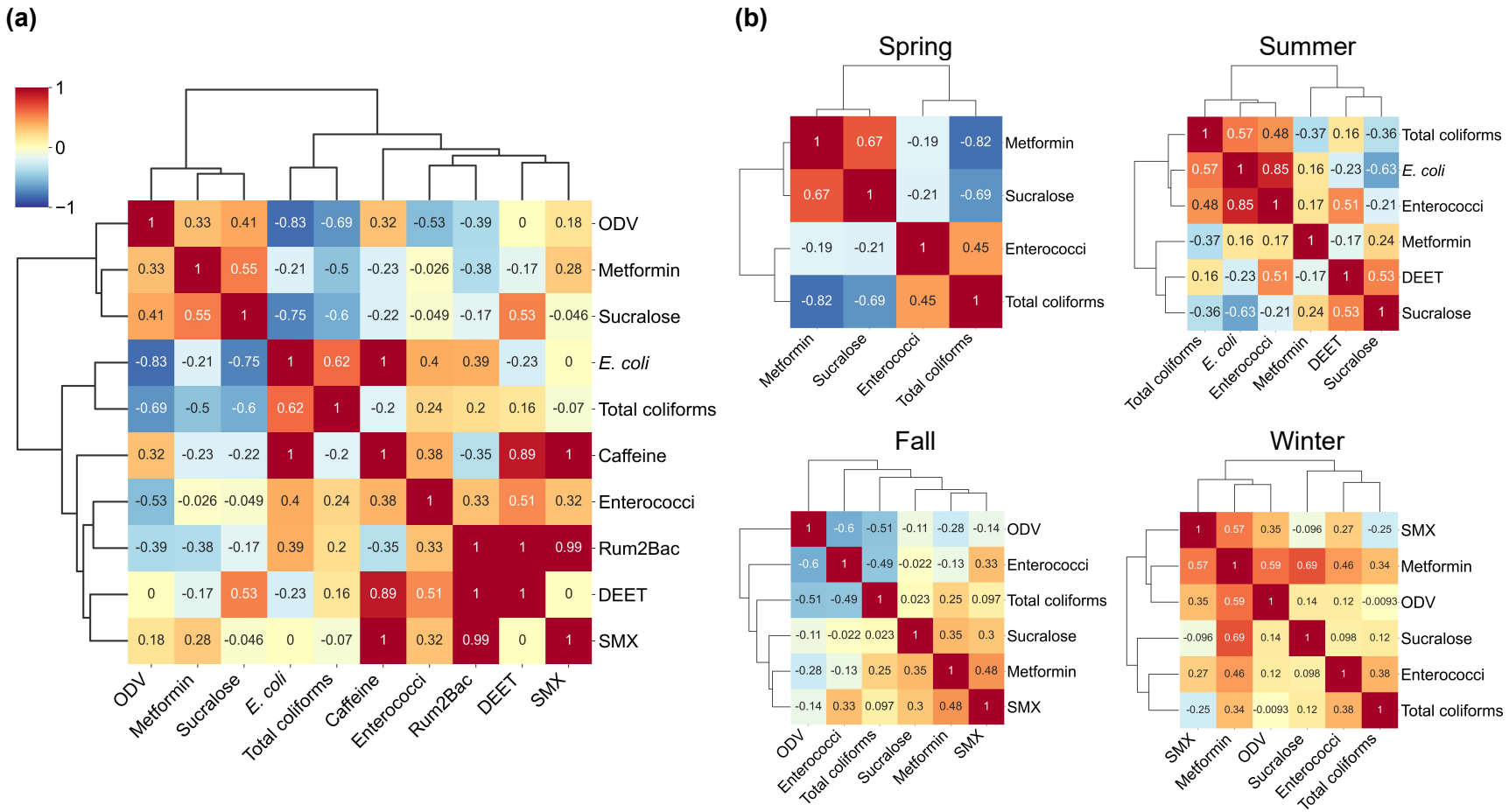


Figure 4.10 *Correlation and clustering of the (a) overall and (b) seasonal data for* ***BR2****. DEET: N,N-diethyl-meta-toluamide. ODV: O-Desmethyl-venlafaxine. SMX: Sulfamethoxazole.*

enterococci and ODV (-0.53) and total coliforms with metformin (-0.50) ODV (-0.69), and sucralose (-0.60). The observed negative correlation between total coliforms and sucralose is noteworthy given the positive moderate correlation between total coliforms and flow (0.60) and the strong negative correlation between sucralose and flow (-0.81) as indicated earlier. In addition, the flow was also negatively correlated with all six chemical indicators at BR2 (caffeine: -0.042, DEET: -0.69, metformin: -0.35, ODV: -0.64, sucralose: -0.81, SMX: -0.33). This suggests point-source origins of the chemical indicators, i.e., low dilution during low-flow events which can elevate dissolved contaminant concentrations. As noted in the literature, contaminants entering a waterway are likely to be retained longer during low-flow conditions compared to the rapid transit and dilution that occur under high-flow conditions [266], [267]. This result further indicates that under low-flow conditions, the detection of chemical substances in source waters may be a more dominant indicator. In contrast, microbial indicators could carry greater significance as markers of fecal contamination during high-flow and runoff events.

These relationships also exist when examining the datasets seasonally (Figure 4.10b). For instance, during the spring, strong negative correlations were found between total coliforms and both metformin (-0.82) and sucralose (-0.69), showing an increased relationship compared to the overall dataset. A strong positive correlation emerged between *E. coli* and enterococci (0.85) during the summer, indicating a strong direct relationship between the two parameters. This finding aligns with a previous study [33] which observed seasonal variations in *E. coli* concentrations, such as an increase in concentrations from the winter to the summer at a water treatment plant intake. However, this seasonal pattern was not observed at a secondary water treatment plant intake site, suggesting that local factors and waterbody characteristics influence the seasonal variations in *E. coli* levels [33]. The correlation between *E. coli* and enterococci during the summer at BR2 may also be influenced by the fact that *E. coli* only met the numerical detection limit during the overall and summer datasets, limiting the ability to assess relationships between the two parameters in other seasons. Moderate correlations were also seen between enterococci and DEET (0.51), *E. coli* and both total coliforms (0.57) and sucralose (-0.63). The fall was characterized by a moderate negative correlation between enterococci and ODV (-0.60). The winter

displayed moderate positive correlations between metformin and ODV (0.59), sucralose (0.69), and SMX (0.57). Finally, it appears that the fall dataset is the primary driver of the overall dataset, as evidenced by the high degree of similarity between the correlation trends observed in the fall and the overall datasets.

Considering the combined dataset at BR2, three primary clusters emerged: 1) a chemical indicator cluster consisting of ODV, metformin, and sucralose, 2) an FIB cluster comprised of *E. coli* and total coliforms, and 3) a combination of microbial and chemical indicators including caffeine, enterococci, Rum2Bac, DEET, and SMX (Figure 4.10a). Upon examining the seasonal cluster maps (Figure 4.10b), the spring dataset exhibited one chemical indicator cluster (metformin and sucralose), and one FIB cluster (enterococci and total coliforms). Similar to the overall dataset, parameters within each cluster demonstrated positive correlations, suggesting potential co-occurrence in the environment. The summer dataset demonstrated an association between all three FIB while metformin, DEET, and sucralose formed a separate cluster, mirroring the clustering observed in the overall and spring datasets. This pattern reinforces the tendency for FIB and chemical indicators to form independent clusters at BR2. In the fall, a gradual expansion of a cluster initiated by metformin and SMX was observed, encompassing additional parameters as the cluster grew. Lastly, the winter dataset displayed clear divisions of clusters, with one group containing SMX, metformin and ODV, and the other comprising sucralose, enterococci, and total coliforms. This is the first instance in the seasonal data where both microbial and chemical indicators are found in an independent cluster.

Despite several relationships observed via clustering, the consistent clustering of metformin with sucralose suggests a potential common source or similar environmental factors influencing their distribution at BR2. Similar to sucralose, studies have found that metformin is persistent in the environment [147], [148], with metformin being reported as the second most frequently detected pharmaceutical in rivers around the world as of 2022 [268]. As observed earlier in the section, metformin is also detected at BR3 and BR4 which supports the inclusion of metformin in fingerprinting approaches. What sets BR2 apart from the other sample sites is the unique relationship between metformin and enterococci. Metformin, an antidiabetic, can be attributed to anthropogenic sources, and the clustering of

metformin with enterococci would strengthen the indication of fecal contamination at the site. This combination of indicators, particularly their frequent negative correlations, suggests a specific interaction (i.e., environmental factors favouring the persistence of one indicator over the other) or pollution source that is not as prevalent at other sites.

4.5.2 BR3 – Urban Site

The correlation analysis of BR3 showed an increase in complexity in the relationships between parameters at the sample site compared to BR2 (Figure 4.11). Overall, strong positive correlations were seen between LeeSG and CG01 (0.77), LeeSG and acetaminophen (0.95), LeeSG and DEET (0.75), enterococci and CG01 (0.72), BPS and ODV (0.76), BPS and DEET (0.70), BPS and Rum2Bac (0.79), acetaminophen and DEET (0.88), and Rum2Bac and cotinine (0.93). A perfect positive correlation was observed between BPS and acetaminophen. On the other hand, strong negative correlations were seen between CG01 and BPS (-0.83), Rum2Bac and ODV (-0.70), as well as between acetaminophen and SMX (-0.92). However, similar to BR2, many of these relationships may be an artifact of the analysis since the detection rates of these parameters ranged from 14.58 to 41.67% (CG01: 45.84%, LeeSG: 29.17%, Rum2Bac: 33.34%, acetaminophen: 14.58%, BPS: 25.00%, cotinine: 29.17%, DEET: 41.67%, ODV: 41.67%, SMX: 37.50%), with the exception of enterococci, which was consistently detected.

Other studies which have analyzed relationships between microbial and chemical indicators at urban sites often express that the water quality at such sites can be influenced by various sources (i.e., stormwater runoff, aging infrastructure, CSOs, cross-connections). Similar to BR2, total coliforms and sucralose were negatively correlated overall (-0.48) while flow and total coliforms were positively correlated (0.39). In addition, the flow was positively correlated with LeeSG (0.50) and had both positive and negative correlations with the chemical indicators. This mirrors the trends seen in the literature, where many factors influence the presence of chemical indicators in addition to flow.

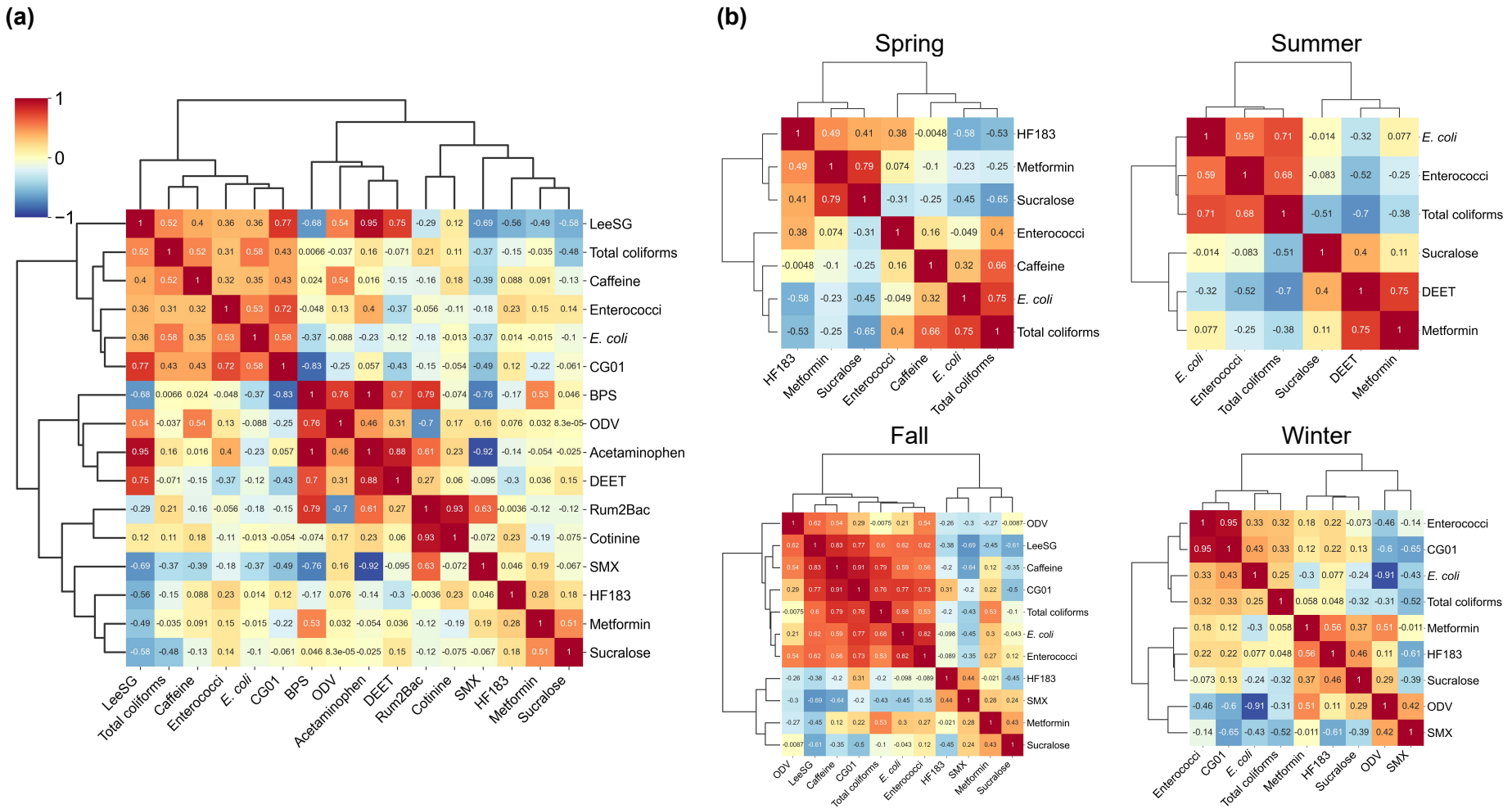


Figure 4.11 *Correlation and clustering of the (a) overall and (b) seasonal data for* ***BR3****. BPS: Bisphenol S. DEET: N,N-diethyl-meta-toluamide. ODV: O-Desmethyl-venlafaxine. SMX: Sulfamethoxazole.*

From the overall dataset, Rum2Bac and LeeSG emerged as MST markers with significant relationships with both FIB and chemical indicators. Rum2Bac exhibited moderate positive relationships with multiple chemical indicators, including acetaminophen (0.61), and SMX (0.63), as well as strong positive correlations with BPS (0.79) and cotinine (0.93). In the BR2 overall correlation matrix (Figure 4.10a), Rum2Bac was strongly correlated with SMX (0.99) and perfectly correlated with DEET, which contrasts the moderate and weak correlations seen in BR3 between Rum2Bac, SMX (0.63), and DEET (0.27), respectively. With a higher number of detections of Rum2Bac at BR3 (33.34%) compared to BR2 (20.41%), the correlation analysis becomes more robust.

These findings also indicate that the relationships between MST markers and micropollutants are site-specific and may be influenced by land use and hydrological conditions. From the literature review, only one study identified Rum2Bac as a microbial indicator of interest when analyzing the relationships between microbial and chemical indicators [136]. The study identified rainfall as a key factor influencing the transport of the ruminant marker, emphasizing the role of catchment characteristics in this process. Though the catchments studied were predominantly in rural settings, the correlations between Rum2Bac and rainfall suggest that rainfall-driven events may contribute to contaminant travel and mobilization within aquatic environments [136].

In contrast, LeeSG was characterized by moderate negative concentrations with HF183 (-0.56), BPS (-0.68), sucralose (-0.58), and SMX (-0.69). CG01, the second marker for a bird host, was also negatively correlated with BPS (-0.83) and SMX (-0.49). The seagull-host marker also showed moderate positive correlations with total coliforms (0.52) and ODV (0.54), as well as strong positive correlations with CG01 (0.77), acetaminophen (0.95) and DEET (0.75). While previous studies have found significant associations between seagull markers and *E. coli* in rivers, one study which investigated sources of fecal contamination in the Humber River, ON, reported Spearman correlation coefficients ranging from 0.23 to 0.75 between the gull marker and several chemical indicators, including acetaminophen, acesulfame, caffeine, carbamazepine, codeine, and cotinine in a river environment [28].

Additional relationships within the overall dataset for BR3 revealed moderate positive correlations between *E. coli* and both enterococci (0.53) and total coliforms (0.58) (Figure 4.11a). Increased correlations between *E. coli* and enterococci from BR2 to BR3 are potential indicators of increased human contamination inputs. Moderate to strong correlations between CG01 with *E. coli* (0.58) and enterococci (0.72) highlight that the site is impacted by human and non-human fecal sources. The increased relationship from BR2 to BR3 between total coliforms and caffeine (BR2: -0.20, BR3: 0.52) is an additional indication of human fecal pollution. As previously mentioned, caffeine has been identified as a potential sewage-associated marker since it is often linked to human contamination and degrades rapidly in the environment. The observed positive correlation between total coliforms and caffeine at BR3 suggests recent inputs of untreated sewage at this site. This correlation alludes to the use of caffeine as a marker for recent untreated effluent discharges as indicated in Table 2.4.

Two of the relationships seen in the overall dataset (LeeSG and CG01, LeeSG and SMX) are seen in the fall since it is the primary driver of the overall correlation and clustering as it has the largest dataset of the seasons (n = 15) (Figure 4.11b). During the spring, strong positive correlations were seen between sucralose and metformin (0.79), and *E. coli* and total coliforms (0.75). Strong negative correlations also emerged between *E. coli* and HF183 (-0.58), total coliforms and HF183 (-0.53), as well as between total coliforms and sucralose (-0.65). The relationship between *E. coli* and total coliforms decreases slightly during the summer (0.71), and metformin and DEET are found to have a strong positive relationship (0.75). This aligns with our understanding that DEET is predominantly used and detected during the summer. However, DEET is negatively correlated with total coliforms (-0.70). The positive relationship between DEET and metformin is an additional indication of anthropogenic impacts at BR3. In the fall, strong positive correlations are seen between LeeSG, caffeine (0.83) and CG01 (0.77), CG01 with caffeine (0.91), *E. coli* (0.77), enterococci (0.73), and total coliforms (0.76). In addition, strong relationships were also found between total coliforms and caffeine (0.79), and *E. coli* and enterococci (0.82). Similar to caffeine, metformin has been identified as a sewage indicator to trace human-related

pharmaceutical contamination in aquatic environments [147], [249], [269], [270]. Therefore, the low to moderate relationships between metformin, *E. coli* (0.30) and total coliforms (0.53) are an additional indication of sewage impacts at BR3. In contrast, moderate negative correlations were observed between LeeSG, SMX (-0.69), and sucralose (-0.61), as well as caffeine and SMX (-0.64). The winter displayed a strong positive correlation between enterococci and CG01 (0.95), and a strong negative correlation between *E. coli* and ODV (-0.91).

The overall cluster analysis at BR3 revealed four main clusters (Figure 4.11a). The first cluster encompasses a diverse group of microbial and chemical indicators including LeeSG, total coliforms, caffeine, enterococci, *E. coli*, and CG01. Within this cluster, a sub-grouping emerged between enterococci, *E. coli*, and CG01, while total coliforms and caffeine formed another distinct sub-cluster. The second cluster consisted of multiple chemical indicators (BPS, ODV, acetaminophen, and DEET). Cotinine and Rum2Bac formed a solitary cluster, and the fourth cluster comprised SMX, HF183, metformin, and sucralose, with metformin and sucralose forming a distinct sub-cluster. Similar to the overall and seasonal cluster maps for BR2, the majority of parameters that formed clusters were positively correlated. In addition, several parameter pairs found in clusters at BR2 were also observed to cluster together at BR3, including total coliforms and *E. coli*, caffeine and enterococci, and sucralose and metformin. The fact that these parameters appeared in clusters at both sites indicates potential similarities regarding environmental conditions and persistent contaminant sources. The detection, correlation, and clustering of compounds such as metformin and sucralose, both known for their persistence in aquatic environments, at both BR2 and BR3 imply that these contaminants may originate from sources located further upstream or from additional inputs (i.e., stormwater outfalls, CSOs, WWTP effluents, anthropogenic activities) [31], [271].

The spring dataset at BR3 yielded two primary clusters: 1) comprising HF183, metformin, and sucralose, and 2) comprising of the FIB and caffeine (Figure 4.11b). During the summer, two primary clusters also emerged: 1) containing the FIB, and 2) consisting of sucralose, DEET, and metformin. Of the

seasonal data, the microbial and chemical indicators cluster separately from each other in the summer and the winter. The fall was characterized by three clusters: 1) with ODV, LeeSG, caffeine, CG01, and the FIB, 2) including HF183 and SMX, and 3) comprising metformin and sucralose. The majority of the parameters in the first cluster of the fall data are also found in the first cluster of the overall data. Lastly, the winter data also contains three clusters: 1) containing the FIB and CG01, 2) with metformin, HF183, and sucralose, and 3) comprising ODV and SMX. The second and third clusters combine to create a larger cluster, and this is similar to the indicators present in the second cluster of the overall dataset. These seasonal cluster maps for the sample site provide a more focused relationship between the parameters, highlighting the relationships between FIB, HF183, metformin, caffeine, and sucralose. These seasonal clustering patterns, in addition to the positive correlations observed, highlight the strong association between microbial and chemical indicators, suggesting potential human sewage contamination as a source of these contaminants at BR3.

While HF183 is not exclusively found at BR3, its presence in an urban site is noteworthy since it is indicative of human fecal contamination. Coupled with its positive correlation with various FIB, MST markers, and chemical indicators overall and seasonally, HF183 is a potential indicator that can be used for fingerprinting at BR3. The clustering of HF183 and sucralose in the overall dataset combined with the larger magnitude of their correlations both overall and seasonally compared to BR4 reinforces their uniqueness as indicators. Therefore, HF183 and sucralose are strong candidates for fingerprinting indicators at BR3.

4.5.3 BR4 – Downstream of WWTPs

Figure 4.12 illustrates the relationships between microbial and chemical indicators within the overall data at BR4, downstream of the City of Calgary's three WWTPs. As previously stated, the increased detection of chemical indicators aligns with the site's characteristics since Calgary's WWTPs are not capable of complete chemical removal prior to discharge into the BR, contributing to their elevated presence in the receiving water [259], [260]. Similar to BR2 and BR3, strong to perfect correlations between Rum2Bac, LeeSG, CG01, and other parameters may again be a result of artifacts of the analysis as a result of their low detection rates (Rum2Bac: 25.93%, LeeSG: 31.48%, CG01: 46.30%) within the overall dataset. Moderate positive correlations were seen between HF183 and enterococci (0.69) and between total coliforms and flow (0.51). Positive correlations were also observed between *E. coli*, erythromycin (0.77) and propranolol (0.55). In contrast, negative moderate to strong correlations were seen between HF183 and citalopram (-0.61), HumM2 and propranolol (-0.58), as well as between total coliforms and multiple chemical indicators. These findings align with previous research at sites downstream of WWTPs, which have reported positive correlations between HF183 and enterococci [141], and positive Spearman correlations between *E. coli*, caffeine, codeine, cotinine, and acetaminophen [28]. The positive correlation between FIB and HF183, as well as the increased presence of chemical indicators is not surprising since BR4 is significantly impacted by wastewater discharges.

In the overall data (Figure 4.12), low positive correlations were found between *E. coli* and both Enterococci (0.36) and total coliforms (0.32). These relationships are lower than those observed in BR2 and BR3, suggesting that diverse pollution sources, rather than human sewage contamination, impact the site. Furthermore, moderate and strong positive correlations were seen between *E. coli,* propranolol (0.55) and erythromycin (0.77), respectively, while a moderate positive correlation was observed between enterococci and HF183 (0.69). The increased correlation between enterococci and HF183 from BR3 (0.23) to BR4 (0.69) may reflect the presence of viable and non-viable bacteria detected in wastewater effluents, given the known association of elevated HF183 concentrations with WWTP effluents [272].

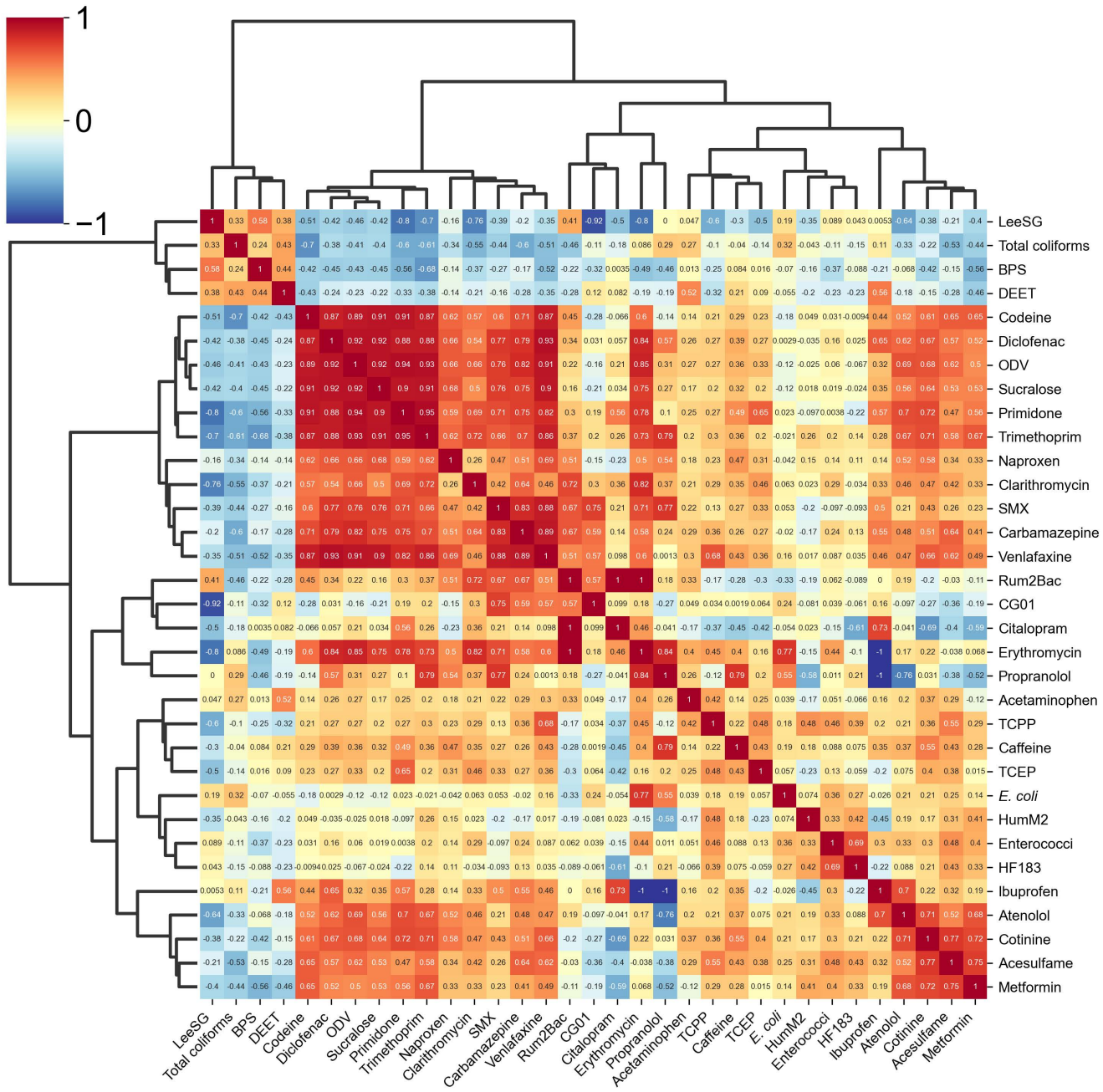


Figure 4.12 *Correlation and clustering of the overall data for* ***BR4****. BPS: Bisphenol S. DEET: N,N-diethyl-meta-toluamide. ODV: O-Desmethyl-venlafaxine. SMX: Sulfamethoxazole. TCEP: Tris(2-carboxyethyl)phosphine. TCPP: Tris(chloropropyl) phosphate.*

Total coliforms exhibited negative correlations with several chemical indicators, including acesulfame (-0.53), carbamazepine (-0.60), clarithromycin (-0.55), codeine (-0.70), primidone (-0.60), trimethoprim (-0.61), and venlafaxine (-0.51). While the relationship between total coliforms and chemical indicators within WWTP effluents has not been studied extensively, these findings suggest that flow conditions may contribute to the relationships. This phenomenon may also apply to the relationship between HumM2 and propranolol (-0.58). In addition, the correlation between total coliforms and caffeine decreased from BR3 (0.52) to BR4 (-0.04). It is currently difficult to assess whether this is mainly due to treated effluent discharge as there are also stormwater outfalls between BR3 and BR4.

Contrasting patterns were observed regarding the correlations between Rum2Bac, LeeSG, and chemical indicators. Rum2Bac displayed moderate to perfect positive correlations with multiple chemical indicators, while LeeSG exhibited numerous moderate to strong negative correlations with the chemical indicators. A strong negative correlation between LeeSG and CG01 (-0.92) was also observed. These differing relationships between the MST markers and chemical indicators point to divergent behaviour of ruminant, seagull, and Canada goose species in the aquatic environment at BR4.

The seasonal correlation matrices for BR4 are shown in Figure 4.13. The general trends observed in the overall correlation analysis persisted across all seasons, however, the strengths of these correlations varied. This highlights the dynamic interactions between indicators at the site throughout the year. Although the summer exhibited strong similarities to the overall dataset, specifically regarding the chemical indicators, the larger sample size in the fall (n = 16) likely contributed more significantly to the overall correlation patterns. *E. coli* exhibited positive correlations with total coliforms across all seasons except for fall (spring: 0.62, summer: 0.67, fall: -0.22, winter: 0.46). The highest positive correlations between *E. coli,* FIB, and MST markers were observed during the summer, coinciding with the most pronounced negative correlations between *E. coli* and several chemical indicators. In contrast, *E. coli* demonstrated positive correlations with a broad range of chemical indicators during the fall.

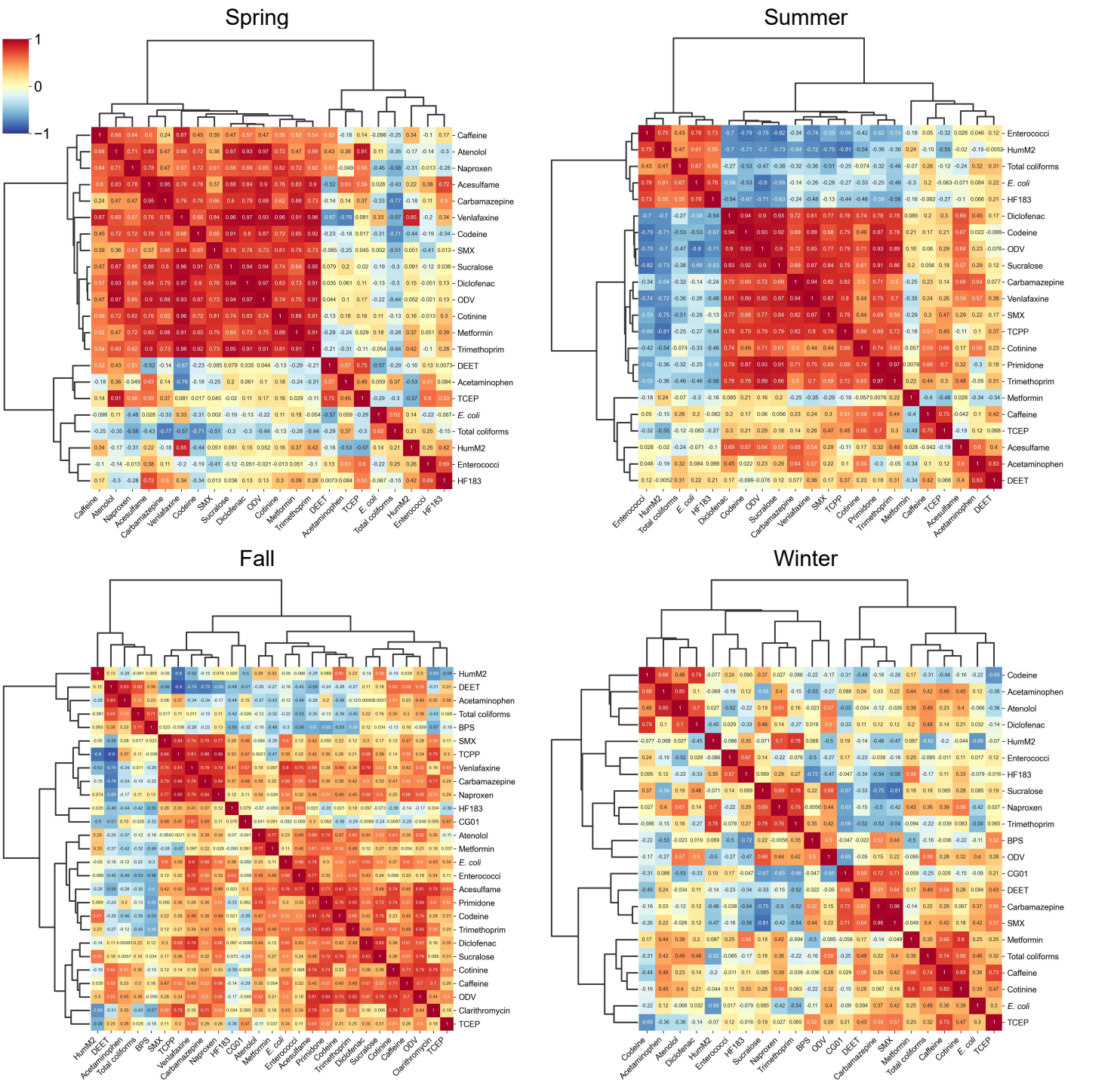


Figure 4.13 *Correlation and clustering of the seasonal data for* ***BR4****. BPS: Bisphenol S. DEET: N,N-diethyl-meta-toluamide. ODV: O-Desmethyl-venlafaxine. SMX: Sulfamethoxazole. TCEP: Tris(2-carboxyethyl)phosphine. TCPP: Tris(chloropropyl) phosphate.*

Throughout all seasons (Figure 4.13), enterococci exhibited consistent moderate to strong positive correlations with HF183 (spring: 0.69, summer: 0.73, fall: 0.63, winter: 0.67). The strongest correlation occurred during the summer, coinciding with the highest negative correlations between enterococci and the majority of the chemical indicators. Similar to *E. coli* and enterococci, total coliforms displayed the highest positive correlations with FIB and MST markers during the summer, coinciding with the most pronounced negative correlations with the chemical indicators. It was also observed that total coliforms were negatively correlated with *E. coli* (-0.22), enterococci (-0.53), and HF183 (-0.42) during the fall, while these parameters were positively correlated during the summer (0.67, 0.43, 0.55, respectively). HF183 and HumM2 displayed similar correlations to the FIB. These relationships may be a result of variations in flow, WWTP effluent discharge, growth rates, as well as environmental impacts (i.e., rainfall, and ice formation).

CG01, primarily detected during the fall and winter, exhibited positive correlations with carbamazepine, SMX, tris(2-carboxyethyl)phosphine (TCEP), tris(chloropropyl) phosphate (TCPP), and venlafaxine, while negative correlations were seen with HumM2, atenolol, ODV, sucralose, and trimethoprim across both seasons to varying degrees, with the highest correlations being seen during the winter. As previously discussed, seasonal environmental factors (i.e., ice formation, snowmelt) may influence the distribution and detection of CG01 [273], [274], [246].

The chemical indicators demonstrated the highest positive correlations amongst themselves during the spring, followed by the summer. While generally positively correlated in the winter, the lowest positive correlation and most negative relationships between chemical indicators were observed during this season. DEET consistently exhibited negative correlations with most parameters, except during the summer.

The cluster analysis for the overall BR4 dataset revealed a complex pattern as a result of the elevated number of microbial and chemical indicators detected at the site (Figure 4.12). Six distinct clusters emerged. The first, which included LeeSG, total coliforms, BPS, and DEET, suggested potential

relationships between both microbial and chemical indicators. The second cluster comprised a wide range of pharmaceuticals in addition to sucralose, highlighting the prevalence of micropollutants at the site. The third cluster consisted of Rum2Bac, CG01, and three pharmaceuticals. The fourth cluster grouped acetaminophen, TCPP, caffeine, and TCEP, all chemical indicators, once again indicating the co-occurrence of chemical indicators at BR4. *E. coli*, HumM2, enterococci, and HF183 were found in the fifth cluster, and the sixth cluster grouped ibuprofen, atenolol, cotinine, acesulfame, and metformin. The presence of multiple clusters encompassing both microbial and chemical indicators at BR4 suggests a higher degree of anthropogenic impact compared to the other sites. This is consistent with the expectation that these compounds are often co-released within WWTP effluents. While some similarities exist, such as the presence of FIB and chemical indicators clustering at all three sites, the co-occurrence of diverse compounds and unique microbial and chemical interactions at BR4 indicates a distinct contaminant profile unique to WWTP effluents.

Considering the seasonal cluster maps for BR4 (Figure 4.13), the spring and summer both contained a dominant cluster encompassing a wide range of chemical indicators, namely pharmaceuticals, suggesting consistent inputs of these compounds into the BR during these seasons. Microbial indicators (*E. coli*, enterococci, total coliforms, HF183, HumM2) clustered together during these two seasons, however, the correlations between them were more pronounced during the summer, showcasing seasonal variations in their behaviours within the environment. Both the fall and winter exhibited more fragmented clustering patterns characterized by multiple small clusters, suggesting increased variability in the presence of microbial and chemical indicators. This is shown through the fact that microbial indicators were present in clusters along with chemical indicators, whereas microbial indicators formed separate clusters during the spring and summer months.

As a marker associated with human fecal contamination, HumM2's exclusive presence at BR4 emphasizes its relevance as a specific indicator for this location. HumM2 is a less sensitive human marker compared to HF183 [232], therefore, its presence at BR4 is indicative of elevated levels of anthropogenic

contamination [168]. While diclofenac was detected in all samples taken at BR4, it is not specific to humans as it is also used as a veterinary drug [275]. Therefore, carbamazepine and metformin are ideal chemical indicators for fingerprinting at this site due to their high detection rate (79.63% and 98.15%, respectively), association with anthropogenic activities, and previous research highlighting the pharmaceuticals as indicators of wastewater in the aquatic environment [140], [142].

4.5.4 Recommended Site-Specific Composite Fingerprints

The analysis of the presence/absence, time series, box plots, correlations and clustering reveal the similarity and uniqueness of microbial and chemical indicators across the three sample sites. The development of baseline trends supports the utility of multiple indicators that can be used to understand fecal contamination in the BR. These **site-specific composite fingerprints** will be useful in future monitoring campaigns and will support decisions related to Calgary's SWP plan and One Water approaches. Many of the parameters recommended are monitored regularly by the City of Calgary as part of regulatory compliance (e.g., FIB) [256]. While 56 micropollutants comprise a comprehensive list, it is economical to focus on a few as high-throughput chemical analysis is an expensive and labour-intensive undertaking. For BR2 (source water site), if chemical indicators are found at elevated concentrations (e.g., ~3x the median concentration) and/or human markers are present, this may trigger concern and can be used to aid decisions related to sewage contamination. The fall was the most influential season due to elevated detections (BR2: 10/49, BR3: 15/48, BR4: 16/54).

The purpose of the site-specific fingerprints is to differentiate between water samples taken at the three sites based on the indicators present. The use of these fingerprints will allow for a more targeted analysis of domestic sewage contamination along the BR by reducing the list of indicators analyzed by the City of Calgary. Deviations from the expected indicators within samples can be used to inform additional monitoring of fecal contamination at the sites. Including a limited set of indicators in sampling campaigns will allow for more frequent sampling (i.e., by potentially eliminating the need for a holding

time) and more effective resource allocation based on the City's needs. Table 4.3 outlines the recommended site-specific fingerprints for BR2, BR3, and BR4 in addition to the rationales for the inclusion of each microbial and chemical indicator. These fingerprints are specific to the three sites analyzed along the BR and are not representative of other watersheds.

The common microbial and chemical indicator candidates for all sites are: 1) three traditional FIB, 2) Rum2Bac, and 3) metformin, sucralose, and ODV. The FIB were chosen for all sites due to their known association with fecal matter, high detection rates, and existing monitoring plans as part of regulatory compliance. Rum2Bac was selected because it was the only MST marker detected at BR2, and CG01 was chosen for BR3 and BR4 because it had the highest median concentration of the non-human MST markers at both sites. The two human markers (HF183 and HumM2) are also included since HF183 had the highest detection rate among the MST markers at BR3 in addition to the highest median concentrations at BR4, while HumM2 was exclusive to BR4. Caffeine was added for BR3 and BR4 fingerprinting due to its frequent detection at the downstream sites. Lastly, carbamazepine and diclofenac were included due to their elevated and unique presence at BR4 as well as their inclusion in existing and proposed water quality guidelines (i.e., Alberta Surface Waters guideline for carbamazepine and proposed EU guideline for diclofenac) [276], [277].

Table 4.3 *Candidate microbial and chemical indicators for site-specific fingerprinting. FIB: Fecal indicator bacteria. MST: Microbial source tracking. ODV: O-Desmethyl-venlafaxine.*

Sample Site	**FIB**		**MST markers**		**Chemical indicators**		**Season of importance**
	Candidates	Rationale	Candidates	Rationale	Candidates	Rationale	
BR2	*E. coli*, total coliforms	Monitored daily by the City of Calgary	Rum2Bac	No human markers present	Metformin, sucralose, ODV	Environmental persistence; relationships between metformin and sucralose indicate similar sources	Fall
	Enterococci	Public health-related guidance					
BR3	*E. coli*, total coliforms	Monitored monthly by the City of Calgary	Rum2Bac CG01	Consistent detections Highest median concentrations of the MST markers	Caffeine	Increased median concentrations from BR2 to BR3 suggest urban inputs	Fall
	Enterococci	Public health-related guidance	HF183	Human marker indicating human fecal contamination	Metformin, sucralose	Consistent positive correlation and clustering; high detection rates	
					ODV	Increased detections from BR2 to BR3	
BR4	*E. coli*, total coliforms	Monitored monthly by the City of Calgary	Rum2Bac CG01	Consistent detections Highest median concentrations of the non-human markers	Caffeine, ODV	Increased concentrations and detections from BR3 to BR4	Fall
	Enterococci	Public health-related guidance	HF183, HumM2	Human markers indicating human fecal contamination and/or wastewater effluent	Carbamazepine	High detection rates; substance of concern in the Environmental Quality Guidelines for Alberta Surface Waters [276]	
					Diclofenac	Consistent detection; often exceeds the EU Water Framework Directive's proposed guidance value (50 ng/L) [277]	
					Metformin, sucralose	High detection rates; consistent positive correlation and clustering	

Chapter 5 – Conclusions and Recommendations

This study focused on the analysis of three FIB, 9 MST markers and 56 micropollutants (i.e., chemical indicators) at the following sites near Calgary along the BR from 2018 to 2023: 1) an upstream source water site (BR2), an intermediate urban site (BR3), and a site which experiences the cumulative effects of the City's three WWTPs (BR4). BR2, being a source water site, had the lowest number of parameters detected, while BR4, located downstream of the WWTPs, exhibited the highest number of parameters, reflecting the accumulation of substances in the BR. FIB and non-human markers were detected at all three sites. Rum2Bac was the only MST marker present at BR2, while Rum2Bac, LeeSG, and CG01 were found at BR3 and BR4. The presence of the ruminant marker at all three sites reflects significant ruminant (i.e., deer) activity. Detections of Canada goose and gull markers at BR3 and BR4 may be influenced by habitat, food availability, and nesting in proximity to the sampling sites. HF183 was detected at BR3 and BR4, while HumM2 was found exclusively at BR4, indicating human impacts at both sites. Six chemical indicators were found across all three sample sites (caffeine, DEET, metformin, ODV, sucralose, SMX), demonstrating persistent anthropogenic impacts within the BR. The increased presence of chemical indicators at BR3 and BR4 is likely due to recent untreated sewage releases (i.e., caffeine) and cumulative impacts of multiple WWTP discharges, respectively. Comparing the most commonly found indicators across all sites (enterococci, total coliforms, HF183, CG01, and sucralose) to the BR flow, total coliform concentrations showed low to moderate positive correlations while sucralose had strong negative correlations. These relationships suggest that total coliforms are enriched while sucralose is diluted during high-flow conditions.

Box plot analysis revealed that enterococci had the highest median concentration magnitude among the FIB. While MST markers varied by site, HF183 was prevalent at BR4, likely indicating wastewater effluent impacts. Rum2Bac was the only MST marker at BR2 and CG01 was the dominant MST marker at BR3, reflecting wildlife activities at both sites. Enterococci often exceeded Health Canada's recreational water quality guidelines at BR3, suggesting a higher likelihood of fecal

contamination and public health risks. Sucralose was the most prevalent chemical indicator at all sites, particularly at BR4, likely due to incomplete removal during the wastewater treatment processes upstream.

Correlation and cluster analysis revealed complex interactions between microbial and chemical indicators at the sample sites. BR2 showed moderate to strong negative correlations between *E. coli*, ODV, and sucralose, enterococci and ODV, as well as between total coliforms, ODV, metformin, and sucralose. These negative relationships likely reflect similar point-source origins. Distinct clusters were often found, with only one cluster containing both microbial and chemical indicators in the overall and winter datasets. At BR3, significant relationships emerged between Rum2Bac, LeeSG, FIB, and chemical indicators. Increased correlations between *E. coli* and enterococci from BR2 to BR3 highlight the impact of various fecal sources, while the increased relationship between total coliforms and caffeine points to human contamination. The majority of the overall clusters were a combination of microbial and chemical indicators, with only one cluster containing only chemical indicators. This was also observed seasonally, with the exception of the summer dataset where microbial and chemical indicators clustered separately. BR4 was characterized by positive correlations between FIB and HF183, as well as between *E. coli*, erythromycin and propranolol. HF183 and total coliforms both demonstrated negative correlations with various chemical indicators. The positive relationship between HF183 and FIB as well as the increased presence of chemical indicators at BR4 suggest wastewater discharge impacts rather than point-source contamination. Clusters containing microbial and chemical indicators as well as a combination of indicators were observed in the overall BR4 dataset. A dominant pharmaceutical cluster was observed in the spring and summer, while more fragmented cluster patterns were seen during the fall and winter. The fall was the seasonal driver of the overall datasets at all three sites due to the elevated number of detections.

This work showcases the potential for the relationships between microbial indicators and micropollutants to be used as indications of domestic sewage pollution in aquatic environments. In

addition, this analysis provides a baseline of microbial and chemical indicators present at three sites along the BR which can be utilized by the City of Calgary when monitoring domestic sewage contamination.

The following recommendations have been summarized for future research related to fecal contamination near the City of Calgary along the BR:

- Monitor three FIB (*E. coli*, total coliforms, enterococci), 4 MST markers (Rum2Bac, CG01, HF183, HumM2), and 6 micropollutants (caffeine, carbamazepine, diclofenac, metformin, sucralose, and ODV) for site-specific fingerprinting due to their presence (unique or persistent) and prevalence.
- Investigate relationships between traditional water quality parameters (i.e., pH, temperature, total suspended solids [TSS], DO).
- Expand monitoring to include sites a) upstream of BR2 to develop a comprehensive understanding of the influences of septic, stormwater, and wastewater impacts, b) along NC to differentiate between NC point-source and stormwater influences at BR3, and c) immediately upstream of the WWTPs to identify the impact of wastewater discharge (including removal efficiencies) on the presence and concentrations of chemical indicators at BR4.

Bibliography

[1] S. Watson and J. Lawrence, "Drinking water quality and sustainability," *Water Qual. Res. J. Can.*, vol. 38, pp. 3–13, Jan. 2003.

[2] M. Salvadori, J. Sontrop, A. Garg, L. Moist, and W. Clark, "Factors that led to the Walkerton tragedy," *Kidney Int. Suppl.*, vol. 75, pp. S33-4, Mar. 2009, doi: 10.1038/ki.2008.616.

[3] P. B. Jameson, Y.-T. Hung, C. Y. Kuo, and P. A. Bosela, "*Cryptosporidium* Outbreak (Water Treatment Failure): North Battleford, Saskatchewan, Spring 2001," *J. Perform. Constr. Facil.*, vol. 22, no. 5, pp. 342–347, Oct. 2008, doi: 10.1061/(ASCE)0887-3828(2008)22:5(342).

[4] R. J. Patrick, "Adapting to Climate Change Through Source Water Protection: Case Studies from Alberta and Saskatchewan, Canada," *Int. Indig. Policy J.*, vol. 9, no. 3, Art. no. 3, Aug. 2018, doi: 10.18584/iipj.2018.9.3.1.

[5] "Source Water Protection Plan." City of Calgary, 2020. Accessed: Feb. 22, 2024. [Online]. Available: https://www.calgary.ca/content/dam/www/ca/city-clerks/documents/council-policy-library/cp2020-04-source-water-protection-policy.pdf

[6] J. I. Price, M. T. Heberling, and C. T. Nietch, "Economic Support for Decisions on Source Water Protection," *J. AWWA*, vol. 110, no. 9, pp. 56–61, 2018, doi: 10.1002/awwa.1153.

[7] G. T. Daigger, S. Sharvelle, M. Arabi, and N. G. Love, "Progress and Promise Transitioning to the One Water/Resource Recovery Integrated Urban Water Management Systems," *J. Environ. Eng.*, vol. 145, no. 10, p. 04019061, Oct. 2019, doi: 10.1061/(ASCE)EE.1943-7870.0001552.

[8] S. R. Pokhrel *et al.*, "Conceptual model of one water community: A path to move forward," *J. Clean. Prod.*, vol. 415, p. 137886, Aug. 2023, doi: 10.1016/j.jclepro.2023.137886.

[9] S. R. Pokhrel, G. Chhipi-Shrestha, K. Hewage, and R. Sadiq, "Sustainable, resilient, and reliable urban water systems: making the case for a 'one water' approach," *Environ. Rev.*, vol. 30, no. 1, pp. 10–29, Mar. 2022, doi: 10.1139/er-2020-0090.

[10] "Water Quality and Wastewater." United Nations, Sep. 2018. Accessed: Jun. 20, 2024. [Online]. Available: https://www.unwater.org/sites/default/files/app/uploads/2018/10/WaterFacts_water_and_watewater_sep2018.pdf

[11] Z. Wang, D. Shao, and P. Westerhoff, "Wastewater discharge impact on drinking water sources along the Yangtze River (China)," *Sci. Total Environ.*, vol. 599–600, pp. 1399–1407, Dec. 2017, doi: 10.1016/j.scitotenv.2017.05.078.

[12] A. Müller, H. Österlund, J. Marsalek, and M. Viklander, "The pollution conveyed by urban runoff: A review of sources," *Sci. Total Environ.*, vol. 709, no. 2020, p. 136125, Mar. 2020, doi: https://doi.org/10.1016/j.scitotenv.2019.136125.

[13] S. Mcgrane, "Impacts of urbanisation on hydrological and water quality dynamics, and urban water management: a review," *Hydrol. Sci. J.*, vol. 61, no. 13, pp. 2295–2311, 2016, doi: https://doi.org/10.1080/02626667.2015.1128084.

[14] B. Wijesiri, E. Bandala, A. Liu, and A. Goonetilleke, "A Framework for Stormwater Quality Modelling under the Effects of Climate Change to Enhance Reuse," *Sustainability*, vol. 12, no. 24, p. 10463, Dec. 2020, doi: 10.3390/su122410463.

[15] J. Pouliotte, H. Sandhu, and L. Gibb, "Source Water Protection Policy," Calgary, UCS2020-1007, Sep. 2020. Accessed: Jul. 28, 2024. [Online]. Available: https://pub-calgary.escribemeetings.com/filestream.ashx?DocumentId=141783

[16] B. van Duin, "Evolution of Stormwater Management in Calgary," presented at the TRIECA 2013, Mar. 26, 2013. Accessed: Apr. 23, 2024. [Online]. Available: https://sourcetostream.com/app/uploads/2016/07/EvolutionofStormwaterManagementinCalgarypresentation.pdf

[17] B. Petrie, "A review of combined sewer overflows as a source of wastewater-derived emerging contaminants in the environment and their management," *Environ. Sci. Pollut. Res.*, vol. 28, no. 25, pp. 32095–32110, Jul. 2021, doi: 10.1007/s11356-021-14103-1.

[18] H. T. Olds, S. R. Corsi, D. K. Dila, K. M. Halmo, M. J. Bootsma, and S. L. McLellan, “High levels of sewage contamination released from urban areas after storm events: A quantitative survey with sewage specific bacterial indicators,” *PLOS Med.*, vol. 15, no. 7, p. e1002614, Jul. 2018, doi: 10.1371/journal.pmed.1002614.
[19] B. Sercu, L. C. Van De Werfhorst, J. L. S. Murray, and P. A. Holden, “Sewage exfiltration as a source of storm drain contamination during dry weather in urban watersheds,” *Environ. Sci. Technol.*, vol. 45, no. 17, pp. 7151–7157, Sep. 2011, doi: 10.1021/es200981k.
[20] O. Panasiuk, A. Hedström, J. Marsalek, R. M. Ashley, and M. Viklander, “Contamination of stormwater by wastewater: A review of detection methods,” *J. Environ. Manage.*, vol. 152, pp. 241–250, Apr. 2015, doi: 10.1016/j.jenvman.2015.01.050.
[21] B. Leveque, J.-B. Burnet, S. Dorner, and F. Bichai, “Impact of climate change on the vulnerability of drinking water intakes in a northern region,” *Sustain. Cities Soc.*, vol. 66, p. 102656, Mar. 2021, doi: 10.1016/j.scs.2020.102656.
[22] N. J. Ashbolt, M. E. Schoen, J. A. Soller, and D. J. Roser, “Predicting pathogen risks to aid beach management: The real value of quantitative microbial risk assessment (QMRA),” *Water Res.*, vol. 44, no. 16, pp. 4692–4703, Sep. 2010, doi: 10.1016/j.watres.2010.06.048.
[23] T. Ahmed, M. Zounemat-Kermani, and M. Scholz, “Climate Change, Water Quality and Water-Related Challenges: A Review with Focus on Pakistan,” *Int. J. Environ. Res. Public. Health*, vol. 17, no. 22, p. 8518, Nov. 2020, doi: 10.3390/ijerph17228518.
[24] V. J. Harwood, C. Staley, B. D. Badgley, K. Borges, and A. Korajkic, “Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes,” *FEMS Microbiol. Rev.*, vol. 38, no. 1, pp. 1–40, Jan. 2014, doi: 10.1111/1574-6976.12031.
[25] K. G. Field and M. Samadpour, “Fecal source tracking, the indicator paradigm, and managing water quality,” *Water Res.*, vol. 41, no. 16, pp. 3517–3538, Aug. 2007, doi: 10.1016/j.watres.2007.06.056.
[26] C. Rock, B. Rivera, and C. P. Gerba, “Microbial Source Tracking,” in *Environmental Microbiology (Third Edition)*, I. L. Pepper, C. P. Gerba, and T. J. Gentry, Eds., San Diego: Academic Press, 2015, pp. 309–317. doi: 10.1016/B978-0-12-394626-3.00014-4.
[27] L. Paruch and A. M. Paruch, “An Overview of Microbial Source Tracking Using Host-Specific Genetic Markers to Identify Origins of Fecal Contamination in Different Water Environments,” *Water*, vol. 14, no. 11, Art. no. 11, Jan. 2022, doi: 10.3390/w14111809.
[28] Z. R. Staley, J. Grabuski, E. Sverko, and T. A. Edge, “Comparison of Microbial and Chemical Source Tracking Markers To Identify Fecal Contamination Sources in the Humber River (Toronto, Ontario, Canada) and Associated Storm Water Outfalls,” *Appl. Environ. Microbiol.*, vol. 82, no. 21, pp. 6357–6366, Nov. 2016, doi: 10.1128/AEM.01675-16.
[29] C. Hagedorn and S. B. Weisberg, “Chemical-Based Fecal Source Tracking Methods,” in *Microbial Source Tracking: Methods, Applications, and Case Studies*, C. Hagedorn, A. R. Blanch, and V. J. Harwood, Eds., New York, NY: Springer, 2011, pp. 189–206. doi: 10.1007/978-1-4419-9386-1_8.
[30] B. Gilpin, T. James, F. Nourozi, D. Saunders, P. Scholes, and M. Savill, “The use of chemical and molecular microbial indicators for faecal source identification,” *Water Sci. Technol.*, vol. 47, no. 3, pp. 39–43, Feb. 2003, doi: https://doi.org/10.2166/wst.2003.0155.
[31] M. J. Arlos *et al.*, “Combining chemical, bioanalytical and predictive tools to assess persistence, seasonality, and sporadic releases of organic micropollutants within the urban water cycle,” *Water Res.*, vol. 244, p. 120454, Oct. 2023, doi: 10.1016/j.watres.2023.120454.
[32] E. Lee, S. Lee, Y. Kim, Y.-J. Huh, K.-S. Kim, and J. Cho, “Wastewater Treatment Plant : Anthropogenic Micropollutant Indicators for Sustainable River Management,” in *Encyclopedia of Sustainability Science and Technology*, R. A. Meyers, Ed., New York, NY: Springer, 2012, pp. 11911–11932. doi: 10.1007/978-1-4419-0851-3_843.
[33] H. Guérineau *et al.*, “Source tracking of leaky sewers: A novel approach combining fecal indicators in water and sediments,” *Water Res.*, vol. 58, pp. 50–61, Jul. 2014, doi: 10.1016/j.watres.2014.03.057.

[34] Government of Alberta, “Calgary - Population.” Accessed: Feb. 23, 2024. [Online]. Available: https://regionaldashboard.alberta.ca/region/calgary/population/#/?from=2001&to=2022
[35] G. Aljoudi, “Bacterial Communities in the Glenmore Reservoir with an emphasis on the Cyanobacteria,” University of Waterloo, Waterloo, Ontario, Canada, 2021. Accessed: Feb. 22, 2024. [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/17315
[36] Water Resources and Water Services, “Calgary’s water supply,” https://www.calgary.ca. Accessed: May 21, 2024. [Online]. Available: https://www.calgary.ca/content/www/en/home/water/drinking-water/water-supply.html
[37] City of Calgary, “Watershed Monitoring Program sampling locations | Open Calgary.” Accessed: Jun. 24, 2024. [Online]. Available: https://data.calgary.ca/Environment/Watershed-Monitoring-Program-sampling-locations/h8q6-8evm
[38] Alberta Landuse Knowledge Network, “Bow River Basin State of the Watershed | Landuse.” Accessed: May 21, 2024. [Online]. Available: https://landusekn.ca/resource/bow-river-basin-state-watershed
[39] City of Calgary, “Land Use Districts | Open Calgary.” Accessed: Jun. 24, 2024. [Online]. Available: https://data.calgary.ca/Base-Maps/Land-Use-Districts/mw9j-jik5
[40] The Town of Crossfield, “Water and Wastewater Bylaw No. 2022-07.” Accessed: Jun. 24, 2024. [Online]. Available: https://crossfieldalberta.com/Home/DownloadDocument?docId=2b9e2ab5-897c-425b-9fb7-3e7d88b61597
[41] Nose Creek Watershed Partnership, “Nose Creek Map Watershed,” NCWP. Accessed: Jun. 24, 2024. [Online]. Available: https://www.nosecreekpartnership.com/watershed
[42] Government of Canada, “Census Profile, 2021 Census of Population,” Statistics Canada. Accessed: Aug. 14, 2024. [Online]. Available: https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E
[43] Government of Alberta, “Code of Practice for Wastewater Systems Using a Wastewater Lagoon.” Accessed: Aug. 07, 2024. [Online]. Available: https://open.alberta.ca/publications/lagoon
[44] J. Holstein, “Town of Crossfield Wastewater Lagoon Discussion,” Aug. 07, 2024.
[45] P. Li and J. Wu, “Drinking Water Quality and Public Health,” *Expo. Health*, vol. 11, no. 2, pp. 73–79, Jun. 2019, doi: 10.1007/s12403-019-00299-8.
[46] J.-M. Davies and A. Mazumder, “Health and environmental policy issues in Canada: the role of watershed management in sustaining clean drinking water quality at surface sources,” *J. Environ. Manage.*, vol. 68, no. 3, pp. 273–286, Jul. 2003, doi: 10.1016/S0301-4797(03)00070-7.
[47] L. Lin, H. Yang, and X. Xu, “Effects of Water Pollution on Human Health and Disease Heterogeneity: A Review,” *Front. Environ. Sci.*, vol. 10, Jun. 2022, doi: 10.3389/fenvs.2022.880246.
[48] C. Furlong, R. Brotchie, R. Considine, G. Finlayson, and L. Guthrie, “Key concepts for Integrated Urban Water Management infrastructure planning: Lessons from Melbourne,” *Util. Policy*, vol. 45, pp. 84–96, Apr. 2017, doi: 10.1016/j.jup.2017.02.004.
[49] P. Mukheibir, C. Howe, and D. Gallet, “What’s getting in the way of a ‘One Water’ approach to water services planning and management?,” *Water Aust.*, vol. 41, pp. 67–73, May 2014.
[50] J. Bolson, M. C. Sukop, M. Arabi, G. Pivo, and A. Lanier, “A Stakeholder-Science Based Approach Using the National Urban Water Innovation Network as a Test Bed for Understanding Urban Water Sustainability Challenges in the U.S.,” *Water Resour. Res.*, vol. 54, no. 5, pp. 3453–3471, 2018, doi: 10.1029/2017WR021191.
[51] National Round Table on the Environment and the Economy, *Charting a course : Sustainable water use by Canada’s natural resource sectors*. Ottawa: Canadian Electronic Library, 2011. Accessed: Jul. 19, 2024. [Online]. Available: https://canadacommons-ca.login.ezproxy.library.ualberta.ca/artifacts/1202571/charting-a-course/1755682/
[52] National Research Council, *Assigning Economic Value to Natural Resources*. Washington, DC: The National Academies Press, 1994. doi: 10.17226/4844.

[53] UNESCO World Water Assessment Programme, "The United Nations world water development report 2021: valuing water," United Nations Educational, Scientific and Cultural Organization, Paris, France, 2021. Accessed: Jul. 19, 2024. [Online]. Available: https://www.unesco.org/reports/wwdr/2021/en/download-report
[54] R. Cardone and C. Howe, "Advancing One Water in Texas," The Cynthia & George Mitchell Foundation, Texas, Feb. 2018. Accessed: Jul. 19, 2024. [Online]. Available: https://cgmf.org/graphics/cgmf_one_water_report_02_14_18_final.pdf
[55] US Water Alliance, "One Water Roadmap: The Sustainable Management of Life's Most Essential Resource." Dec. 2016. Accessed: Jul. 21, 2024. [Online]. Available: https://uswateralliance.org/resources/one-water-roadmap-the-sustainable-management-of-lifes-most-essential-resource/
[56] Health Canada, "Canadian Guidelines for Domestic Reclaimed Water for Use in Toilet and Urinal Flushing," Government of Canada. Accessed: Jul. 30, 2024. [Online]. Available: https://www.canada.ca/en/health-canada/services/publications/healthy-living/canadian-guidelines-domestic-reclaimed-water-use-toilet-urinal-flushing.html
[57] Government of Alberta, "Public health guidelines for water reuse and stormwater use." Accessed: Jul. 21, 2024. [Online]. Available: https://open.alberta.ca/publications/public-health-guidelines-water-reuse-stormwater-use
[58] A. G. Capodaglio, "Fit-for-purpose urban wastewater reuse: Analysis of issues and available technologies for sustainable multiple barrier approaches," *Crit. Rev. Environ. Sci. Technol.*, vol. 51, no. 15, pp. 1619–1666, Jun. 2021, doi: 10.1080/10643389.2020.1763231.
[59] J. Stoler *et al.*, "Modular, adaptive, and decentralised water infrastructure: promises and perils for water justice," *Curr. Opin. Environ. Sustain.*, vol. 57, p. 101202, Aug. 2022, doi: 10.1016/j.cosust.2022.101202.
[60] S. Fletcher, M. Lickley, and K. Strzepek, "Learning about climate change uncertainty enables flexible water infrastructure planning," *Nat. Commun.*, vol. 10, no. 1, p. 1782, Apr. 2019, doi: 10.1038/s41467-019-09677-x.
[61] U. Wehn, K. Collins, K. Anema, L. Basco-Carrera, and A. Lerebours, "Stakeholder engagement in water governance as social learning: lessons from practice," in *OECD Principles on Water Governance*, Routledge, 2019.
[62] Chicago Metropolitan Agency for Planning, "On To 2050." Accessed: Jul. 02, 2024. [Online]. Available: https://cmap.illinois.gov/wp-content/uploads/dlm_uploads/ON-TO-2050-Comprehensive-Regional-Plan-FINAL.pdf
[63] City of Copenhagen, "CPH 2025 Climate Plan," 2020. Accessed: Jul. 02, 2024. [Online]. Available: https://kk.sites.itera.dk/apps/kk_pub2/index.asp?mode=detalje&id=2062
[64] E. Lloyd and A. Licata, "One New York City: One Water," Water, Megacities and Global Change, New York, NY. [Online]. Available: https://www.nyc.gov/assets/dep/downloads/pdf/climate-resiliency/one-nyc-one-water.pdf
[65] Phoenix City Council, "PlanPHX," Phoenix, AZ, Mar. 2015. Accessed: Jul. 02, 2024. [Online]. Available: https://issuu.com/jbednarek/docs/pla-001_general_plan_full_print_com?e=15431004/12969882
[66] G. Dunn, R. R. Brown, J. J. Bos, and K. Bakker, "The role of science-policy interface in sustainable urban water transitions: Lessons from Rotterdam," *Environ. Sci. Policy*, vol. 73, pp. 71–79, Jul. 2017, doi: 10.1016/j.envsci.2017.04.013.
[67] H. Lee and T. P. Tan, "Singapore's experience with reclaimed water: NEWater," in *Water Reuse Policies for Potable Use*, Routledge, 2017.
[68] "Sustainable Singapore Blueprint," Singapore, 2015. Accessed: Jul. 02, 2024. [Online]. Available: https://www.clc.gov.sg/docs/default-source/books/ssbcombined-cover-text.pdf

[69] P. Mukheibir and L. Currie, “A whole of water approach for the city of Sydney,” May 2016. Accessed: Jul. 02, 2024. [Online]. Available: https://www.semanticscholar.org/paper/A-whole-of-water-approach-for-the-city-of-Sydney-Mukheibir-Currie/f3bc066e5516a7ec15c5630645764406ff9876b5
[70] City of Calgary, “One Calgary One Water.” Dec. 2019. Accessed: Jul. 02, 2024. [Online]. Available: https://pub-calgary.escribemeetings.com/filestream.ashx?DocumentId=120329
[71] City of Calgary, “One Calgary One Water A framework for Calgary’s water secure future.” Jan. 2020. Accessed: Jul. 30, 2024. [Online]. Available: https://www.calgary.ca/content/dam/www/uep/water/documents/water-documents/Water_Security_Report_2020.pdf
[72] R. Plummer, J. Velaniškis, D. de Grosbois, R. D. Kreutzwiser, and R. de Loë, “The development of new environmental policies and processes in response to a crisis: the case of the multiple barrier approach for safe drinking water,” *Environ. Sci. Policy*, vol. 13, no. 6, pp. 535–548, Oct. 2010, doi: 10.1016/j.envsci.2010.05.004.
[73] D. R. O’Connor, “Part One Report of the Walkerton Inquiry: The Events of May 2000 and Related Issues,” Ontario Ministry of the Attorney General, Jan. 2002. Accessed: Jul. 02, 2024. [Online]. Available: https://www.archives.gov.on.ca/en/e_records/walkerton/report1/index.html
[74] D. R. O’Connor, “Part Two Report of the Walkerton Inquiry: A Strategy for Safe Drinking Water,” Ontario Ministry of the Attorney General, May 2002. Accessed: Jul. 02, 2024. [Online]. Available: https://www.archives.gov.on.ca/en/e_records/walkerton/report2/index.html
[75] S. K. Ram, H. Panidepu, V. Cheernam, and R. D. Tyagi, “Multibarrier approach for the removal of organic micro-pollutants,” in *Current Developments in Biotechnology and Bioengineering*, S. Varjani, A. Pandey, R. D. Tyagi, H. H. Ngo, and C. Larroche, Eds., Elsevier, 2020, pp. 141–159. doi: 10.1016/B978-0-12-819594-9.00006-1.
[76] Federal-Provincial-Territorial Committee on Drinking Water and CCME Water Quality Task Group, “From Source to Tap: Guidance on the Multi-Barrier Approach to Safe Drinking Water.” 2004. Accessed: Jul. 02, 2024. [Online]. Available: https://ccme.ca/en/res/sourcetotap-finalenglish03-28-04.pdf
[77] T. Bereskie, I. Delpla, M. J. Rodriguez, and R. Sadiq, “Drinking-water management in Canadian provinces and territories: a review and comparison of management approaches for ensuring safe drinking water,” *Water Policy*, vol. 20, no. 3, pp. 565–596, Jun. 2018, doi: 10.2166/wp.2018.040.
[78] M. H. Dore, “Waterborne Disease Outbreaks and the Multi-barrier Approach to Protecting Drinking Water,” in *Global Drinking Water Management and Conservation: Optimal Decision-Making*, M. H. Dore, Ed., Cham: Springer International Publishing, 2015, pp. 13–31. doi: 10.1007/978-3-319-11032-5_2.
[79] N. Islam, R. Sadiq, M. J. Rodriguez, and A. Francisque, “Reviewing source water protection strategies: A conceptual model for water quality assessment,” *Environ. Rev.*, vol. 19, pp. 68–105, Mar. 2011, doi: 10.1139/a11-001.
[80] N.-B. Chang, “Hydrological Connections between Low-Impact Development, Watershed Best Management Practices, and Sustainable Development,” *J. Hydrol. Eng.*, vol. 15, no. 6, pp. 384–385, Jun. 2010, doi: https://doi.org/10.1061/(ASCE)HE.1943-5584.0000236.
[81] Y. Liu, B. A. Engel, D. C. Flanagan, M. W. Gitau, S. K. McMillan, and I. Chaubey, “A review on effectiveness of best management practices in improving hydrology and water quality: Needs and opportunities,” *Sci. Total Environ.*, vol. 601–602, pp. 580–593, Dec. 2017, doi: 10.1016/j.scitotenv.2017.05.212.
[82] World Health Organization, “Guidelines for Drinking-water Quality Third Edition Volume 1 Recommendations.” 2004. Accessed: Jul. 03, 2024. [Online]. Available: https://iris.who.int/bitstream/handle/10665/42852/9241546387.pdf
[83] J. L. Ivey, R. de Loë, R. Kreutzwiser, and C. Ferreyra, “An institutional perspective on local capacity for source water protection,” *Geoforum*, vol. 37, no. 6, pp. 944–957, Nov. 2006, doi: 10.1016/j.geoforum.2006.05.001.

[84] J. Cyr-Gagnon and M. J. Rodriguez, “Optimizing data management for municipal source water protection,” *Land Use Policy*, vol. 100, p. 103788, Jan. 2021, doi: 10.1016/j.landusepol.2018.12.035.
[85] T. Zlender and M. Rupnik, “An overview of molecular markers for identification of non-human fecal pollution sources,” *Front. Microbiol.*, vol. 14, Nov. 2023, doi: 10.3389/fmicb.2023.1256174.
[86] D. Stoeckel and V. J. Harwood, “Performance, Design, and Analysis in Microbial Source Tracking Studies,” *Appl. Environ. Microbiol.*, vol. 73, no. 8, pp. 2405–15, Jan. 2007, doi: https://doi.org/10.1128/AEM.02473-06.
[87] C. Hagedorn, A. R. Blanch, and V. J. Harwood, Eds., *Microbial Source Tracking: Methods, Applications, and Case Studies*. New York, NY: Springer, 2011. doi: 10.1007/978-1-4419-9386-1.
[88] E. E. Geldreich and W. Litsky, “Fecal coliform and fecal streptococcus density relationships in waste discharges and receiving waters,” *C R C Crit. Rev. Environ. Control*, vol. 6, no. 4, pp. 349–369, Oct. 1976, doi: 10.1080/10643387609381645.
[89] C. Tartera, F. Lucena, and J. Jofre, “Human origin of Bacteroides fragilis bacteriophages present in the environment,” *Appl. Environ. Microbiol.*, vol. 55, no. 10, pp. 2696–2701, Oct. 1989, doi: 10.1128/aem.55.10.2696-2701.1989.
[90] B. A. Wiggins, “Discriminant analysis of antibiotic resistance patterns in fecal streptococci, a method to differentiate human and animal sources of fecal pollution in natural waters,” *Appl. Environ. Microbiol.*, vol. 62, no. 11, pp. 3997–4002, Nov. 1996, doi: 10.1128/aem.62.11.3997-4002.1996.
[91] A. E. Bernhard and K. G. Field, “Identification of Nonpoint Sources of Fecal Pollution in Coastal Waters by Using Host-Specific 16S Ribosomal DNA Genetic Markers from Fecal Anaerobes,” *Microb. Ecol.*, vol. 66, no. 4, pp. 1587–1594, Apr. 2000, doi: 10.1128/AEM.66.4.1587-1594.2000.
[92] C. L. Meays, K. Broersma, R. Nordin, and A. Mazumder, “Source tracking fecal bacteria in water: a critical review of current methods,” *J. Environ. Manage.*, vol. 73, no. 1, pp. 71–79, Oct. 2004, doi: 10.1016/j.jenvman.2004.06.001.
[93] T. A. Edge and A. B. Boehm, “Classical and Molecular Methods to Measure Fecal Bacteria,” in *The Fecal Bacteria*, John Wiley & Sons, Ltd, 2010, pp. 241–273. doi: 10.1128/9781555816865.ch11.
[94] H. Šubrtová Salmonová and V. Bunesova, “Methods of Studying Diversity of Bacterial Comunities: A Review,” *Sci. Agric. Bohem.*, vol. 48, Sep. 2017, doi: 10.1515/sab-2017-0022.
[95] US EPA, “Using Microbial Source Tracking to Support TMDL Development and Implementation.” Accessed: Jul. 05, 2024. [Online]. Available: https://www.epa.gov/tmdl/using-microbial-source-tracking-support-tmdl-development-and-implementation
[96] C. Zapka *et al.*, “Comparison of Standard Culture-Based Method to Culture-Independent Method for Evaluation of Hygiene Effects on the Hand Microbiome,” *mBio*, vol. 8, no. 2, pp. e00093-17, Mar. 2017, doi: 10.1128/mBio.00093-17.
[97] T. Unno, C. Staley, C. M. Brown, D. Han, M. J. Sadowsky, and H.-G. Hur, “Fecal pollution: new trends and challenges in microbial source tracking using next-generation sequencing,” *Environ. Microbiol.*, vol. 20, no. 9, pp. 3132–3140, 2018, doi: 10.1111/1462-2920.14281.
[98] J. M. Simpson, J. W. Santo Domingo, and D. J. Reasoner, “Microbial Source Tracking: State of the Science,” *Environ. Sci. Technol.*, vol. 36, no. 24, pp. 5279–5288, Dec. 2002, doi: 10.1021/es026000b.
[99] W. Ahmed, “Limitations of library-dependent microbial source tracking methods,” *Water J. Aust. Water Assoc.*, vol. 34, no. 1, Art. no. 1, 2007.
[100] M. Soule, E. Kuhn, F. Loge, J. Gay, and D. R. Call, “Using DNA Microarrays To Identify Library-Independent Markers for Bacterial Source Tracking,” *Appl. Environ. Microbiol.*, vol. 72, no. 3, pp. 1843–1851, Mar. 2006, doi: 10.1128/AEM.72.3.1843-1851.2006.
[101] Service Canada, “Guidelines for Canadian recreational water quality technical document - Microbiological sampling and analysis,” Government of Canada. Accessed: Jul. 10, 2024. [Online]. Available: https://www.canada.ca/en/services/health/publications/healthy-living/guidelines-canadian-recreational-water-quality-technical-document-microbiological-sampling-analysis.html
[102] Health Canada, “Publications – Healthy living,” Government of Canada. Accessed: Jul. 10, 2024. [Online]. Available: https://www.canada.ca/en/services/health/publications/healthy-living.html

[103] ORD Science Inventory, “Microbial Source Tracking Guide Document.” US EPA, Jun. 2005. Accessed: Jul. 05, 2024. [Online]. Available: https://cfpub.epa.gov/si/si_public_record_Report.cfm?Lab=NRMRL&dirEntryID=133523
[104] S. Seurinck, W. Verstraete, and S. D. Siciliano, “Microbial Source Tracking for Identification of Fecal Pollution,” *Rev. Environ. Sci. Biotechnol.*, vol. 4, no. 1, pp. 19–37, May 2005, doi: 10.1007/s11157-005-4997-7.
[105] C. Hagedorn and S. B. Weisberg, “Chemical-based fecal source tracking methods: current status and guidelines for evaluation,” *Rev. Environ. Sci. Biotechnol.*, vol. 8, no. 3, pp. 275–287, Sep. 2009, doi: 10.1007/s11157-009-9162-2.
[106] J. P. S. Sidhu *et al.*, “Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers,” *Sci. Total Environ.*, vol. 463–464, pp. 488–496, Oct. 2013, doi: 10.1016/j.scitotenv.2013.06.020.
[107] J. O. Osuoha, B. O. Anyanwu, and C. Ejileugha, “Pharmaceuticals and personal care products as emerging contaminants: Need for combined treatment strategy,” *J. Hazard. Mater. Adv.*, vol. 9, p. 100206, Feb. 2023, doi: 10.1016/j.hazadv.2022.100206.
[108] A. Küster and N. Adler, “Pharmaceuticals in the environment: scientific evidence of risks and its regulation,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 369, no. 1656, p. 20130587, Nov. 2014, doi: 10.1098/rstb.2013.0587.
[109] S.-K. Kim and Q. V. Ta, “Potential Beneficial Effects of Marine Algal Sterols on Human Health,” in *Advances in Food and Nutrition Research*, vol. 63, 2011, pp. 191–198. Accessed: Jul. 08, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780123876690000144
[110] C. Juste and P. Gérard, “Cholesterol-to-Coprostanol Conversion by the Gut Microbiota: What We Know, Suspect, and Ignore,” *Microorganisms*, vol. 9, no. 9, p. 1881, Sep. 2021, doi: 10.3390/microorganisms9091881.
[111] R. Leeming, A. Ball, N. Ashbolt, and P. Nichols, “Using faecal sterols from humans and animals to distinguish faecal pollution in receiving waters,” *Water Res.*, vol. 30, no. 12, pp. 2893–2900, Dec. 1996, doi: 10.1016/S0043-1354(96)00011-5.
[112] I. Macdonald, V. Bokkenheuser, J. Winter, A. McLernon, and E. Mosbach, “Degradation of steroids in the human gut,” *J. Lipid Res.*, vol. 24, pp. 675–700, Jul. 1983, doi: 10.1016/S0022-2275(20)37944-X.
[113] I. D. Bull, M. J. Lockheart, M. M. Elhmmali, D. J. Roberts, and R. P. Evershed, “The origin of faeces by means of biomarker detection,” *Environ. Int.*, vol. 27, no. 8, pp. 647–654, Mar. 2002, doi: 10.1016/s0160-4120(01)00124-6.
[114] E. S. Reichwaldt, W. Y. Ho, W. Zhou, and A. Ghadouani, “Sterols indicate water quality and wastewater treatment efficiency,” *Water Res.*, vol. 108, pp. 401–411, Jan. 2017, doi: 10.1016/j.watres.2016.11.029.
[115] M. E. Tavares, M. I. H. Spivey, M. R. McIver, and M. Mallin, “Testing for Optical Brighteners and Fecal Bacteria to Detect Sewage Leaks in Tidal Creeks,” *J. N. C. Acad. Sci.*, vol. 124, no. 3, pp. 91–97, 2008.
[116] M. Devane, D. Saunders, and B. Gilpin, “Devane M, Saunders D, Gilpin B. Faecal sterols and fluorescent whiteners as indicators of the source of faecal contamination. Chemistry in New Zealand 2006; 70: 74-77.,” *Chem. N. Z.*, vol. 70, pp. 74–77, Jan. 2006.
[117] S. Li, J. Wen, B. He, J. Wang, X. Hu, and J. Liu, “Occurrence of caffeine in the freshwater environment: Implications for ecopharmacovigilance,” *Environ. Pollut.*, vol. 263, p. 114371, Aug. 2020, doi: 10.1016/j.envpol.2020.114371.
[118] R. L. Seiler, S. D. Zaugg, J. M. Thomas, and D. L. Howcroft, “Caffeine and Pharmaceuticals as Indicators of Waste Water Contamination in Wells,” *Groundwater*, vol. 37, no. 3, pp. 405–410, 1999, doi: 10.1111/j.1745-6584.1999.tb01118.x.
[119] R. Siegener and R. F. Chen, “Caffeine in Boston Harbor seawater,” *Mar. Pollut. Bull.*, vol. 44, no. 5, pp. 383–387, May 2002, doi: 10.1016/s0025-326x(00)00176-4.

[120] N. A. I. Tawfik, Z. A. El-Bakary, and K. F. Abd El-Wakeil, “Determination of caffeine in treated wastewater discharged in the Nile River with emphasis on the effect of zinc and physicochemical factors,” *Environ. Sci. Pollut. Res.*, vol. 31, no. 19, pp. 28124–28138, Apr. 2024, doi: 10.1007/s11356-024-32918-6.
[121] M. T. Pham, T. T. H. Chu, and D. C. Vu, “Mitigation of caffeine micropollutants in wastewater through Ag-doped ZnO photocatalyst: mechanism and environmental impacts,” *Environ. Geochem. Health*, vol. 46, no. 5, p. 168, Apr. 2024, doi: 10.1007/s10653-024-01952-1.
[122] N. Z. Dvory *et al.*, “Caffeine vs. carbamazepine as indicators of wastewater pollution in a karst aquifer,” *Hydrol. Earth Syst. Sci.*, vol. 22, no. 12, pp. 6371–6381, Dec. 2018, doi: 10.5194/hess-22-6371-2018.
[123] T. S. Bianchi and E. A. Canuel, *Chemical Biomarkers in Aquatic Ecosystems*. Princeton, NJ: Princeton University Press, 2011. Accessed: Jul. 09, 2024. [Online]. Available: https://web-p-ebscohost-com.login.ezproxy.library.ualberta.ca/ehost/ebookviewer/ebook/ZTAwMHhuYV9fMzU2NDgyX19BTg2?sid=ce93a1e5-c498-4751-91f3-7260c2ae136e@redis&vid=0&format=EB&rid=1
[124] J. L. Wilkinson, J. Swinden, P. S. Hooda, J. Barker, and S. Barton, “Markers of anthropogenic contamination: A validated method for quantification of pharmaceuticals, illicit drug metabolites, perfluorinated compounds, and plasticisers in sewage treatment effluent and rain runoff,” *Chemosphere*, vol. 159, pp. 638–646, Sep. 2016, doi: 10.1016/j.chemosphere.2016.06.039.
[125] M. Scheurer *et al.*, “Correlation of six anthropogenic markers in wastewater, surface water, bank filtrate, and soil aquifer treatment,” *J. Environ. Monit.*, vol. 13, no. 4, pp. 966–973, Apr. 2011, doi: 10.1039/C0EM00701C.
[126] International Commission for the Protection of the Danube River, “Chemical and immunochemical analysis of anthropogenic markers and organic contaminants,” Vienna, Austria, 2014. Accessed: Jul. 09, 2024. [Online]. Available: https://www.danubesurvey.org/jds3/jds3-files/nodes/documents/anthropogenic_markers.pdf
[127] F. Pinongcos *et al.*, “Chemical and Microbial Markers for Discriminating Sanitary Sewer Contamination in Coastal, Urban Streams,” *ACS EST Water*, vol. 2, no. 10, pp. 1747–1759, Oct. 2022, doi: 10.1021/acsestwater.2c00265.
[128] A. R. Blanch *et al.*, “Integrated Analysis of Established and Novel Microbial and Chemical Methods for Microbial Source Tracking,” *Appl. Environ. Microbiol.*, vol. 72, no. 9, pp. 5915–5926, Sep. 2006, doi: 10.1128/AEM.02453-05.
[129] S. K. Haack *et al.*, “Comparing Wastewater Chemicals, Indicator Bacteria Concentrations, and Bacterial Pathogen Genes as Fecal Pollution Indicators,” *J. Environ. Qual.*, vol. 38, no. 1, pp. 248–258, 2009, doi: 10.2134/jeq2008.0173.
[130] L. W. Sinton, R. K. Finlay, and D. J. Hannah, “Distinguishing human from animal faecal contamination in water: A review,” *N. Z. J. Mar. Freshw. Res.*, vol. 32, no. 2, pp. 323–348, Jun. 1998, doi: 10.1080/00288330.1998.9516828.
[131] S. T. Glassmeyer *et al.*, “Transport of Chemical and Microbial Compounds from Known Wastewater Discharges: Potential for Use as Indicators of Human Fecal Contamination,” *Environ. Sci. Technol.*, vol. 39, no. 14, pp. 5157–5169, Jul. 2005, doi: 10.1021/es048120k.
[132] J. I. Oragui and D. D. Mara, “A Selective Medium for the Enumeration of Streptococcus bovis by Membrane Filtration,” *J. Appl. Bacteriol.*, vol. 51, no. 1, pp. 85–93, Aug. 1981, doi: 10.1111/j.1365-2672.1981.tb00911.x.
[133] S. Parveen, R. L. Murphree, L. Edmiston, C. W. Kaspar, K. M. Portier, and M. L. Tamplin, “Association of multiple-antibiotic-resistance profiles with point and nonpoint sources of Escherichia coli in Apalachicola Bay,” *Appl. Environ. Microbiol.*, vol. 63, no. 7, pp. 2607–2612, Apr. 1997, doi: 10.1128/aem.63.7.2607-2612.1997.

[134] C. Hagedorn, S. L. Robinson, J. R. Filtz, S. M. Grubbs, T. A. Angier, and R. B. Reneau Jr, "Determining Sources of Fecal Pollution in a Rural Virginia Watershed with Antibiotic Resistance Patterns in Fecal Streptococci," *Environ. Public Health Microbiol.*, vol. 65, no. 12, pp. 5522–5531, Dec. 1999, doi: 10.1128/aem.65.12.5522-5531.1999.
[135] A. Musolff, S. Leschik, F. Reinstorf, G. Strauch, and M. Schirmer, "Micropollutant Loads in the Urban Water Cycle," *Environ. Sci. Technol.*, vol. 44, no. 13, pp. 4877–4883, Jul. 2010, doi: 10.1021/es903823a.
[136] E. Jardé *et al.*, "Application of a microbial source tracking based on bacterial and chemical markers in headwater and coastal catchments," *Sci. Total Environ.*, vol. 610–611, pp. 55–63, Jan. 2018, doi: 10.1016/j.scitotenv.2017.07.235.
[137] M. L. Devane *et al.*, "Relationships between chemical and microbial faecal source tracking markers in urban river water and sediments during and post-discharge of human sewage," *Sci. Total Environ.*, vol. 651, pp. 1588–1604, Feb. 2019, doi: 10.1016/j.scitotenv.2018.09.258.
[138] L. Paruch and A. M. Paruch, "Cross-tracking of faecal pollution origins, macronutrients, pharmaceuticals and personal care products in rural and urban watercourses," *Water Sci. Technol.*, vol. 83, no. 3, pp. 610–621, Dec. 2020, doi: 10.2166/wst.2020.603.
[139] M. Waso, T. Ndlovu, P. H. Dobrowsky, S. Khan, and W. Khan, "Presence of microbial and chemical source tracking markers in roof-harvested rainwater and catchment systems for the detection of fecal contamination," *Environ. Sci. Pollut. Res.*, vol. 23, no. 17, pp. 16987–17001, Sep. 2016, doi: 10.1007/s11356-016-6895-7.
[140] K. N. Tyre, R. A. Brewton, L. B. Kreiger, and B. E. Lapointe, "Widespread human waste pollution in surface waters observed throughout the urbanized, coastal communities of Lee County, Florida, USA," *Sci. Total Environ.*, vol. 879, p. 162716, Jun. 2023, doi: 10.1016/j.scitotenv.2023.162716.
[141] M. Hachad *et al.*, "β-D-glucuronidase activity triggered monitoring of fecal contamination using microbial and chemical source tracking markers at drinking water intakes," *Water Res.*, vol. 254, p. 121374, May 2024, doi: 10.1016/j.watres.2024.121374.
[142] M. Hajj-Mohamad *et al.*, "Fecal contamination of storm sewers: Evaluating wastewater micropollutants, human-specific *Bacteroides* 16S rRNA, and mitochondrial DNA genetic markers as alternative indicators of sewer cross connections," *Sci. Total Environ.*, vol. 659, pp. 548–560, Apr. 2019, doi: 10.1016/j.scitotenv.2018.12.378.
[143] S. Richards, P. J. A. Withers, E. Paterson, C. W. McRoberts, and M. Stutter, "Potential tracers for tracking septic tank effluent discharges in watercourses," *Environ. Pollut.*, vol. 228, pp. 245–255, Sep. 2017, doi: 10.1016/j.envpol.2017.05.044.
[144] A.-S. Madoux-Humery *et al.*, "Temporal variability of combined sewer overflow contaminants: Evaluation of wastewater micropollutants as tracers of fecal contamination," *Water Res.*, vol. 47, no. 13, pp. 4370–4382, Sep. 2013, doi: 10.1016/j.watres.2013.04.030.
[145] K. Pongmala *et al.*, "Modelling total suspended solids, *E. coli* and carbamazepine, a tracer of wastewater contamination from combined sewer overflows," *J. Hydrol.*, vol. 531, pp. 830–839, Dec. 2015, doi: 10.1016/j.jhydrol.2015.10.042.
[146] A. C. Cabral, J. S. Stark, H. E. Kolm, and C. C. Martins, "An integrated evaluation of some faecal indicator bacteria (FIB) and chemical markers as potential tools for monitoring sewage contamination in subtropical estuaries," *Environ. Pollut.*, vol. 235, pp. 739–749, Apr. 2018, doi: 10.1016/j.envpol.2017.12.109.
[147] P. L. Lenaker *et al.*, "A multi-marker assessment of sewage contamination in streams using human-associated indicator bacteria, human-specific viruses, and pharmaceuticals," *Sci. Total Environ.*, vol. 930, p. 172505, Jun. 2024, doi: 10.1016/j.scitotenv.2024.172505.
[148] "Joint Danube Survey 3: A Comprehensive Analysis of Danube Water Quality," International Commission for the Protection of the Danube River, Vienna, Austria, 2015. Accessed: Jul. 17, 2024. [Online]. Available: https://www.danubesurvey.org/jds3/results

[149] K. J. Miller and J. Meek, “Helena Valley Ground Water: Pharmaceuticals, Personal Care Products, Endocrine Disruptors (PPCPs) and Microbial Indicators of Fecal Contamination.” Montana Department of Environmental Quality, Mar. 2006. Accessed: Jul. 17, 2024. [Online]. Available: https://www.regulations.gov/document/EPA-HQ-OW-2002-0061-0125
[150] M. Serville-Tertullien *et al.*, “Sources of microbial contamination in the watershed and coastal zone of Soufriere, St. Lucia,” *Environ. Monit. Assess.*, vol. 194, no. 3, p. 225, Feb. 2022, doi: 10.1007/s10661-022-09862-4.
[151] D. Roser *et al.*, “Application of physico-chemical parameters and particle-bound biomarkers to indicate microbial contamination of aquifers,” *Water Sci. Technol. Water Supply*, vol. 2, no. 5–6, pp. 419–426, Dec. 2002.
[152] M. Digaletos, C. J. Ptacek, J. Thomas, and Y. Liu, “Chemical and biological tracers to identify source and transport pathways of septic system contamination to streams in areas with low permeability soils,” *Sci. Total Environ.*, vol. 870, p. 161866, Apr. 2023, doi: 10.1016/j.scitotenv.2023.161866.
[153] L. C. Van De Werfhorst, J. L. S. Murray, S. Reynolds, K. Reynolds, and P. A. Holden, “Canine Scent Detection and Microbial Source Tracking of Human Waste Contamination in Storm Drains,” *Water Environ. Res.*, vol. 86, no. 6, pp. 550–558, 2014, doi: 10.2175/106143013X13807328848496.
[154] L. Paruch, A. M. Paruch, A.-G. B. Blankenberg, K. Haarstad, and T. Mæhlum, “Norwegian study on microbial source tracking for water quality control and pollution removal in constructed wetland treating catchment run-off,” *Water Sci. Technol.*, vol. 76, no. 5, pp. 1158–1166, Sep. 2017, doi: 10.2166/wst.2017.303.
[155] T. A. Edge, R. J. Boyd, P. Shum, and J. L. Thomas, “Microbial source tracking to identify fecal sources contaminating the Toronto Harbour and Don River watershed in wet and dry weather,” *J. Gt. Lakes Res.*, vol. 47, no. 2, pp. 366–377, Apr. 2021, doi: 10.1016/j.jglr.2020.09.002.
[156] M. Hachad *et al.*, “Locating illicit discharges in storm sewers in urban areas using multi-parameter source tracking: Field validation of a toolbox composite index to prioritize high risk areas,” *Sci. Total Environ.*, vol. 811, p. 152060, Mar. 2022, doi: 10.1016/j.scitotenv.2021.152060.
[157] C. J. Berg, J. P. Alderete, and E. A. Alderete, “Human wastewater tracking in tropical Hawaiian island streams using qualitative and quantitative assessments of combined fecal indicating bacteria and sucralose, an organic micropollutant of emerging concern,” *Environ. Monit. Assess.*, vol. 195, no. 8, p. 966, Jul. 2023, doi: 10.1007/s10661-023-11545-7.
[158] S. Tolouei *et al.*, “Temporal variability of parasites, bacterial indicators, and wastewater micropollutants in a water resource recovery facility under various weather conditions,” *Water Res.*, vol. 148, pp. 446–458, Jan. 2019, doi: 10.1016/j.watres.2018.10.068.
[159] US EPA, “Method 1611: Enterococci in Water by TaqMan.” Oct. 2012. Accessed: Aug. 01, 2024. [Online]. Available: https://www.epa.gov/sites/default/files/2015-08/documents/method_1611_2012.pdf
[160] Office of Water, “Evaluation of the Suitability of Individual Combinations of Indicators and Methods for Different Clean Water Act Programs.” US EPA, Dec. 2010. Accessed: Aug. 01, 2024. [Online]. Available: https://www.epa.gov/sites/default/files/2015-11/documents/evaluation-suitability-cwa-programs.pdf
[161] M. Ranjan *et al.*, “Python: Empowering Data Science Applications and Research,” *J. Oper. Syst. Dev. Trends*, vol. 10, pp. 27–33, Aug. 2023, doi: 10.37591/joosdt.v10i1.576.
[162] T. P. D. Team, *pandas: Powerful data structures for data analysis, time series, and statistics*. Cython, Python. Accessed: May 22, 2024. [OS Independent]. Available: https://pandas.pydata.org
[163] M. Waskom, “seaborn: statistical data visualization,” *J. Open Source Softw.*, vol. 6, no. 60, p. 3021, Apr. 2021, doi: 10.21105/joss.03021.
[164] J. D. H. Droettboom Michael, *matplotlib: Python plotting package*. Python. Accessed: May 22, 2024. [Online]. Available: https://matplotlib.org
[165] “os — Miscellaneous operating system interfaces — Python 3.12.3 documentation.” Accessed: May 22, 2024. [Online]. Available: https://docs.python.org/3/library/os.html

[166] B. Fremaux, T. Boa, and C. K. Yost, “Quantitative Real-Time PCR Assays for Sensitive Detection of Canada Goose-Specific Fecal Pollution in Water Sources,” *Appl. Environ. Microbiol.*, vol. 76, no. 14, pp. 4886–4889, Jul. 2010, doi: 10.1128/AEM.00110-10.
[167] S. Zhi, G. Banting, and N. F. Neumann, “Development of a qPCR assay for the detection of naturalized wastewater E. coli strains,” *J. Water Health*, vol. 20, no. 4, pp. 727–736, Apr. 2022, doi: 10.2166/wh.2022.014.
[168] M. Beaudry, “From Nuisance to Resource: Understanding Microbial Sources of Contamination in Urban Stormwater-Impacted Bodies of Water Intended for Water Reuse Activities,” University of Alberta, 2019. doi: 10.7939/r3-0pqf-5r96.
[169] O. C. Shanks *et al.*, “Performance Assessment PCR-Based Assays Targeting Bacteroidales Genetic Markers of Bovine Fecal Pollution,” *Appl. Environ. Microbiol.*, vol. 76, no. 5, pp. 1359–1366, Mar. 2010, doi: 10.1128/AEM.02033-09.
[170] L. E. Brooks and K. G. Field, “Global model fitting to compare survival curves for faecal indicator bacteria and ruminant-associated genetic markers,” *J. Appl. Microbiol.*, vol. 122, no. 6, pp. 1704–1713, Jun. 2017, doi: 10.1111/jam.13454.
[171] H. C. Green, K. M. White, C. A. Kelty, and O. C. Shanks, “Development of Rapid Canine Fecal Source Identification PCR-Based Assays,” *Environ. Sci. Technol.*, vol. 48, no. 19, pp. 11453–11461, Oct. 2014, doi: 10.1021/es502637b.
[172] C. Lee, J. W. Marion, and J. Lee, “Development and application of a quantitative PCR assay targeting *Catellicoccus marimammalium* for assessing gull-associated fecal contamination at Lake Erie beaches,” *Sci. Total Environ.*, vol. 454–455, pp. 1–8, Jun. 2013, doi: 10.1016/j.scitotenv.2013.03.003.
[173] D. R. Helsel, “Fabricating data: How substituting values for nondetects can ruin results, and what can be done about it,” *Chemosphere*, vol. 65, no. 11, pp. 2434–2439, Dec. 2006, doi: 10.1016/j.chemosphere.2006.04.051.
[174] B. Cantoni, R. Delli Compagni, A. Turolla, I. Epifani, and M. Antonelli, “A statistical assessment of micropollutants occurrence, time trend, fate and human health risk using left-censored water quality data,” *Chemosphere*, vol. 257, p. 127095, Oct. 2020, doi: 10.1016/j.chemosphere.2020.127095.
[175] L. Lee and D. Helsel, “Statistical analysis of water-quality data containing multiple detection limits II: S-language software for nonparametric distribution modeling and hypothesis testing,” *Comput. Geosci.*, vol. 33, no. 5, pp. 696–704, May 2007, doi: 10.1016/j.cageo.2006.09.006.
[176] S. M. Abdel-Megeed, “Accuracy of Correlation Coefficient with Limited Number of Points,” *J. Exp. Educ.*, vol. 52, no. 4, pp. 188–191, 1984.
[177] R. Dawson, “How Significant is a Boxplot Outlier?,” *J. Stat. Educ.*, vol. 19, no. 2, 2011, doi: https://doi.org/10.1080/10691898.2011.11889610.
[178] Y. Dodge, “Box Plot,” in *The Concise Encyclopedia of Statistics*, New York, NY: Springer, 2008, pp. 55–57. doi: 10.1007/978-0-387-32833-1_43.
[179] G. Hatem, J. Zeidan, M. Goossens, and C. Moreira, “Normality Testing Methods and the Importance of Skewness and Kurtosis in Statistical Analysis,” *BAU J. - Sci. Technol.*, vol. 3, no. 2, Jun. 2022, doi: 10.54729/KTPE9512.
[180] M. Hubert and E. Vandervieren, “An adjusted boxplot for skewed distributions,” *Comput. Stat. Data Anal.*, vol. 52, no. 12, pp. 5186–5201, Aug. 2008, doi: 10.1016/j.csda.2007.11.008.
[181] C. Choonpradub and D. McNeil, “Can the box plot be improved?,” *Songklanakarin J. Sci. Technol. SJST*, vol. 27, no. 3, pp. 649–657, May 2005.
[182] D. Bohlender, “When do the seasons start?,” Government of Canada. Accessed: Jun. 28, 2024. [Online]. Available: https://nrc.canada.ca/en/certifications-evaluations-standards/canadas-official-time/3-when-do-seasons-start
[183] B. W. Yap and C. H. Sim, “Comparisons of various types of normality tests,” *J. Stat. Comput. Simul.*, vol. 81, no. 12, pp. 2141–2155, Dec. 2011, doi: 10.1080/00949655.2010.520163.

[184] P. Mishra, C. M. Pandey, U. Singh, A. Gupta, C. Sahu, and A. Keshri, “Descriptive Statistics and Normality Tests for Statistical Data,” *Ann. Card. Anaesth.*, vol. 22, no. 1, pp. 67–72, 2019, doi: 10.4103/aca.ACA_157_18.
[185] A. Ghasemi and S. Zahediasl, “Normality Tests for Statistical Analysis: A Guide for Non-Statisticians,” *Int. J. Endocrinol. Metab.*, vol. 10, no. 2, pp. 486–489, 2012, doi: 10.5812/ijem.3505.
[186] S. S. Shapiro and M. B. Wilk, “An Analysis of Variance Test for Normality (Complete Samples),” *Biometrika*, vol. 52, no. 3/4, pp. 591–611, 1965, doi: 10.2307/2333709.
[187] A. P. King and R. J. Eckersley, “Inferential Statistics IV: Choosing a Hypothesis Test,” in *Statistics for Biomedical Engineers and Scientists*, A. P. King and R. J. Eckersley, Eds., Academic Press, 2019, pp. 147–171. doi: 10.1016/B978-0-08-102939-8.00016-5.
[188] D. Melbourne, “A New method for Testing Normality based upon a Characterization of the Normal Distribution,” University Graduate School at FIU Digital Commons, Miami, Fl, 2014. Accessed: Jun. 27, 2024. [Online]. Available: https://digitalcommons.fiu.edu/cgi/viewcontent.cgi?article=2288&context=etd
[189] R. D’Agostino and E. Pearson, “Tests for Departure from Normality. Empirical Results for the Distributions of b2 and √ b1,” *Biometrika*, vol. 60, no. 3, pp. 613–622, Dec. 1973, doi: 10.2307/2335012.
[190] T. Thadewald and H. Buning, “Jarque-Bera test and its competitors for testing normality - A power comparison,” *J. Appl. Stat.*, vol. 34, no. 1, pp. 87–105, Jan. 2007, doi: 10.1080/02664760600994539.
[191] C. M. Jarque and A. K. Bera, “Efficient tests for normality, homoscedasticity and serial independence of regression residuals,” *Econ. Lett.*, vol. 6, no. 3, pp. 255–259, Jan. 1980, doi: 10.1016/0165-1765(80)90024-5.
[192] K. Das, “A Brief Review of Tests for Normality,” *Am. J. Theor. Appl. Stat.*, vol. 5, p. 5, Jan. 2016, doi: 10.11648/j.ajtas.20160501.12.
[193] M. Ahmad and A. Al-Mutairi, “Normality Tests: A Brief Review,” *J. ISOSS*, vol. 3, no. 1, pp. 35–44, 2017.
[194] G. Di Leo and F. Sardanelli, “Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach,” *Eur. Radiol. Exp.*, vol. 4, no. 1, p. 18, Mar. 2020, doi: 10.1186/s41747-020-0145-y.
[195] A. Hayter, *Probability and Statistics for Engineers and Scientists*, Statistics 235, Custom Edition. Alberta: Nelson Education, 2013.
[196] R. Aggarwal and P. Ranganathan, “Common pitfalls in statistical analysis: The use of correlation techniques,” *Perspect. Clin. Res.*, vol. 7, no. 4, pp. 187–190, 2016, doi: 10.4103/2229-3485.192046.
[197] V. Bewick, L. Cheek, and J. Ball, “Statistics review 7: Correlation and regression,” *Crit. Care*, vol. 7, no. 6, pp. 451–459, 2003.
[198] N. S. Chok, “Pearson’s Versus Spearman’s and Kendall’s Correlation Coefficients for Continuous Data.” Accessed: Jun. 27, 2024. [Online]. Available: https://d-scholarship.pitt.edu/8056/
[199] A. Barakat, M. El Baghdadi, J. Rais, B. Aghezzaf, and M. Slassi, “Assessment of spatial and seasonal water quality variation of Oum Er Rbia River (Morocco) using multivariate statistical techniques,” *Int. Soil Water Conserv. Res.*, vol. 4, no. 4, pp. 284–292, Dec. 2016, doi: 10.1016/j.iswcr.2016.11.002.
[200] S. N. Shamsudin, M. H. F. Rahman, M. N. Taib, W. R. W. A. Razak, A. H. Ahmad, and M. M. Zain, “Analysis between Escherichia Coli growth and physical parameters in water using Pearson correlation,” in *2016 7th IEEE Control and System Graduate Research Colloquium, ICSGRC 2016 - Proceeding*, Institute of Electrical and Electronics Engineers Inc., 2017. Accessed: Jun. 27, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7813315
[201] F. Zhou, Y. Liu, and H. Guo, “Application of Multivariate Statistical Methods to Water Quality Assessment of the Watercourses in Northwestern New Territories, Hong Kong,” *Environ. Monit. Assess.*, vol. 132, no. 1, pp. 1–13, Sep. 2007, doi: 10.1007/s10661-006-9497-x.

[202] V. Z. Antonopoulos, D. M. Papamichail, and K. A. Mitsiou, “Statistical and trend analysis of water quality and quantity data for the Strymon River in Greece,” *Hydrol. Earth Syst. Sci.*, vol. 5, no. 4, pp. 679–692, Dec. 2001, doi: 10.5194/hess-5-679-2001.
[203] J. Tu, Z.-G. Xia, K. C. Clarke, and A. Frei, “Impact of Urban Sprawl on Water Quality in Eastern Massachusetts, USA,” *Environ. Manage.*, vol. 40, no. 2, pp. 183–200, Aug. 2007, doi: 10.1007/s00267-006-0097-x.
[204] A. G. Asuero, A. Sayago, and A. G. González, “The Correlation Coefficient: An Overview,” *Crit. Rev. Anal. Chem.*, vol. 36, no. 1, pp. 41–59, Mar. 2006, doi: 10.1080/10408340500526766.
[205] B. Ratner, “The correlation coefficient: Its values range between +1/−1, or do they?,” *J. Target. Meas. Anal. Mark.*, vol. 17, no. 2, pp. 139–142, Jun. 2009, doi: 10.1057/jt.2009.5.
[206] Water Quality Australia, “Correlation between water quality variables,” Commonwealth of Australia. Accessed: Aug. 08, 2024. [Online]. Available: https://www.waterquality.gov.au/anz-guidelines/monitoring/data-analysis/correlation-between-variables
[207] H. Li *et al.*, “Index Correlation Analysis in Water Quality Monitoring Big Data,” *J. Phys. Conf. Ser.*, vol. 2504, no. 1, p. 012059, May 2023, doi: 10.1088/1742-6596/2504/1/012059.
[208] *Cluster Analysis*, 5th ed. John Wiley & Sons, Ltd, 2011. doi: 10.1002/9780470977811.
[209] P. C. Wong, “Unsupervised Machine Learning,” in *Computational Psychometrics: New Methodologies for a New Generation of Digital Learning and Assessment: With Examples in R and Python*, A. A. von Davier, R. J. Mislevy, and J. Hao, Eds., Cham: Springer International Publishing, 2021, pp. 173–193. doi: 10.1007/978-3-030-74394-9_10.
[210] V. Chahar, J. Chhabra, and D. Kumar, “Impact of Distance Measures on the Performance of Clustering Algorithms,” in *Advances in Intelligent Systems and Computing*, vol. 243, 2014, pp. 183–190. doi: 10.1007/978-81-322-1665-0_17.
[211] J. Chen, A. M. MacEachren, and D. J. Peuquet, “Constructing Overview + Detail Dendrogram-Matrix Views,” *IEEE Trans. Vis. Comput. Graph.*, vol. 15, no. 6, pp. 889–896, Nov. 2009, doi: 10.1109/TVCG.2009.130.
[212] F. Nielsen, “Hierarchical Clustering,” in *Introduction to HPC with MPI for Data Science*, F. Nielsen, Ed., Cham: Springer International Publishing, 2016, pp. 195–211. doi: 10.1007/978-3-319-21903-5_8.
[213] K. Backhaus, B. Erichson, S. Gensler, R. Weiber, and T. Weiber, “Cluster Analysis,” in *Multivariate Analysis: An Application-Oriented Introduction*, K. Backhaus, B. Erichson, S. Gensler, R. Weiber, and T. Weiber, Eds., Wiesbaden: Springer Fachmedien, 2021, pp. 451–530. doi: 10.1007/978-3-658-32589-3_8.
[214] F. Murtagh and P. Contreras, “Algorithms for hierarchical clustering: an overview, II,” *WIREs Data Min. Knowl. Discov.*, vol. 7, no. 6, p. e1219, 2017, doi: 10.1002/widm.1219.
[215] G. Gan, C. Ma, and J. Wu, “Hierarchical Clustering Techniques,” in *Data Clustering: Theory, Algorithms, and Applications*, in ASA-SIAM Series on Statistics and Applied Mathematics. , Philadelphia, Pa: Society for Industrial and Applied Mathematics, 2007, pp. 109–149. Accessed: May 23, 2024. [Online]. Available: https://epubs-siam-org.login.ezproxy.library.ualberta.ca/doi/epdf/10.1137/1.9780898718348.ch7
[216] M. Roux, “A Comparative Study of Divisive and Agglomerative Hierarchical Clustering Algorithms,” *J. Classif.*, vol. 35, no. 2, pp. 345–366, Jul. 2018, doi: 10.1007/s00357-018-9259-9.
[217] N. Erman, A. Korosec, and J. Suklan, “Performance of Selected Agglomerative Hierarchical Clustering Methods,” *Innov. Issues Approaches Soc. Sci.*, vol. 8, no. 1, pp. 180–204, Jan. 2015, doi: 10.12959/issn.1855-0541.IIASS-2015-no1-art11.
[218] B. S. Everitt, S. Landau, M. Leese, and D. Stahl, “Hierarchical Clustering,” in *Cluster Analysis*, 5th ed., in Wiley Series in Probability and Statistics. , John Wiley & Sons, Ltd, 2011, pp. 71–110. Accessed: Jun. 30, 2024. [Online]. Available: https://onlinelibrary-wiley-com.login.ezproxy.library.ualberta.ca/doi/abs/10.1002/9780470977811.ch4

[219] C. K. Reddy and C. C. Aggarwal, *A Survey of Partitional and Hierarchical Clustering Algorithms*. Chapman and Hall/CRC, 2013. Accessed: Jun. 28, 2024. [Online]. Available: https://learning.oreilly.com/library/view/data-clustering/9781466558229/K15510_C004.xhtml
[220] Y. Yang, "Temporal Data Clustering," in *Temporal Data Mining Via Unsupervised Ensemble Learning*, Elsevier, 2017, pp. 19–34. doi: 10.1016/B978-0-12-811654-8.00003-8.
[221] Z. Trábert *et al.*, "Effect of Land Use on the Benthic Diatom Community of the Danube River in the Region of Budapest," *Water*, vol. 12, p. 479, Feb. 2020, doi: 10.3390/w12020479.
[222] A. Tiri, N. Lahbari, and A. Boudoukha, "Assessment of the quality of water by hierarchical cluster and variance analyses of the Koudiat Medouar Watershed, East Algeria," *Appl. Water Sci.*, vol. 7, no. 8, pp. 4197–4206, Dec. 2017, doi: 10.1007/s13201-014-0261-z.
[223] J. Ukpatu, E. Udoinyang, and J. Udoh, "The Use of Agglomerative Hierarchical Cluster Analysis for the Assessment of Mangrove Water Quality of Okoro River Estuary, Southeastern Nigeria," vol. 3, pp. 17–24, Dec. 2015.
[224] C. Balderas Guzman, R. Wang, O. Muellerklein, M. Smith, and C. G. Eger, "Comparing stormwater quality and watershed typologies across the United States: A machine learning approach," *Water Res.*, vol. 216, p. 118283, Jun. 2022, doi: 10.1016/j.watres.2022.118283.
[225] N. Kahal, T. Lee, and T. Creetch, "Calgary Captured Year One Analysis: Technical Report May 2017 - May 2018," Miistakis Institute, Calgary, Dec. 2020. Accessed: Aug. 20, 2024. [Online]. Available: https://www.rockies.ca/files/reports/Calgary%20Captured_Reporting_May2017-2018.pdf
[226] "Gulls flock on Calgary's Bow River," *CBC News*, Apr. 23, 2010. Accessed: Aug. 20, 2024. [Online]. Available: https://www.cbc.ca/news/canada/calgary/gulls-flock-on-calgary-s-bow-river-1.894399
[227] A. Slater and J. McDonald, "Winter Birding along the Bow River." 1993. Accessed: Aug. 20, 2024. [Online]. Available: https://naturecalgary.com/wp-content/uploads/2018/07/location_alongthebowriver.pdf
[228] H. Pike, "What's behind the 'amazing' surge of animal activity?," *CBC News*, Apr. 10, 2020. Accessed: Aug. 20, 2024. [Online]. Available: https://www.cbc.ca/news/canada/calgary/calgary-animal-return-birds-pandemic-1.5528817
[229] C. Fisher and J. H. Acorn, *Birds of Alberta*. Lone Pine Pub., 2019.
[230] Water Resources Business Unit, "2020 Watershed Management Report," City of Calgary, Calgary, 2020. Accessed: Aug. 20, 2024. [Online]. Available: https://www.calgary.ca/content/dam/www/uep/water/documents/water-documents/watershed-management-report.pdf
[231] Palliser Environmental Services Ltd., "Updated Nose Creek Watershed Water Management Plan." Nose Creek Watershed Partnership, Apr. 2018. Accessed: Aug. 20, 2024. [Online]. Available: https://www.alidp.org/assets/pdfs/NCWWMP-Summary-Document.pdf
[232] L. R. Carson, C. Goodman, B. van Duin, and N. F. Neumann, "Application of a microbial and pathogen source tracking toolbox to identify infrastructure problems in stormwater drainage networks: a case study," *Microbiol. Spectr.*, vol. 12, no. 9, pp. e00337-24, Aug. 2024, doi: 10.1128/spectrum.00337-24.
[233] P. Bottoni, S. Caroli, and A. B. Caracciolo, "Pharmaceuticals as priority water contaminants," *Toxicol. Environ. Chem.*, vol. 92, no. 3, pp. 549–565, Mar. 2010, doi: 10.1080/02772241003614320.
[234] K. Kümmerer, "Pharmaceuticals in the Environment," *Annu. Rev. Environ. Resour.*, vol. 35, no. Volume 35, 2010, pp. 57–75, Nov. 2010, doi: 10.1146/annurev-environ-052809-161223.
[235] L.-H. Wu *et al.*, "Occurrence of bisphenol S in the environment and implications for human exposure: A short review," *Sci. Total Environ.*, vol. 615, pp. 87–98, Feb. 2018, doi: 10.1016/j.scitotenv.2017.09.194.
[236] S. Nicolardi, S. Herrera, M. J. M. Bueno, and A. R. Fernández-Alba, "Two new competitive ELISA methods for the determination of caffeine and cotinine in wastewater and river waters," *Anal. Methods*, vol. 4, no. 10, pp. 3364–3371, 2012, doi: 10.1039/C2AY25359C.

[237] K. A. Saffran, “Temporal and Spatial Trend for Coliform Bacteria in the Bow River 1951 - 1994.” Surface Water Assessment Branch, Technical Services and Monitoring Division, Alberta Environmental Protection, Jan. 1996. Accessed: Aug. 21, 2024. [Online]. Available: https://open.alberta.ca/dataset/7e8368bb-39d3-4639-9c22-210779f02d8b/resource/aa6f4883-48e1-4e6b-ac94-0206d7e4e8b8/download/7893.pdf
[238] I. Herrig *et al.*, “Prediction of fecal indicator organism concentrations in rivers: the shifting role of environmental factors under varying flow conditions,” *Environ. Sci. Eur.*, vol. 31, no. 1, p. 59, Sep. 2019, doi: 10.1186/s12302-019-0250-9.
[239] E. Rochelle-Newall, T. M. H. Nguyen, T. P. Q. Le, O. Sengtaheuanghoung, and O. Ribolzi, “A short review of fecal indicator bacteria in tropical aquatic ecosystems: knowledge gaps and future directions,” *Front. Microbiol.*, vol. 6, p. 308, Apr. 2015, doi: 10.3389/fmicb.2015.00308.
[240] S. Tamai, H. Shimamoto, K. Nukazawa, and Y. Suzuki, “Growth and Decay of Fecal Indicator Bacteria and Changes in the Coliform Composition on the Top Surface Sand of Coastal Beaches during the Rainy Season,” *Microorganisms*, vol. 11, no. 4, p. 1074, Apr. 2023, doi: 10.3390/microorganisms11041074.
[241] J. K. Bradshaw, B. Snyder, D. Spidle, R. C. Sidle, K. Sullivan, and M. Molina, “Sediment and Fecal Indicator Bacteria Loading in a Mixed Land Use Watershed: Contributions from Suspended Sediment and Bedload Transport,” *J. Environ. Qual.*, vol. 50, no. 3, pp. 598–611, May 2021, doi: 10.1002/jeq2.20166.
[242] Y. Liu, C. Zarfl, N. B. Basu, and O. A. Cirpka, “Modeling the Fate of Pharmaceuticals in a Fourth-Order River Under Competing Assumptions of Transient Storage,” *Water Resour. Res.*, vol. 56, no. 3, p. e2019WR026100, 2020, doi: 10.1029/2019WR026100.
[243] B. M. Mueller, H. Schulz, A. Höhne, A. Putschew, and J. Lewandowski, “Seasonal Differences in the Attenuation of Polar Trace Organics in the Hyporheic Zone of an Urban Stream,” *Water Resour. Res.*, vol. 58, no. 1, p. e2021WR031272, 2022, doi: 10.1029/2021WR031272.
[244] Y. Yang *et al.*, “Sucralose, a persistent artificial sweetener in the urban water cycle: insights into occurrence, chlorinated byproducts formation, and human exposure,” *J. Environ. Chem. Eng.*, vol. 9, no. 4, p. 105293, Aug. 2021, doi: https://doi.org/10.1016/j.jece.2021.105293.
[245] S. V. Sladkova, S. V. Kholodkevich, G. H. Olsen, P. Geraudie, and L. Camus, “Acute and long-term effects of sucralose on the water flea Daphnia magna on mobility, survival and reproduction at different temperature regimes,” *Chem. Ecol.*, vol. 32, no. 10, pp. 952–960, Nov. 2016, doi: 10.1080/02757540.2016.1212851.
[246] “Assessing microbial safety of drinking-water: Improving approaches and methods.” World Health Organization, Feb. 21, 2003. Accessed: Aug. 11, 2024. [Online]. Available: https://www.who.int/publications/i/item/9241546301
[247] L. P. de Souza *et al.*, “Photochemical environmental persistence of venlafaxine in an urban water reservoir: A combined experimental and computational investigation,” *Process Saf. Environ. Prot.*, vol. 166, pp. 478–490, Oct. 2022, doi: 10.1016/j.psep.2022.08.049.
[248] R. H. Jack *et al.*, “Seasonal trends in antidepressant prescribing, depression, anxiety and self-harm in adolescents and young adults: an open cohort study using English primary care data,” *BMJ Ment. Health*, vol. 26, no. 1, p. e300855, Nov. 2023, doi: 10.1136/bmjment-2023-300855.
[249] A. Balakrishnan, M. Sillanpää, M. M. Jacob, and D.-V. N. Vo, “Metformin as an emerging concern in wastewater: Occurrence, analysis and treatment methods,” *Environ. Res.*, vol. 213, p. 113613, Oct. 2022, doi: 10.1016/j.envres.2022.113613.
[250] M. Llorca *et al.*, “Fungal biodegradation of the *N*-nitrosodimethylamine precursors venlafaxine and *O*-desmethylvenlafaxine in water,” *Environ. Pollut.*, vol. 246, pp. 346–356, Mar. 2019, doi: 10.1016/j.envpol.2018.12.008.
[251] P. Grenni *et al.*, “Sulfamethoxazole persistence in a river water ecosystem and its effects on the natural microbial community and *Lemna minor* plant,” *Microchem. J.*, vol. 149, p. 103999, Sep. 2019, doi: 10.1016/j.microc.2019.103999.

[252] "Guidelines for drinking-water quality, 4th edition, incorporating the 1st addendum." World Health Organization, Apr. 24, 2017. Accessed: Aug. 22, 2024. [Online]. Available: https://www.who.int/publications/i/item/9789241549950
[253] E. P. Sauer, J. L. VandeWalle, M. J. Bootsma, and S. L. McLellan, "Detection of the human specific Bacteroides genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment," *Water Res.*, vol. 45, no. 14, pp. 4081–4091, Aug. 2011, doi: https://doi.org/10.1016/j.watres.2011.04.049.
[254] A. K. Salmore, E. J. Hollis, and S. L. McLellan, "Delineation of a chemical and biological signature for stormwater pollution in an urban river," *J. Water Health*, vol. 4, no. 2, pp. 247–262, Jun. 2006.
[255] K. Skiendzielewski *et al.*, "Two risk assessments: Evaluating the use of indicator HF183 *Bacteroides* versus pathogen measurements for modelling recreational illness risks in an urban watershed," *Water Res.*, vol. 259, p. 121852, Aug. 2024, doi: 10.1016/j.watres.2024.121852.
[256] Health Canada, "Canadian recreational water quality guidelines - Indicators of fecal contamination: E. coli and enterococci in recreational waters," Government of Canada. Accessed: Aug. 23, 2024. [Online]. Available: https://www.canada.ca/en/health-canada/services/publications/healthy-living/recreational-water-quality-guidelines-indicators-fecal-contamination/e-coli-enterococci.html
[257] J. T. Bunce, A. Robson, and D. W. Graham, "Seasonal influences on the use of genetic markers as performance indicators for small wastewater treatment plants," *Sci. Total Environ.*, vol. 739, p. 139928, Oct. 2020, doi: 10.1016/j.scitotenv.2020.139928.
[258] J. Oppenheimer, A. Eaton, M. Badruzzaman, A. W. Haghani, and J. G. Jacangelo, "Occurrence and suitability of sucralose as an indicator compound of wastewater loading to surface waters in urbanized regions," *Water Res.*, vol. 45, no. 13, pp. 4019–4027, Jul. 2011, doi: 10.1016/j.watres.2011.05.014.
[259] O. Golovko *et al.*, "Occurrence and removal of chemicals of emerging concern in wastewater treatment plants and their impact on receiving water systems," *Sci. Total Environ.*, vol. 754, p. 142122, Feb. 2021, doi: 10.1016/j.scitotenv.2020.142122.
[260] M. Sörengård, H. Campos-Pereira, M. Ullberg, F. Y. Lai, O. Golovko, and L. Ahrens, "Mass loads, source apportionment, and risk estimation of organic micropollutants from hospital and municipal wastewater in recipient catchments," *Chemosphere*, vol. 234, pp. 931–941, Nov. 2019, doi: 10.1016/j.chemosphere.2019.06.041.
[261] E. R. V. Dickenson, S. A. Snyder, D. L. Sedlak, and J. E. Drewes, "Indicator compounds for assessment of wastewater effluent contributions to flow and water quality," *Water Res.*, vol. 45, no. 3, pp. 1199–1212, Jan. 2011, doi: 10.1016/j.watres.2010.11.012.
[262] S. D. W. Comber, M. J. Gardner, and B. Ellor, "Seasonal variation of contaminant concentrations in wastewater treatment works effluents and river waters," *Environ. Technol.*, vol. 41, no. 21, pp. 2716–2730, Sep. 2020, doi: 10.1080/09593330.2019.1579872.
[263] N. M. Vieno, T. Tuhkanen, and L. Kronberg, "Seasonal Variation in the Occurrence of Pharmaceuticals in Effluents from a Sewage Treatment Plant and in the Recipient Water," *Environ. Sci. Technol.*, vol. 39, no. 21, pp. 8220–8226, Nov. 2005, doi: 10.1021/es051124k.
[264] Y. Yu, L. Wu, and A. C. Chang, "Seasonal variation of endocrine disrupting compounds, pharmaceuticals and personal care products in wastewater treatment plants," *Sci. Total Environ.*, vol. 442, pp. 310–316, Jan. 2013, doi: 10.1016/j.scitotenv.2012.10.001.
[265] A. Daneshvar *et al.*, "Evaluating pharmaceuticals and caffeine as indicators of fecal contamination in drinking water sources of the Greater Montreal region," *Chemosphere*, vol. 88, no. 1, pp. 131–139, Jun. 2012, doi: 10.1016/j.chemosphere.2012.03.016.
[266] S. C. Birrer *et al.*, "Legacy Metal Contaminants and Excess Nutrients in Low Flow Estuarine Embayments Alter Composition and Function of Benthic Bacterial Communities," *Front. Microbiol.*, vol. 12, p. 661177, Oct. 2021, doi: 10.3389/fmicb.2021.661177.

[267] "Water Quality Standards Handbook Chapter 5: General Policies." US EPA, Sep. 2014. Accessed: Aug. 24, 2024. [Online]. Available: https://nepis.epa.gov/Exe/ZyPDF.cgi/P100ODE3.PDF?Dockey=P100ODE3.PDF
[268] J. L. Wilkinson *et al.*, "Pharmaceutical pollution of the world's rivers," *Proc. Natl. Acad. Sci.*, vol. 119, no. 8, p. e2113947119, Feb. 2022, doi: 10.1073/pnas.2113947119.
[269] U. Lertxundi *et al.*, "Is the Environmental Risk of Metformin Underestimated?," *Environ. Sci. Technol.*, vol. 57, no. 23, pp. 8463–8466, Jun. 2023, doi: 10.1021/acs.est.3c02468.
[270] Y. He, Y. Zhang, and F. Ju, "Metformin Contamination in Global Waters: Biotic and Abiotic Transformation, Byproduct Generation and Toxicity, and Evaluation as a Pharmaceutical Indicator," *Environ. Sci. Technol.*, vol. 56, no. 19, pp. 13528–13545, Oct. 2022, doi: 10.1021/acs.est.2c02495.
[271] N. Morin-Crini *et al.*, "Worldwide cases of water pollution by emerging contaminants: a review," *Environ. Chem. Lett.*, vol. 20, no. 4, pp. 2311–2338, Aug. 2022, doi: 10.1007/s10311-022-01447-4.
[272] J. P. Nshimyimana, E. Ekklesia, P. Shanahan, L. H. C. Chua, and J. R. Thompson, "Distribution and abundance of human-specific Bacteroides and relation to traditional indicators in an urban tropical catchment," *J. Appl. Microbiol.*, vol. 116, no. 5, pp. 1369–1383, May 2014, doi: 10.1111/jam.12455.
[273] "Canada Goose - Abundance Map - eBird Status and Trends." Accessed: Aug. 06, 2024. [Online]. Available: https://science.ebird.org/status-and-trends/species/cangoo/abundance-map
[274] "Waterfowl Population Status, 2021." U.S. Fish & Wildlife Service, Aug. 20, 2021. Accessed: Aug. 06, 2024. [Online]. Available: https://www.fws.gov/sites/default/files/documents/WaterfowlPopulationStatusReport21.pdf
[275] M. K. McLean and S. A. Khan, "Toxicology of Frequently Encountered Nonsteroidal Anti-inflammatory Drugs in Dogs and Cats: An Update," *Vet. Clin. Small Anim. Pract.*, vol. 48, no. 6, pp. 969–984, Nov. 2018, doi: 10.1016/j.cvsm.2018.06.003.
[276] Government of Alberta, "Environmental quality guidelines for Alberta surface waters [2018] - Open Government." Water Policy Branch, Mar. 28, 2018. Accessed: Sep. 05, 2024. [Online]. Available: https://open.alberta.ca/publications/9781460138731
[277] T. Di Lorenzo, M. Cifoni, M. Baratti, G. Pieraccini, W. D. Di Marzio, and D. M. P. Galassi, "Four scenarios of environmental risk of diclofenac in European groundwater ecosystems," *Environ. Pollut.*, vol. 287, p. 117315, Oct. 2021, doi: 10.1016/j.envpol.2021.117315.

Appendix A: Sample Python Codes for BR2

Importing and processing the data

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import os
import numpy as np

os.chdir(r"G:\Shared drives\Arlos Lab\5_StudentFolders\JaimeHicks\Databases")
AllData = pd.read_excel('FINAL Micro and ESOC data (2018-2024).xlsx',
sheet_name='BR2')
AllData = pd.DataFrame(AllData).set_index('Sample Date')
drop_columns = []
for x in AllData:
    if AllData[x]["Numerical detections:"] < 7:
        drop_columns.append(x)
for x in drop_columns:
    AllData = AllData.drop(columns=x)
FilteredData =AllData.drop('Numerical detections:')
FilteredData=FilteredData.apply(pd.to_numeric, errors='coerce')
```

Data transformation ($\log_{10}$)

```
FilteredData =AllData.drop('Numerical detections:')
FilteredData=FilteredData.apply(pd.to_numeric, errors='coerce').apply(lambda
x: np.log10(x+1))
```

Data transformation (1/x)

```
FilteredData =AllData.drop('Numerical detections:')
FilteredData=FilteredData.apply(pd.to_numeric, errors='coerce').apply(lambda
x: 1/(x + 1))
```

Shapiro-Wilk Normality test

```
import scipy
from scipy import stats
normality_test=[]
normal_count=0
notnormal_count=0
import pprint

for col in FilteredData.columns:
    stat,p=stats.shapiro(FilteredData[col].dropna())
    results='%s,Statistics(W)= %e, p = %e' % (col,stat, p )
    if p>0.05:
        comment='Normal'
        normal_count+=1
```

```
    else:
        comment='Not Normal'
        notnormal_count+=1
    total=normal_count+notnormal_count
    results='%s,Statistics(W)= %e, p = %e,%s' % (col,stat, p,comment )
    normality_test.append(results)

pprint.pprint(list(normality_test))
print('The number of substances that are normal based on Shapiro-Wilk test is
%i out of %i.'% (normal_count,total))
```

Presence/absence plot

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np
import os
from matplotlib.colors import LinearSegmentedColormap

os.chdir(r"G:\Shared drives\Arlos Lab\5_StudentFolders\JaimeHicks\Databases")
data = pd.read_excel('Final Micro and ESOC data (2018-2024).xlsx',
sheet_name='BR2 for PresenceAbsence')
data['Sample Date'] = pd.to_datetime(data['Sample Date'], errors='coerce')

data.rename(columns={'E. coli MPN': 'E. coli', 'Total Coliforms MPN': 'Total
coliforms', 'O-Desmethyl-venlafaxine': 'ODV',
                     'Sulfamethoxazole':'SMX', 'Entero1':'Enterococci',
'Bisphenol S': 'BPS'}, inplace=True)

desired_order = ['E. coli', 'Enterococci', 'Total coliforms', 'HF183',
'HumM2', 'Rum2Bac', 'LeeSG', 'CG01', 'Acesulfame',
                 'Acetaminophen', 'Atenolol', 'BPS', 'Caffeine',
'Carbamazepine', 'Citalopram', 'Clarithromycin',
                 'Codeine', 'Cotinine', 'DEET', 'Diclofenac', 'Erythromycin',
'Ibuprofen', 'Metformin', 'Naproxen',
                 'ODV', 'Primidone', 'Propranolol', 'Sucralose', 'SMX',
'TCEP', 'TCPP',
                 'Trimethoprim', 'Venlafaxine']
variables = desired_order
data[variables] = data[variables].apply(pd.to_numeric, errors='coerce')

def convert_to_binary(x):
    if isinstance(x, str) or pd.isna(x):
        return 0  # Set strings or NaNs to 0
    elif x > 0:
        return 1  # Set values greater than 0 to 1
    elif x == 0:
        return 0  # Special marker for 0 values
    return 0  # Default case
```

```
def determine_overlay(x):
    if x == 0:
        return "NS"
    return ""

data_binary = data[variables].applymap(lambda x: convert_to_binary(x))
overlay_info = data[variables].applymap(lambda x: determine_overlay(x))
data[variables] = data_binary
data_binary['Sample Date'] = data['Sample Date']
data_binary.set_index('Sample Date', inplace=True)
data_binary.index = data_binary.index.strftime('%Y-%m-%d')

colors = ["#A4BDFF", "#F25E5E"]  # NaN = blue, Numerical = red
n_bins = 2  # Number of bins
custom_cmap = LinearSegmentedColormap.from_list("custom_colormap", colors,
N=n_bins)

plt.figure(figsize=(35, 15))
ax = sns.heatmap(data_binary.T, cmap=custom_cmap, cbar=False,
annot=overlay_info.T, fmt='', annot_kws={"size": 20,
"fontname": "Arial"})

yticks = ax.get_yticks()
ax.set_yticklabels(['E. coli', 'Enterococci', 'Total coliforms', 'HF183',
'HumM2', 'Rum2Bac', 'LeeSG', 'CG01', 'Acesulfame',
                    'Acetaminophen', 'Atenolol', 'BPS', 'Caffeine',
'Carbamazepine', 'Citalopram', 'Clarithromycin',
                    'Codeine', 'Cotinine', 'DEET', 'Diclofenac', 'Erythromycin',
'Ibuprofen', 'Metformin', 'Naproxen',
                    'ODV', 'Primidone', 'Propranolol', 'Sucralose', 'SMX',
'TCEP', 'TCPP',
                    'Trimethoprim', 'Venlafaxine'],
                      fontsize=25, fontname='Arial')
yticklabels = plt.yticks()[1]
for label in yticklabels:
    if label.get_text() in ['E. coli']:
        label.set_fontstyle('italic')
plt.xticks(rotation=45, ha='right', fontsize=20, fontname='Arial')
plt.xlabel('')
plt.grid(color='gray', linestyle='-', linewidth=0.5)
plt.show()
```

Enterococci time/series plot

```
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.dates as mdates

data = pd.read_excel('Final Micro and ESOC data (2018-2024).xlsx',
sheet_name='BR2 + flow')
data = data[data['Sample Date'] != 'Numerical detections:']
```

```
data['Sample Date'] = pd.to_datetime(data['Sample Date'], errors='coerce')

concentration_variables = [
     'Enterococci',
    # 'Total coliforms',
    # 'HF183',
    # 'CG01',
    # 'Sucralose'
]
flow_variable = 'Flow'
unit_mapping = { 'HF183': 'Copies/100 mL', 'CG01': 'Copies/100 mL',
'Total coliforms': 'MPN/100 mL', 'Flow': 'm³/s', 'Enterococci': 'CCE/100 mL'}

data[concentration_variables + [flow_variable]] =
data[concentration_variables + [flow_variable]].apply(pd.to_numeric,
errors='coerce').fillna(0)
data.set_index('Sample Date', inplace=True)
width = 35
height = 9
plt.rcParams['font.family'] = 'Arial'
fig, ax = plt.subplots(figsize=(width, height))
filtered_data = data[[concentration_variables[0], flow_variable]]


if not filtered_data.empty:
    ax.plot(filtered_data.index, filtered_data[concentration_variables[0]],
marker='o', label=concentration_variables[0], color='gray', linewidth=3,
markersize=10)
    ax.set_yscale('log')
    ax.set_ylabel(f'Concentration (log
{unit_mapping.get(concentration_variables[0], "ng/L")})', fontsize=35,
color='gray')
    ax.set_ylim(0.1, 1e6)
    ticks = [0.1, 1, 10, 100, 1000, 10000]
    ax.set_yticks(ticks)
    ax.get_yaxis().set_major_formatter(plt.ScalarFormatter())
    ax.get_yaxis().set_minor_formatter(plt.NullFormatter())
    ax.tick_params(axis='y', which='both', labelsize=35)
    ax2 = ax.twinx()
    ax2.plot(filtered_data.index, filtered_data[flow_variable], marker='o',
label='Flow', color='#4c7cfc', linewidth=3, markersize=10)
    ax2.set_ylabel(f'{flow_variable} ({unit_mapping.get(flow_variable)})',
fontsize=35, color='#4c7cfc')
    ax2.set_ylim(0, 340)
    ax2.tick_params(axis='y', which='both', labelsize=35)
    ax.xaxis.set_major_formatter(mdates.DateFormatter('%Y-%m'))
    ax.xaxis.set_major_locator(mdates.MonthLocator())
    ax.tick_params(axis='x', labelsize=25, rotation=45)

    for tick in ax.get_xticklabels():
```

```
        tick.set_rotation(45)
        tick.set_ha('right')
        tick.set_fontsize('25')
    ax2.legend(['Flow'], loc='upper left', fontsize=35)

fig.tight_layout()
plt.subplots_adjust(bottom=0.2)
plt.show()
```

Overall box plot

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np
import os

os.chdir(r"G:\Shared drives\Arlos Lab\5_StudentFolders\JaimeHicks\Databases")
AllData = pd.read_excel('Final Micro and ESOC data (2018-2024).xlsx',
sheet_name='BR2')
AllData = pd.DataFrame(AllData).set_index('Sample Date')
drop_columns = []

for x in AllData:
    if AllData[x]["Numerical detections:"] < 7:
        drop_columns.append(x)

for x in drop_columns:
    AllData = AllData.drop(columns=x)

FilteredData =AllData.drop('Numerical detections:')
FilteredData = FilteredData.apply(pd.to_numeric, errors='coerce')
FilteredData = FilteredData.fillna(0)
melted_data = pd.melt(FilteredData.reset_index(), id_vars=['Sample Date'],
var_name='Variable', value_name='Concentration')
melted_data['Variable'] = melted_data['Variable'].replace({
    'E. coli MPN': 'E. coli',
    'Total Coliforms MPN': 'Total coliforms',
    'Entero1': 'Enterococci',
    'O-Desmethyl-venlafaxine': 'ODV',
    'Sulfamethoxazole': 'SMX'
})

min_value = melted_data['Concentration'].min(skipna=True)  # Get minimum non-
negative value
melted_data['Concentration'] = np.where(melted_data['Concentration'] < 0,
min_value, melted_data['Concentration'])
```

```
y_min = -0.1
y_max = 1e6

palette = {'E. coli': '#a4bdff', 'Total coliforms': '#a4bdff', 'Enterococci':
'#a4bdff', 'HF183': '#ffd71a', 'Rum2Bac': '#ffd71a', 'CG01': '#ffd71a',
'Caffeine': '#91ebb2', 'DEET': '#91ebb2', 'Metformin': '#91ebb2', 'ODV':
'#91ebb2', 'Sucralose': '#91ebb2', 'SMX': '#91ebb2'}

order = ['E. coli', 'Total coliforms', 'Enterococci', 'Rum2Bac', 'Caffeine',
'DEET', 'Metformin', 'ODV', 'Sucralose', 'SMX']

plt.figure(figsize=(25, 12))
plt.rcParams['font.family'] = 'Arial'
sns.boxplot(data=melted_data, x='Variable', y='Concentration',
palette=palette, order=order, boxprops={'edgecolor': 'black'},
            medianprops={'color': 'black'},
            whiskerprops={'color': 'black'})

sns.stripplot(data=melted_data, x='Variable', y='Concentration',
color='gray', order=order, size=5, jitter=True,
            edgecolor='grey')

legend_handles = [plt.Rectangle((0, 0), 1, 1, color='#a4bdff', label='FIB'),
                  plt.Rectangle((0, 0), 1, 1, color='#ffd71a', label='MST
markers'),
                  plt.Rectangle((0, 0), 1, 1, color='#91ebb2', label='Chemical
indicators')]

plt.legend(handles=legend_handles, loc='upper right', prop={'size': 25,
'family': 'Arial'})

plt.yscale('symlog')
plt.ylim(y_min, y_max)
plt.xlabel('')
plt.ylabel('Concentration*', fontsize=25, fontname='Arial')

xticklabels = plt.xticks()[1]
for label in xticklabels:
    if label.get_text() in ['E. coli']:
        label.set_fontstyle('italic')

plt.xticks(np.arange(len(order)), xticklabels, rotation=0, fontsize=25,
fontname='Arial')
plt.yticks(fontsize=25, fontname='Arial')
plt.tight_layout()
plt.show()
```

Overall correlation and clustering

```
import seaborn as sns
import matplotlib.pyplot as plt
correlation_matrix = FilteredData.corr(method='pearson')

replaced_labels = {
    'E. coli MPN': 'E. coli',
    'Entero1': 'Enterococci',
    'Total Coliforms MPN': 'Total coliforms',
    'O-Desmethyl-venlafaxine': 'ODV',
    'Sulfamethoxazole': 'SMX',
    'Bisphenol S': 'BPS'
}
correlation_matrix = correlation_matrix.rename(columns=replaced_labels,
index=replaced_labels)

plt.figure(figsize=(5, 5))
heatmap = sns.heatmap(correlation_matrix, annot=True, cmap='RdYlBu_r', vmin=-
1, vmax=1, center=0, linewidths=.5, annot_kws={"size": 10})
colorbar = heatmap.collections[0].colorbar
colorbar.ax.tick_params(labelsize=20)
plt.xticks(fontsize=20, rotation=45, ha='right')
plt.yticks(fontsize=20, rotation=0)
plt.show()

plt.rcParams['font.family'] = "arial"
plt.figure(figsize=(20, 20))
correlation_matrix = correlation_matrix.fillna(0)
cluster_map = sns.clustermap(correlation_matrix, vmin=-1, vmax=1,
method="ward", annot=True, cmap="RdYlBu_r", annot_kws={"size": 15})

for ax in [cluster_map.ax_row_dendrogram, cluster_map.ax_col_dendrogram]:
    for line in ax.collections:
        line.set_linewidth(2)

plt.setp(cluster_map.ax_heatmap.yaxis.get_majorticklabels(), fontsize=20,
rotation=0)
plt.setp(cluster_map.ax_heatmap.xaxis.get_majorticklabels(), fontsize=20,
rotation=45, ha="right")

for label in cluster_map.ax_heatmap.get_xticklabels():
    if label.get_text() in ['E. coli']:
        label.set_fontstyle('italic')
for label in cluster_map.ax_heatmap.get_yticklabels():
    if label.get_text() in ['E. coli']:
        label.set_fontstyle('italic')
colorbar = cluster_map.cax
colorbar.yaxis.set_tick_params(labelsize=20)
colorbar.yaxis.label.set_size(20)
plt.show()
```

Appendix B: Supplementary Information

Table B.1 *Sample sizes for all sites across all datasets. DEET: N,N-diethyl-meta-toluamide. PFOA: Perfluorooctanoic acid. PFOS: Perfluorooctanesulfonic acid. TCEP: Tris(2-carboxyethyl)phosphine. TCPP: Tris(chloropropyl) phosphate.*

		BR2					BR3					BR4				
		All	**Spring**	**Summer**	**Fall**	**Winter**	**All**	**Spring**	**Summer**	**Fall**	**Winter**	**All**	**Spring**	**Summer**	**Fall**	**Winter**
Fecal indicator bacteria (FIB)	*E. coli*	22	6	13	2	1	48	11	11	15	11	54	14	14	16	10
	Enterococci	49	12	15	10	12	48	11	11	15	11	54	14	14	16	10
	Total coliforms	49	12	15	10	12	48	11	11	15	11	54	14	14	16	10
Microbial source tracking (MST) markers	CG01	4	0	1	2	1	22	3	0	9	10	25	5	1	9	10
	CowM3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Dog3	0	0	0	0	0	2	1	1	0	0	6	1	2	1	2
	LeeSG	0	0	0	0	0	14	3	4	7	0	17	6	6	4	1
	MuBac	0	0	0	0	0	1	0	0	1	0	1	0	0	1	0
	Pig2Bac	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	Rum2Bac	10	4	3	1	2	16	3	6	4	3	14	3	3	5	3
	HF183	5	1	1	0	3	38	8	6	14	10	52	13	13	16	10
	HumM2	2	0	0	0	2	4	1	1	1	1	46	13	12	11	10
Micropollutants	17a-Estradiol	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	17a-Ethynylestradiol	0	0	0	0	0	2	1	0	0	1	2	2	0	0	0
	17b-Estradiol	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4-n-Nonylphenol	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4-t-Octylphenol	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Acesulfame	0	0	0	0	0	1	0	0	0	1	30	7	8	11	4
	Acetaminophen	0	0	0	0	0	7	1	0	4	2	42	9	10	15	8
	Androstenedione	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
	Atenolol	0	0	0	0	0	1	0	0	1	0	35	8	6	14	7
	Benzyl butyl phthalate	0	0	0	0	0	2	1	1	0	0	2	0	2	0	0
	Bisphenol A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Bisphenol S	0	0	0	0	0	12	4	4	3	1	28	6	4	9	9
	Caffeine	11	3	6	1	1	30	9	5	11	5	52	13	13	16	10
	Carbamazepine	1	0	0	0	1	0	0	0	0	0	43	10	11	15	7

		BR2					BR3					BR4				
		All	Spring	Summer	Fall	Winter	All	Spring	Summer	Fall	Winter	All	Spring	Summer	Fall	Winter
Micropollutants	Citalopram	0	0	0	0	0	0	0	0	0	0	11	1	1	5	4
	Clarithromycin	0	0	0	0	0	0	0	0	0	0	24	5	6	9	4
	Codeine	0	0	0	0	0	1	0	0	1	0	38	8	10	13	7
	Cotinine	0	0	0	0	0	14	3	1	6	4	51	13	14	14	10
	DEET	8	0	8	0	0	20	2	10	6	2	45	10	13	12	10
	Di(2-ethylhexyl) phthalate	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
	Dibutyl phthalate	2	0	2	0	0	2	0	2	0	0	3	1	2	0	0
	Diclofenac	0	0	0	0	0	0	0	0	0	0	54	14	14	16	10
	Diethyl phthalate	0	0	0	0	0	0	0	0	0	0	2	1	1	0	0
	Dimethyl phthalate	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Di-n-octyl phthalate	1	0	0	1	0	0	0	0	0	0	1	1	0	0	0
	Equilenin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Equilin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Erythromycin	0	0	0	0	0	0	0	0	0	0	7	1	1	3	2
	Estriol	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Estrone	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Fluoxetine	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Gemfibrozil	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2
	Ibuprofen	0	0	0	0	0	1	0	1	0	0	9	2	1	2	4
	Metformin	47	12	13	10	12	44	11	9	14	10	53	13	14	16	10
	Naproxen	0	0	0	0	0	0	0	0	0	0	34	9	5	13	7
	Nifedipine	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Norethindrone	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Norfluoxetine	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Norgestimate	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	O-Desmethyl-venlafaxine	26	4	3	7	12	20	2	0	8	10	54	14	14	16	10
	Pantoprazole	0	0	0	0	0	0	0	0	0	0	4	0	1	3	0
	PFOA	2	0	2	0	0	1	0	1	0	0	3	0	2	1	0
	PFOS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Primidone	0	0	0	0	0	0	0	0	0	0	30	6	8	12	4

		BR2					BR3					BR4				
		All	**Spring**	**Summer**	**Fall**	**Winter**	**All**	**Spring**	**Summer**	**Fall**	**Winter**	**All**	**Spring**	**Summer**	**Fall**	**Winter**
Micropollutants	Progesterone	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Propranolol	0	0	0	0	0	0	0	0	0	0	7	0	0	4	3
	Salbutamol	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Sucralose	49	12	15	10	12	46	10	11	15	10	54	14	14	16	10
	Sulfamethoxazole	16	2	0	7	7	18	1	0	10	7	42	11	10	14	7
	TCEP	3	0	2	1	0	4	1	2	1	0	45	8	12	15	10
	TCPP	0	0	0	0	0	0	0	0	0	0	23	3	7	9	4
	Testosterone	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Triclosan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Trimethoprim	0	0	0	0	0	0	0	0	0	0	44	11	11	15	7
	Venlafaxine	5	1	0	2	2	2	0	0	1	1	32	7	9	12	4
	Zopiclone	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0

Table B.2 *BR2 statistics. DEET: N,N-diethyl-meta-toluamide. IQR: Interquartile range.*

Parameter	Data type	Median	First quartile (Q1)	Third quartile (Q3)	IQR	Skewness
E. coli	Overall	3.100	1.000	7.950	6.950	2.039
	Summer	3.100	2.000	9.700	7.700	2.207
Enterococci	Overall	79.004	36.571	159.770	123.199	5.889
	Spring	53.066	33.800	212.015	178.215	3.325
	Summer	62.100	32.350	136.298	103.948	2.394
	Fall	193.373	129.887	356.463	226.576	1.284
	Winter	74.752	37.712	129.161	91.450	0.095
Total coliforms	Overall	22.800	13.200	67.700	54.500	2.975
	Spring	21.250	16.525	44.000	27.475	2.779
	Summer	69.100	46.150	167.150	121.000	2.166
	Fall	22.350	14.275	45.150	30.875	2.224
	Winter	10.300	7.875	11.550	3.675	-0.063
Rum2Bac	Overall	3715.000	2805.000	5335.000	2530.000	0.454
Caffeine	Overall	6.700	6.143	13.300	7.157	3.052
DEET	Overall	12.900	11.850	15.375	3.525	2.402
	Summer	12.900	11.850	15.375	3.525	2.402
Metformin	Overall	17.000	13.250	20.945	7.695	0.984
	Spring	21.964	13.050	31.850	18.800	0.077
	Summer	13.900	12.700	15.400	2.700	0.553
	Fall	17.050	12.650	20.400	7.750	-0.303
	Winter	17.450	16.850	24.025	7.175	1.198
O-Desmethyl-venlafaxine	Overall	2.000	1.425	2.300	0.875	-0.146
	Fall	1.700	1.377	2.100	0.723	0.404
	Winter	2.347	2.100	2.499	0.399	-1.036
Sucralose	Overall	46.600	34.700	60.100	25.400	0.657
	Spring	47.300	32.000	53.200	21.200	0.252
	Summer	31.600	20.250	42.650	22.400	1.037
	Fall	48.450	44.525	58.987	14.462	0.565
	Winter	58.375	49.125	64.748	15.623	1.968
Sulfamethoxazole	Overall	1.700	1.450	1.925	0.475	0.915
	Fall	1.800	1.550	1.850	0.300	2.109
	Winter	1.900	1.450	2.207	0.757	0.169

Table B.3 *BR3 statistics. DEET: N,N-diethyl-meta-toluamide. IQR: Interquartile range.*

Parameter	Data type	Median	First quartile (Q1)	Third quartile (Q3)	IQR	Skewness
E. coli	Overall	44.100	20.000	82.825	62.825	6.448
	Spring	18.700	14.100	28.650	14.550	2.174
	Summer	47.300	32.900	69.900	37.000	0.808
	Fall	44.100	22.350	88.050	65.700	3.852
	Winter	62.000	47.650	124.000	76.350	1.914
Enterococci	Overall	1411.156	295.508	2942.436	2646.928	6.629
	Spring	326.000	170.500	1310.396	1139.896	0.722
	Summer	681.839	227.500	1627.332	1399.832	1.027
	Fall	2412.075	690.895	4294.083	3603.188	3.829
	Winter	2560.000	1712.780	4215.000	2502.220	2.585
Total coliforms	Overall	290.900	140.925	475.125	334.200	3.413
	Spring	218.700	112.100	476.600	364.500	3.224
	Summer	648.800	346.800	1251.500	904.700	1.533
	Fall	248.100	138.750	347.450	208.700	3.131
	Winter	218.700	148.150	317.750	169.600	3.225
HF183	Overall	3480.000	1941.000	5700.000	3759.000	2.613
	Spring	3030.000	1876.500	4984.500	3108.000	0.571
	Fall	3180.000	2092.500	5765.000	3672.500	1.101
	Winter	4065.000	3306.000	6741.000	3435.000	2.250
Rum2Bac	Overall	2235.000	1555.500	3391.500	1836.000	3.318
LeeSG	Overall	3165.000	1665.000	4379.500	2714.500	0.880
	Fall	2570.000	1710.000	4198.000	2488.000	1.278
CG01	Overall	15189.000	6141.000	30280.500	24139.500	3.069
	Fall	31800.000	18462.000	43600.000	25138.000	2.510
	Winter	13130.000	6610.500	17923.500	11313.000	3.071
Acetaminophen	Overall	8.700	7.112	10.650	3.538	-0.262
Bisphenol S	Overall	11.650	7.500	104.870	97.370	2.877
Caffeine	Overall	20.750	10.219	30.650	20.431	2.198
	Spring	16.600	9.400	32.766	23.366	1.969
	Fall	23.000	16.100	34.752	18.652	1.740
Cotinine	Overall	2.150	1.266	2.733	1.467	1.516
DEET	Overall	13.442	9.700	23.850	14.150	0.879
	Summer	15.250	10.050	22.300	12.250	0.962
Metformin	Overall	14.550	11.525	18.700	7.175	1.191
	Spring	15.400	11.634	23.200	11.566	0.859
	Summer	12.100	11.000	12.700	1.700	1.567
	Fall	14.100	11.575	21.100	9.525	0.380
	Winter	17.400	15.200	18.425	3.225	1.056
O-Desmethyl-venlafaxine	Overall	1.505	1.275	1.925	0.650	0.135
	Fall	1.400	1.175	1.607	0.432	0.471
	Winter	1.800	1.525	2.100	0.575	-0.476

Parameter	Data type	Median	First quartile (Q1)	Third quartile (Q3)	IQR	Skewness
Sucralose	Overall	39.021	27.175	50.571	23.396	0.194
	Spring	33.321	21.000	50.600	29.600	0.414
	Summer	26.400	17.050	32.100	15.050	1.692
	Fall	44.300	36.200	52.302	16.102	0.015
	Winter	45.450	40.775	51.049	10.274	0.727
Sulfamethoxazole	Overall	1.600	1.400	1.980	0.580	0.956
	Fall	1.550	1.400	1.775	0.375	1.588
	Winter	1.700	1.500	2.440	0.940	0.247

Table B.4 *BR4 statistics. DEET: N,N-diethyl-meta-toluamide. IQR: Interquartile range. TCEP: Tris(2-carboxyethyl)phosphine. TCPP: Tris(chloropropyl) phosphate.*

Parameter	Data type	Median	First quartile (Q1)	Third quartile (Q3)	IQR	Skewness
E. coli	Overall	40.700	22.875	84.350	61.475	4.036
	Spring	32.350	18.675	55.250	36.575	2.535
	Summer	38.200	17.850	86.975	69.125	0.883
	Fall	36.150	20.925	70.100	49.175	0.963
	Winter	80.450	54.850	97.350	42.500	-0.554
Enterococci	Overall	9385.000	3890.000	17184.470	13294.470	1.926
	Spring	9015.000	3998.143	18406.956	14408.814	1.833
	Summer	2580.000	1953.912	8559.735	6605.823	1.754
	Fall	6901.949	4585.567	14567.933	9982.366	1.758
	Winter	17022.980	15087.599	27917.626	12830.028	1.697
Total coliforms	Overall	596.250	312.075	965.500	653.425	5.179
	Spring	532.350	277.900	1264.300	986.400	3.547
	Summer	925.400	770.100	2229.075	1458.975	3.096
	Fall	448.150	299.575	703.925	404.350	1.813
	Winter	387.700	257.350	604.675	347.325	0.906
HF183	Overall	40107.000	17370.000	88447.000	71077.000	2.712
	Spring	69800.000	34068.000	133260.000	99192.000	0.381
	Summer	24222.000	11200.000	38970.000	27770.000	1.458
	Fall	45010.000	15639.000	65268.000	49629.000	1.745
	Winter	53750.000	36744.000	157237.500	120493.500	1.603
HumM2	Overall	5478.000	3350.000	15057.000	11707.000	3.338
	Spring	6620.000	4890.000	40980.000	36090.000	2.582
	Summer	3884.000	2940.000	5049.000	2109.000	3.047
	Fall	5460.000	2585.000	11408.000	8823.000	2.098
	Winter	12134.000	7127.000	15057.000	7930.000	2.779
Rum2Bac	Overall	2541.000	1911.000	4392.500	2481.500	1.938
LeeSG	Overall	2990.000	1800.000	6820.000	5020.000	1.298
CG01	Overall	5460.000	2676.000	20040.000	17364.000	2.217
	Fall	16100.000	8178.000	28920.000	20742.000	0.452
	Winter	3832.000	2697.000	11461.500	8764.500	1.602
Acesulfame	Overall	245.500	142.500	539.000	396.500	1.401
	Spring	488.000	392.651	843.095	450.443	0.902
	Summer	139.500	109.625	160.000	50.375	-0.102
	Fall	205.000	145.000	319.500	174.500	1.886
Acetaminophen	Overall	10.000	6.825	13.900	7.075	2.157
	Spring	8.900	8.200	13.600	5.400	0.949
	Summer	6.850	5.925	16.800	10.875	1.880
	Fall	11.000	7.554	14.250	6.696	2.451
	Winter	11.700	10.100	13.775	3.675	1.729
Atenolol	Overall	11.720	8.800	15.550	6.750	0.359
	Spring	12.374	9.831	13.541	3.710	-1.111
	Fall	11.060	9.075	14.700	5.625	0.422
	Winter	17.200	16.300	18.550	2.250	0.271

Parameter	Data type	Median	First quartile (Q1)	Third quartile (Q3)	IQR	Skewness
Bisphenol S	Overall	11.750	7.450	45.475	38.025	2.735
	Fall	9.000	6.300	23.300	17.000	1.183
	Winter	10.400	7.700	11.900	4.200	1.048
Caffeine	Overall	38.650	14.500	72.025	57.525	6.070
	Spring	45.500	13.464	71.500	58.036	3.379
	Summer	18.300	12.000	48.100	36.100	1.258
	Fall	26.900	20.425	72.275	51.850	2.196
	Winter	46.800	23.200	81.025	57.825	0.172
Carbamazepine	Overall	25.200	17.000	37.700	20.700	0.376
	Spring	22.300	14.903	25.895	10.992	0.251
	Summer	16.700	9.000	21.600	12.600	1.370
	Fall	33.600	26.650	41.400	14.750	0.320
	Winter	39.100	22.050	42.500	20.450	-0.174
Citalopram	Overall	25.058	23.900	26.900	3.000	2.508
Clarithromycin	Overall	15.950	13.175	43.552	30.377	1.315
	Fall	20.800	13.900	51.400	37.500	1.119
Codeine	Overall	59.600	41.775	79.950	38.175	1.088
	Spring	61.066	27.544	79.980	52.435	0.848
	Summer	23.150	12.175	42.525	30.350	1.966
	Fall	60.300	54.400	77.100	22.700	2.957
	Winter	82.800	77.100	95.100	18.000	1.590
Cotinine	Overall	5.106	3.100	9.300	6.200	1.563
	Spring	5.300	2.800	10.200	7.400	0.422
	Summer	3.100	1.700	4.000	2.300	0.707
	Fall	5.550	3.850	8.099	4.249	1.867
	Winter	15.000	8.925	17.975	9.050	0.393
DEET	Overall	19.616	14.400	39.000	24.600	3.384
	Spring	19.958	16.521	26.350	9.829	2.181
	Summer	45.300	24.900	106.000	81.100	1.654
	Fall	14.650	9.250	17.725	8.475	2.975
	Winter	17.000	15.750	20.700	4.950	2.934
Diclofenac	Overall	209.500	139.356	305.000	165.644	0.486
	Spring	153.712	69.650	190.058	120.408	0.235
	Summer	118.000	69.050	200.500	131.450	0.612
	Fall	270.500	208.000	368.000	160.000	0.763
	Winter	368.000	274.750	420.250	145.500	-0.399
Erythromycin	Overall	2.600	2.400	4.750	2.350	1.944
Ibuprofen	Overall	11.400	10.600	15.300	4.700	0.892
Metformin	Overall	696.000	442.000	1200.000	758.000	0.630
	Spring	920.000	509.100	1430.000	920.900	0.367
	Summer	395.500	335.500	489.250	153.750	-0.867
	Fall	714.000	504.000	1003.750	499.750	1.113
	Winter	1445.000	1227.500	1537.500	310.000	-0.526
Naproxen	Overall	12.500	9.560	16.150	6.590	1.287
	Spring	8.722	8.300	16.200	7.900	0.556
	Fall	12.600	10.300	16.000	5.700	0.124
	Winter	13.800	12.350	18.700	6.350	1.718

Parameter	Data type	Median	First quartile (Q1)	Third quartile (Q3)	IQR	Skewness
O-Desmethyl-venlafaxine	Overall	157.000	105.250	230.250	125.000	2.062
	Spring	142.590	58.300	157.000	98.700	-0.455
	Summer	90.900	54.525	135.750	81.225	1.906
	Fall	195.000	149.000	246.250	97.250	3.106
	Winter	245.500	201.000	290.250	89.250	0.582
Primidone	Overall	4.700	3.025	6.250	3.225	-0.035
	Summer	2.350	1.575	3.225	1.650	1.029
	Fall	5.795	4.375	6.800	2.425	-0.296
Propranolol	Overall	5.700	5.400	7.100	1.700	2.638
Sucralose	Overall	2300.000	1726.898	2832.000	1105.102	0.583
	Spring	2025.000	835.500	2460.000	1624.500	-0.108
	Summer	1680.000	795.250	2202.500	1407.250	0.311
	Fall	2653.000	1972.500	2942.500	970.000	0.352
	Winter	3135.000	2702.500	3832.500	1130.000	1.745
Sulfamethoxazole	Overall	15.401	9.800	27.300	17.500	0.175
	Spring	10.100	4.321	13.451	9.129	0.610
	Summer	10.550	7.475	18.650	11.175	0.581
	Fall	28.050	24.352	29.800	5.448	-0.507
	Winter	24.900	11.750	30.000	18.250	-0.256
TCEP	Overall	14.400	10.500	16.800	6.300	1.289
	Spring	12.500	10.850	16.735	5.885	1.439
	Summer	12.450	7.550	15.825	8.275	1.930
	Fall	15.400	10.650	17.000	6.350	0.740
TCPP	Overall	121.000	107.500	161.000	53.500	3.550
	Summer	106.000	79.300	111.000	31.700	-0.917
	Fall	129.000	114.000	158.000	44.000	0.879
	Winter	14.650	12.725	17.775	5.050	-0.042
Trimethoprim	Overall	17.773	9.775	24.800	15.025	0.184
	Spring	12.914	7.000	17.773	10.773	0.748
	Summer	9.700	5.200	11.050	5.850	1.512
	Fall	19.700	17.874	27.150	9.277	0.360
	Winter	25.200	24.900	26.050	1.150	1.224
Venlafaxine	Overall	36.949	23.700	54.050	30.350	0.441
	Spring	25.509	13.850	45.849	31.999	0.643
	Summer	20.700	15.100	29.500	14.400	0.890
	Fall	50.100	35.400	57.325	21.925	1.007