
# Investigating the public health risk of urban flooding events

**Sophie Scutt**

**Thesis submitted for the degree of**
**Doctor of Philosophy**

The University of Sheffield
Department of Civil and Structural Engineering
April 2024

# Abstract

In recent years, the frequency and intensity of flooding has increased. With flooding, comes a plethora of detrimental effects, including damage to infrastructure and utilities. However, often overlooked is the risk floods can pose to public health. Urban flood water can act as a vector for the dissemination of disease causing bacteria- such bacteria can be present as a risk suspended in the floodwater, or via deposition onto surfaces. Potentially causing illness should a member of the public encounter such bacteria. To mitigate the risk an urban flood can pose to public health, it is important that the dynamics and behaviour of the bacterial communities present in a flood scenario are understood.

In this research the dynamics of bacterial communities at urban flood sites have been investigated. This includes: identifying the bacterial community structure of urban flood water and affected urban soils and how this changes, understanding the bacterial transfer between the flood water and urban soil, and also determining how this changes over long and short term scales. In this study a field study was performed, sampling two sites with a high surface flood risk over a long term period- several months, and also in the short term- sampling for 3 consecutive days following a heavy rainfall and thus flood event. Collected soil and floodwater samples, were analysed using heterotrophic plate counts, but also by sequencing the V4 region of the 16s rRNA gene. Complementing the field study, a laboratory based soil column study was also carried out, similarly analysing samples using HPCs and 16s rRNA sequencing.

Bacterial communities in both soil and floodwater showed seasonal changes in the long term study, including changes in diversity, and bacterial abundance. The short term field study, similarly detected faecal contamination at one sample site, but found that no disease causing or pathogenic bacteria were present in flood affected soils in the 3 days following a flood event at either site. This study also found that the altered soil bacterial community at flood sites is thought to be further impacted by the presence of faecal contamination, finding lower diversity at the site though to be faecally contaminated. During the column study, it was found that most bacteria remained within the top 0.4m of the soil column, with no pathogenic species found to have moved into the soil. This study also compared the distribution and transport of the bacteria to that of a solute in order to hydraulically characterise the soil. Naturally, bacterial transport and distribution was different to that of a solute, with soil characteristics such as grain and pore size, along with bacterial behaviours, such as biofilm formation, thought to influence bacterial transport and distribution .

This research has provided new insights on the presence, risk and behaviour of bacteria at urban flood sites. It also emphasises how ‘real life’ data from the field is still of utmost importance when assessing risk, as each flood site is unique with different physico-chemical factors, bacterial community compositions, and also flood origins, which can all influence the ‘risk’ to public health.

# Table of contents

# Acknowledgments

The research reported here was supported by the UK Engineering and Physical Sciences Research Council (EPSRC). The laboratory work was performed at the NERC Biomolecular Analysis Facility at the University of Sheffield and was supported by the Natural Environment Research Council, UK.

I would like to thank my supervisors, Dr James Shucksmith, Dr Henriette Jensen and Dr Isabel Douterelo for their unwavering support during the last 3 and a half years, for encouraging me and believing in me when I didn't believe in myself. You have shown me the true value of collaboration, not only academically but also personally and this is something I hope to carry forward.
Isabel- thank you for being more than a supervisor, for being a friend and role model. I do not have enough space here to write how much you have helped and inspired me so I will leave it at that for now.

Of course, I would like to thank my mum and dad for always supporting me and pushing me to achieve my best, even if you didn't always understand what exactly you were supporting me in! You always told me I can do anything I set my mind to, even if it hasn't always been easy. Thankyou for giving me a life in which I had choices and space to dream.

Thankyou to my best friend Charlotte. The kindest and most understanding person- what would I have done without you? This thesis wouldn't have been possible without all of our nights sat in your car with a Starbucks, laughing, or crying, or just being- together. "Kindred spirits are not so scarce as I used to think. It's splendid to find out there are so many of them in the world."

A special and long overdue thankyou to my partner- Elliot. You have really been here through all of the tears, the late nights and all of the worry, and through it all, your support never wavered. You kept me grounded and reminded me of the things I have to be excited about in my life outside of this PhD! I don't know how you've put up with me but I will always be grateful that you have.

I would like to dedicate this thesis to every woman who persevered, when they were told they could not. You do not need to be less outspoken, or less friendly, or less yourself to succeed. During my time studying for this I have overcome things I truly never thought I could, and there were moments when I did not believe I could get this far. So finally, thank you to me. I did this!

# Declaration of Authorship

I ,Sophie Scutt, hereby declare that I am the sole author of this thesis titled “Investigating the public health risk of Urban Flooding events” and that the contents of this research are the result of my own work.

I confirm that:

- The dissertation is original and has not been submitted for any other degree, diploma or other qualifications to this university or any other institution.
- This work was done wholly while in candidature for a research degree at this University. ▪
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

April 2024,

Sophie Scutt

# List of Figures

**Supplementary figures within appendices**

# List of Tables

**Supplementary tables within appendices**



# List of Equations

# Chapter 1
# Introduction

## 1.1 Introduction

The risks urban floods pose to public health, is yet to be fully understood. There are several factors to consider, all of which merit investigation. Some factors are far reaching- i.e. the root cause, or contributor, of flooding- climate change, and extreme weather events. Other factors are more specific, such as the processes involved in wastewater treatment, hydraulic performance of infrastructure and movement of water through these systems. To fully consider this problem, social and psychological studies could even be carried out on human behaviour around floods. This topic is multi-faceted with lots of uncertainties and many paths to investigate.
Understanding the presence and behaviour of pathogens in floodwater itself is also essential to better understand public health risk. Although the potential for other contaminants to cause illness is of importance, it is contact with a wide variety of pathogens that can potentially cause severe, infectious, and potentially epidemic illness among the population. Not only do we need to investigate the pathogenic and disease-causing species present, - but also how long they survive after a flood, how many are present, how they move during a flood, and how they might travel between the flood water and urban surfaces. Past studies tend to focus on pathogenic *E. coli*- due to ease of detection and its levels being regulated in legislations for microbial surveillance of rivers and waterways, such as the Water Framework Directive (DEFRA, 2014). This faecal bacterium is traditionally used as an indicator for the presence of faeces and thus disease (Odonkor and Ampofo, 2013; Price and Wildeboer, 2017; Shah et al., 2016; Ten Veldhuis et al., 2010; Yard et al., 2014). However, this study will focus on identifying the actual microbial community structure present in urban floodwater.
To build a better understanding of these processes this project will focus on the specific aspect of transport and quantification of bacterial species, including potentially pathogenic microorganisms, in and between urban soils and flood water. A better understanding of microbial dynamics will allow to determine the safety of a recently flooded area in the short and long term.

## 1.2 Research aims and objectives.

The main aim of the project is to determine to what extent urban flood events, due to direct exposure to polluted water, pose a risk to public health and/or the environment. This will be investigated by focusing on the specific aspect of transport and survival of potential pathogenic microorganisms in and between urban soils and flood water.

A series of research objectives are required to achieve the project aim and are outlined below:

1. Determine the bacterial community structure, including potential pathogens, which are likely to be present in urban flood situations. This will be achieved by combining field work (short- and long-term sampling) and laboratory soil column experiments.
2. Investigate the dynamics of potential pathogens at the soil\water interface following a contaminated flood event, quantify and understand the processes and timescales

involved. This will involve the use of a laboratory-based soil column in which samples will be taken at varying depths, and at different time points after a contaminated water flooding event.

3. Determine whether characteristics of the urban environment (such as soil composition, pH, temperature, rainfall, etc.) have an impact on bacterial distribution between soil and floodwater. The impact of such factors will be investigated by combining field work and soil column experiments.

### 1.3 Dissertation Overview

This thesis consists of 7 chapters which include a review of the literature, methodology, analyses, data, results and conclusions. The thesis also incorporates a collection of three manuscripts that are suitable for publication in a peer-reviewed journal.
This thesis is organised as follows:

- ***Chapter 1*** corresponds to this introduction.
- ***Chapter 2*** shows a general overview of the literature and background of the topics.
- ***Chapter 3*** gives an overview of the methods used for this research.
- ***Chapter 4*** presents the first manuscript derived from this research titled '*Investigating the bacterial communities of urban flood water, and the impacts on soils at urban flood sites*.'
- ***Chapter 5*** presents the second manuscript derived from this research titled *'Exploring Short-Term Public Health Risks of Urban Floods in the UK Through Bacterial Community Analysis'*
- ***Chapter 6*** presents the third manuscript derived from this research titled '*Spatial and temporal variation of bacteria associated with faecal polluted soils after a flooding event'*.
- ***Chapter 7*** summarises the main conclusions from this research.
- ***Appendices*** includes supplementary information for chapters 4, 5, and 6.

# Chapter 2
# Background and Literature Review

## 2.1 Climate change and urban flood events

Temporally, the frequency and occurrence of extreme hydrological events – such as intense precipitation -- has steadily been increasing due to climate change (Pall *et al.,* 2011). This precipitation has increased in intensity - volume, as well as frequency in some cases, consequently leading to an increase in flood events (Waters *et al*., 2010).
Many sewer systems in densely populated urban areas are becoming unable to cope with sudden influxes of rainfall runoff, resulting from intense precipitation. Combined sewer systems (CSOs) - the main sewer type used in most UK towns and cities - are designed to discharge diluted wastewater, storm water, and sewage usually, via a 'release valve'. The discharge flows out into a river or water body, when the sewer becomes 'full' to prevent damage to the sewer infrastructure, as well as to prevent uncontrolled flooding of homes and buildings (Phillips *et al.,* 2012).

CSOs are not the only system that discharges water via outlets into the environment. Surface water outflow pipes also discharge rainwater that runs-off of roads, gutters, fields, etc. into water bodies, including groundwater via soakaways. Surface water systems discharge regularly, as the discharge is rainwater runoff, and so it is not usually thought of as harmful, in terms of the risk it poses to public health (Webber *et al.,* 2018). However, in some cases, misconnection of pipes can also lead to wastewater being discharged, along with surface runoff rainwater (Chandler and Lerner, 2015).

Many studies have investigated how CSOs spills affect the ecology of rivers (Veronesi *et al.,* 2013; Wang,2014; Wu *et al.,* 2019), as well as the immediate risk posed to the public using the river recreationally (de Man *et al.,* 2013; Gray, 2008; Ten Veldhuis *et al.,* 2010). However, contaminated water can also end up discharging into the urban environment – before it reaches the 'release valve'- due to misconnections, old infrastructure, blockages or from the lack of local capacity in the pipe network under intense rainfall, leading to an urban flood or exceedance event. This may result in contaminated water surcharging from manholes and gulleys onto streets, roads, and recreational areas (Abdellatif *et al*., 2013; National Academies of Sciences, Engineering, and Medicine, 2019; Olds *et al.,* 2018).

Although more infrequent, flood water can also reach people's homes – presenting a more immediate risk of contact with contaminated water. As urban areas tend to comprise a variety of impermeable surfaces, as well as permeable but easily saturated areas, such as parkland, the excess water cannot be easily absorbed into soil. On the one hand, the permeable properties of such areas can work favourably as a designated area to saturate and store flood water, preventing damage to homes and buildings (Förster *et al.,* 2008; Schultz, 2006). However, when this is not the intention for the area, the water can sit in the area for days or even several weeks, potentially posing a microbiological risk (Gholami *et al.,* 2010; Serinaldi *et al.,* 2018).

## 2.2 Defining the risk and identifying hazards associated to flooding

A risk can be defined as the combination of 'potential harm and likelihood of exposure' (Šotić and Rajić, 2015). In the field of public health, 'risk' has a robust definition, and is usually evaluated using a selection of specific methods and tools, creating a multifaceted risk matrix or risk assessment strategy. For example, a typical tool used to examine risk in a flooding scenario is a microbiological risk assessment (MRA) or quantitative microbial risk assessment (QMRA). This method estimates the risk to public health, whilst accounting for any influencing factors by: identifying the microorganisms likely to cause harm, assessing exposure/dose exposed to, applying a dose-response model, and characterising risk in terms of probability of infection and illness (Haas *et al*., 2014; WHO, 2016). This method has been used by other large-scale flood studies including that of ten Veldhuis *et al.,* and Fewtrell *et al* (Fewtrell et al., 2011; ten Veldhuis et al., 2010). However, this approach requires a researcher to focus on a select number of microorganisms/species, and doesn't account for the whole bacterial community present. Thus, this study, instead of focusing on 'risk' discretely, will instead look at characterising the microbiological hazards (likely bacteria and/or pathogens) present during a flood event, and factors that influence the behaviour of such hazards. Instead of carrying out a full QMRA, the first stage 'identifying the microorganisms likely to cause harm' is the focus here.

In the case of the public health risk of urban floods, it is vital to understand the environment where the 'potential harm' exists in. The pathogens in wastewater, have the potential to cause harm if they are able to survive in soils after the floodwater withdraws (the point this study will investigate). The likelihood of this potential harm is dependent on the frequency of flood occurrence, and the probability of someone encountering the pathogens. Naturally, the floodwater, along with the contaminants are only a risk if members of the public physically come into contact with it. Scenarios in which this is likely to happen may include the following: children playing in floodwater, vehicles driving through floodwater and splashing pedestrians, or the floodwater being widespread in areas that may be the sole access route- and so there is no other option other than to walk through (Schets *et al*.,2008; Ten Veldhuis *et al.*, 2010). If the flood is more severe, there is also the possibility the floodwater may enter homes and gardens, making direct contact with residents unavoidable.

### 2.3 The microbial contaminants of urban floodwater

Microbial contaminants – pathogens- are important to consider. Flood water in urban environments is usually a mixture of diluted sewage, rainwater runoff- containing dirt and animal faeces, from streets and roads, or faeces from livestock, chemicals from fertilisers and pesticides/herbicides, resulting from agricultural run-off, as well as metals, often originating from industrial and domestic practices via leaching (Ghane *et al.,* 2016; Scoullos *et al*.,2020; Yard *et al*., 2014). Floodwater, therefore, has the potential to induce illness (Muirhead *et al*., 2004). This is due to bacterial pathogens, intestinal parasites, and enteric viruses, present in the floodwater, leading to outbreaks of waterborne diseases (Fewtrell *et al*., 2011; Ohl and Tapsell, 2000). Thus, causing health problems including, gastrointestinal and dermatological issues, and even more serious illnesses such as septicaemia (Baker-Austin *et al*., 2018; de Man *et al*., 2014; Haake and Levett, 2015).

The most common waterborne pathogens present in urban flood water globally, after an intense hydrological event, such as heavy rainfall include *Vibrio, Leptospira., Campylobacter*, and

*Cryptosporidium* (Goodfellow and Taube, 2016). With *E. coli* and *Giardia* also being frequently detected in UK specific studies of urban floodwater (Fewtrell et al., 2011). In UK drinking water the presence of any bacterial pathogens, deems the water unsafe for consumption (Drinking water inspectorate, 2019), therefore bacterial pathogens present at any level - as seen in floodwater- could also pose a risk. Despite *Cryptosporidium* being responsible for 69% of waterborne disease outbreaks in the UK over an 11-year period, it is *E. coli* that is routinely used as an indicator for faecal contamination in drinking water and freshwater by water companies and the Environment Agency, who are legally bound to do so by legislation (DEFRA, 2014). This is because using *E. coli* detection is affordable and easier to perform than other methods (Odonkor and Ampofo, 2013). This leaves a gap in the literature, with less information available on the behaviour of pathogens present in surface floods.

**2.4 Fate and transport of pathogens from contaminated water to and from soil**

It is widely understood and known that pathogens are present in flood water itself - this is visible as pools of water in urban areas following a flood event- presenting an obvious risk (de Man *et al.*, 2013). However, it is worth noting that contaminated water can serve as a vehicle for transporting pathogens, such as through infiltration into the ground. This process poses a significant threat to public health by potentially contaminating soil and surfaces in the aftermath of floods (Rubinato *et al.*, 2019). With this in mind, the infiltration ability of a soil is an important factor to consider. The slope of the ground, as well as physical properties of soil, including composition, particle density, and texture can significantly impact the water holding capacity of the soil, and affect the infiltration rate of water into the soil (Patle *et al.*, 2019). Coarser soils, usually have sand as a component, and tend to have higher infiltration rates, due to few, large, pore spaces, whereas fine soils, usually with clay as a component, have lower infiltration rates, but higher water holding capacity, due to smaller, but more numerous pore spaces (Adeniji *et al.,* 2013).

Another important factor, when dealing with wastewater containing pathogens, is the sorption of the microorganisms from the water column, into the soil. This creates the potential for contaminated water to pose a longer-term health risk, as the pathogens may be left behind, attached to the soil, after the floodwater withdraws. Although not an urban scenario, one study found levels of *E. coli, Salmonella,* and *Listeria,* detected via PCR, only significantly declined in the soil of a lettuce field 3 weeks after a flood (Castro-Ibáñez *et al.,* 2015). Many studies have investigated the sorption and dispersion behaviour of bacteria and viruses, with factors such as soil type, pH, porosity, temperature, and moisture all having an effect. However, many of these studies focus on organisms such as *E. coli*, *Campylobacter*, and intestinal enterococci- all traditional indicators of faecal contamination (Ten Veldhuis *et al*., 2010).Therefore less is currently known about the sorption abilities of the true range of bacterial species present at urban flood sites. As well as this, the characteristics of the pathogen can also have an effect on their motility in soils- cell type, hydrophobicity, cell shape and size (Brennan *et al.,* 2014; Gannon *et al.,* 1991; Gerba, 1983; Wang *et al.,* 2014).

This also means there is the potential for the pathogens to detach from the soil, when the water becomes saturated again, and become suspended back into the water-, which could then transport

the pathogens to wider areas. This is a specific risk, especially in parkland or recreational land, which are often used as 'flood storage' areas (Collender *et al.,* 2016).

**2.5 Movement of pathogens between the soil-floodwater interface**

Having discussed the sorption of pathogens from water and into the soil, it is important to also discuss the opposite movement of the pathogens- from the soil, to water. Soil dwelling pathogens are often split into two categories: 'soil borne' (edaphic) and soil transmitted. Soil borne pathogens (e.g., *Bacillus anthracis*), are those whose natural habitat is the soil, whereas soil transmitted pathogens (e.g. *Cryptosporidium parvum*) can live for extended lengths of time in the soil, but this is not their usual or optimal habitat- a stepping stone between infecting hosts (Jefferey and van der Putten, 2011). Most of the pathogens originating from sewage overflows, are soil transmitted pathogens, and attach to the soil particles, using floodwater as a 'vehicle' to transport the pathogen into the soil (Bett *et al.,* 2021).

Following the original flood event and attachment, if another flood or heavy rainfall event occurs- re-saturating the soil, it is possible that the attached pathogens could detach from the soil and back into the water. Generally, it is more likely that the pathogens will travel into groundwater- potentially being carried to another area (Jefferey and van der Putten, 2011). This is less likely to happen to pathogens that attach well to soil particles, for example Poliovirus. Pathogens that do not attach as readily to soil particles, such as Enteroviruses, are more likely to become mobilised once the soil becomes saturated again and be flushed further into the soil profile (Landry *et al.,* 1979). If the soil is fully saturated with contaminated water, and another intense hydrological event occurs, there is the possibility of movement of the pathogens, into the 'new' surface water.
Not all pathogens have to rely on the movement of water to carry them into the soil. Pathogens can be spread via aerosols- fine particles or droplets that are suspended in the air (Cambra-López et al., 2010), which can contain bacteria and viruses- often called 'bio-aerosols' (Islam et al., 2019) Bacteria such as *Staphylococcus aureus* and *Mycobacterium tuberculosis*, have both been known to exist , and be transported via bio-aerosols (Fennelly et al., 2004). Other pathogens have the ability to 'swim' in fluid, and usually move via chemotaxis. Moving towards areas of desirable chemistry and being repelled from areas with undesirable chemistry (Webre *et al.,* 2003).

For example, *E. coli* shows positive chemotaxis towards L-amino acids, sugars, dipeptides, pyrimidines and electron acceptors (oxygen, nitrate, and fumarate), and thus will move towards them (Colin *et al.,* 2019; Mesibov and Adler, 1972). Therefore, if 'clean' water passes over the soil surface, which is already contaminated with *E.coli*, this bacteria could move from soil to clean water, if it contained higher levels of attractive chemicals (than the saturated soil). Factors that complicate this by interfering with bacterial transfer/transport include: sorption to soil particles, hydrostatic force of the incoming water onto the soil, viscosity of the fluid/saturated soil, as well as how deep into the soil profile the pathogen exists (Kumar and Philominathan, 2010).
This is theoretical and it is hard to assume how the pathogens will behave, as there is a gap in the literature in regards to this. Especially, regarding how pathogens move between environments such as those found in an urban flood situation, e.g. between water and soil of different saturations and contamination levels.

## 2.6 Mobility and survival of specific pathogenic species typically found in flood water

### 2.6.1 Faecal indicator bacteria: *E.coli*

*E. coli* has been found by several studies to migrate from surface water into top soil layers- despite the unfavourable conditions, such as increased exposure to UV and reduced soil moisture content, (Xing *et al.,* 2017; van Elsas *et al*., 2010). For example, Jiang et al., found that after application of wastewater or manure onto a field, *E.coli* was found to survive under oligotrophic conditions, and at temperatures below 30˚C (application of wastewater or manure, on an intact sandy loam soil column, only 70cm in depth) (Jiang *et al.*, 2005).

The soil type and its physical-properties are important to determine the fate and transport of *E.coli* in soils. *E.coli* showed higher dispersivity values in gravel (1.93$cm^2$/s) as opposed to sand (0.14-0.82cm), after application of contaminated water to experimental soil columns - (Bai *et al*, 2016; Lamy *et al.,* 2013). This is thought to be due to the pore size heterogeneity, as well as the grain size of the soil - with smaller pores and smaller grain sizes (sand), retarding the dispersion of *E.coli* -. It should be noted that the distribution and variety of pore size, as well as their connectivity to each other is also an important factor that would affect bacterial transport through the soil (Kasel *et al.*, 2013; Wang *et al.*, 2014). However, this is an area that requires further study in order to draw solid conclusions (Bai *et al.,* 2016).

It has also been found that *E.coli* attach particularly well to clay soils- with a study over 8 days finding that *E.coli* survived the longest in clays (clay-loam soil with natural organic matter), with the bacterial levels (determined by measuring -ATP l) found to be almost two times that found in silts or sands (Liu *et al.,* 2017). It is thought this is due to the mineralogy of the clay, enhancing the pathogen's survival rate (Brennan *et al.*, 2014).

The minerals alter the surface area and cation exchange capacity of the soil. However, it is not yet understood how this directly helps bacterial survival. Despite this, it has been suggested that the presence of the clay minerals increases moisture and nutrient content, as well as decreasing soil particle size, which is thought to protect the bacteria from desiccation, UV light, and toxins (England *et al.*, 1993; Roper and Marshall, 1972; Rosenzweig and Stotzky, 1979).

### 2.6.2 Common pathogens associated to flooding: *Campylobacter* and *Cryptosporidium*

Much like *E.coli*, *Campylobacter* species (*C.jejuni*) has been found to be twice as likely to be present in clay soils as opposed to other soil types (Sanderson *et al*., 2018). *C.jejuni* is extremely sensitive to desiccation, high temperatures and UV damage. It is thought that clay offers more 'protection' to the bacteria due to its layered structure, confining water molecules, creating an affinity with water- thus, preventing desiccation, as well as having light scattering and light absorption properties that further protect the bacteria from UV damage (Bitton *et al.,* 1972; Fomina and Skorochod, 2020; Marshall, 1975). Clays -accumulate nutrients, reversibly bound $NH_4^+$ cations, which can then be passed to bacteria via ion exchange- therefore providing bacteria with nutrition to aid survival in the soil (Burford *et al*., 2003; Cuadros, 2017).

*Campylobacter* also tend to have higher survival rates at lower temperatures of 10-20°C, than higher temperatures (25-30°C)- with optimum temperature (10°C) increasing survival of *Campylobacter* in faeces by up to 5 days (Ahmed *et al.,* 2013; Moriarty *et al*., 2011). However, in sandy soils, *Campylobacter* attach less strongly to the soil particles- with one study finding, by measuring recovery rates through 9-10cm long, wet packed soil columns, that *C. jejuni* is transported further into soils, than *E. coli* (Bolster *et al.*, 2006). It is thought that this is due to various cell properties, such as size, shape, size distribution, surface charge, and hydrophobicity- however it is difficult to isolate and study each property and so there is uncertainty as to which cell property is most important in terms of transport through soil (Gannon *et al.,* 1991; Weiss *et al.,* 1995).

*Cryptosporidium* has proved to be difficult to treat in wastewater via usual disinfection, and so many water companies – use UV light as well as sand filtration (Rose, 1997; Timms *et al.*, 1995; West, 1991). Using packed sand has shown to significantly filter out *Cryptosporidium* oocysts, as well as *Giardia* oocysts. One study investigated the behaviour of *Cryptosporidium* oocysts in a column packed with quartz sand - although on a small scale (column height = 7.1cm, diameter = 1cm) (Tufenkji *et al.*, 2004). - This study, focused on the ionic strength of the oocysts and sand grains. Both oocysts and sand had negative charges, and thus it would be expected that they would repulse each other, creating low deposition of oocysts into the sand (Hsu *et al*., 2001). However, the opposite was found, with a significant oocyst deposition rate of ≈0.41, even at the lowest ionic strength - 1 mM KCl (Tufenkji, *et al*., 2004). It is thought that this is due to other non-ionic factors, such as grain and cell size, and how this causes straining of oocysts when the ratio of the particle/cell diameter to median grain diameter is <0.05 (Sakthivadivel, 1966). These factors together influence the straining/deposition rate of *C*. *parvum* oocysts, and other pathogens with larger cell sizes, such as *Giardia* (Tufenkji, *et al.*, 2004). This study by Tufenkji *et al*., year indicates that in soil types with fine grains (sands) the aforementioned wastewater parasites are likely to remain in the soil- however this study did not investigate the length of time the pathogens remained in the soil, or how they would behave in soil types with larger or varied grain sizes.

*Cryptosporidium* is generally associated with faecal matter, as is the case for many wastewater pathogens. *Cryptosporidium* has been shown to survive for up to 3 days in faeces (with no other moisture present) (Robertson *et al.,* 1992; Mawdsley *et al.,* 1996). In a flood situation- where the soil is saturated with water, *Cryptosporidium* can survive for much longer, and has been found to diffuse into soil, up to 30cm,(experiment carried out in clay loam, silty loam and loamy sand, over 21 days). The majority (72.8%) of the pathogen was found in the top 2cm of soil- for all soil types (Mawdsley *et al.*, 1996).

Movement of typical pathogens found in contaminated waters has been demonstrated and discussed, however less is currently known on the movement of the wider range of diverse urban floodwater species through a typical urban soil. In order to characterise the bacterial hazards at flood sites, movement behaviour of the complete bacterial community is important for determining the vector of the hazard, i.e. soil, or floodwater.

**2.7 Research Questions**

Upon reflecting on the review of the literature, a number of knowledge gaps were found in the current research. This included a lack of information regarding: the true range and diversity of

bacterial species present in urban floodwater and affected soils, whether pathogens likely to cause disease are present at an urban flood site, and whether such pathogens are 'common' flood related pathogens. Information on the movement of bacteria from urban floods between the soil water interface and into the soil profile was also lacking, and thus warranted further investigation.

Identification of knowledge gaps led to the formation of a series of research questions that could - be answered through this projects findings. A list of the aforementioned research questions have been formulated below:

- How diverse is the bacterial community found in the soil and water following a flood event, if more than just 'indicator organisms' are tested for?
- Post flood, are the bacteria present in the soil or water likely to cause disease? Is the risk higher in the short- or long term?
- Which factors (e.g. soil structure, temperature, etc) are important in determining the distribution and dynamics of bacteria after a flooding event?
- How deep into the soil profile do different bacteria originating from floodwater move during and following a flood event?

To answer these questions, this study combines field monitoring and laboratory experiments that will improve our understanding of bacterial hazards and dynamics in contaminated flooded areas.

# Chapter 3
# Methodology

## 3.1 Experimental overview

In order to investigate and meet the outlined research objectives a dual approach was taken of sampling at urban flood sites in the field, along with laboratory-based investigations. On the one hand, this approach allows the 'real-life' bacterial microbiome of urban floods to be determined temporally *in-situ*, whilst accounting for seasonal variation. On the other hand, the laboratory investigations provide a controlled environment to more closely investigate the movement of bacteria through the soil column, as well as the movement of said bacteria across and between the soil-water interface.
Detailed descriptions of the methods used for each study are described in the relevant chapters, but this chapter served as a general overview.

## 3.2 Field work overview

In order to investigate the range of bacterial pathogens often found at urban floodwater sites, two field studies were carried out. One long term study (chapter 4) and one short term study (chapter 5). The long-term study involved sampling over a one year period. This allowed the bacterial microbiome of urban flooding to be determined *in-situ*, whilst accounting for seasonal variation. Whereas the short-term study, sampling in the 3 days immediately after a flood event allowed for any shifts in the bacterial community structure to be determined and allowed for the most 'critical' time following a flood- on which day, if any, do bacteria from urban floodwater pose a potential risk- to be identified.
Previous studies tend to focus on the analysis of a small number of species-usually faecal indicators, such as *E.coli*, and coliform species (Edberg et al., 2000). Thus, the true range of the microbiological community present, is often overlooked.
Naturally, due to the nature of flood studies allowing only for 'opportunistic' sampling, collecting data is certainly more challenging. However, such data will be extremely valuable in helping to expand the data ranges available on not only bacterial diversity, but also how this can change seasonally and temporally in the long term, but also in the short-term, and how this varies over the soil-water interfaces influenced by urban flooding.
The following methods relating to field work were used for the studies outlined in chapters 4 and 5.

### 3.2.1 Sample site selection

To determine appropriate sites the Environment Agency's surface flood risk map was used in conjunction with The Rivers Trust's 'Is my river fit to play in?' map tool, along with advice from Sheffield City Council to identify sites at annual risk of surface flooding, that are also nearby to Combined Sewer Overflows (CSOs) or outlets that spill regularly (Environment Agency, 2019; The Rivers Trust, 2021). Word of mouth in the local area, as well as advice from volunteer groups such as 'Rivelin Valley Conservation Group', and 'Shiregreen Litter Pickers' were also great sources of information for the selection of potential sample sites. After preliminary visits and tests during the

first year of the study, it was decided a maximum of two sample sites was feasible, in order to ensure that sampling and analysis were conducted in an accurate and reproducible way, without the deterioration of samples. As a consequence of this, and after initial visits, two sampling sites were chosen in Sheffield, for routine sampling.

*Tongue Gutter*

This site is located by a river – Hartley Brook Dike - (53°25'43.3"N 1°28'08.1"W) , (also known as Sheffield Lane Dike and Tongue Gutter), which runs parallel to residential housing areas, and feeds into Blackburn Brook, a stream that eventually meets the River Don. Alongside the river and sample point, a public footpath is also frequented by dog walkers and cyclists. At the sample site there is an outfall pipe discharging into a creek/brook. On every visit to the site the outfall pipe was consistently discharging and a strong smell was present, independently to the presence of rainfall The site had a lot of aesthetic pollution, some of which appeared to flow out of the outfall pipe, while another source of aesthetic pollution was via fly tipping from members of the public (Figure 1).

At Tongue Gutter, separate sewer systems are used in the homes nearby. Misconnections to systems like this have previously been reported as the source of contamination (Chandler and Lerner, 2015), and could be the case at this site. This would mean foul water was connected to a surface water sewer system, and thus wastewater may be discharged via the surface water outlet. On several site visits the discharging water appeared frothy and bubbly, which could have been an indicator that the water discharging was mixed with a soap or detergent- supporting the theory of misconnections. This site has a surface flood risk categorised as 'high', and so has an annual risk of surface floods of <3.3% (Environment Agency, 2019).

**Figure 1.** Sampling site at Tongue Gutter on a visit in May 2022.

*Endcliffe Park*

The second sample site is a large city park, Endcliffe Park, which lies along the course of the Porter Brook River – (53°22'06.4"N 1°30'21.0"W). The area consists of grassy recreational parkland, along with woodland- with the chosen sample site being located in the former. The sample location is an area that frequently has a large pool of standing water- especially visible after rain. The site is also close to two gullies-which are thought to contribute to the water pool, with water being seen flowing out of one several times after rainfall.

This site is particularly interesting due to the proximity to recreational ground as well as a children's playground (Figure 2). During particularly heavy rain, the pool of potentially contaminated water has been known to reach the playground itself. Similarly to site 1, this site has a surface flood risk categorised as 'high', and so has an annual risk of surface floods of <3.3% (Environment Agency, 2019).

**Figure 2.** Sampling site at Endcliffe Park during a visit after heavy rain in February 2022.

3.2.2 Sampling regime

On each visit, triplicate soil samples were taken, from 0-10cm depth, at both sites using a basic handheld soil auger, as well as control (background) soil samples located close by to the sample site but in an area not known to flood. The sampling frequency is outlined in chapters 4 and 5. Samples were placed in Ziploc plastic bags and sealed to prevent drying-out. The samples were stored in the dark and transported within 4 hours of collection to the laboratory and stored at 4 ºC (again in the dark) until analysis was undertaken. This minimised any potential changes in the microbial communities in the soil samples. For the molecular genomic work, 50 g sub-samples of soil were sealed in small plastic bags and were preserved at –20 ºC for DNA analysis.

Flood water samples were also taken in triplicate taken using 1L polyethylene containers and 50ml falcon tubes. Again, like with the soil samples, samples were stored in the dark and transported within 4 hours of collection to the laboratory and stored at 4 ºC until analysis was undertaken.

3.2.3 Rainfall Monitoring

To understand the influence of precipitation on bacterial dynamics, rainfall was monitored using a remote GPRS rain gauge located less than 5 miles from both sites (Detectronic, CIVICA, United Kingdom). The gauge automatically logs rainfall data every hour. This gauge uses a 'tipping bucket' design accompanied by a GPRS data logger (Detectronic, CIVICA, United Kingdom). This rain gauge complies with the 'Guide to Meteorological Instruments No. 8 of the WMO', has an unlimited capacity and has an accuracy of +/- 1% at 26mm/hour (Detectronic, 2020; WMO, 2020).

**3.3 Sample analysis specific to field studies**

The following sample analyses were performed for samples from field sites only- Chapters 4 and 5.

### 3.3.1 pH and Temperature

The first parameters measured for every soil and water sample were pH and temperature. This was done in the field using a portable probe (Sension+ PH1 Portable pH Meter, Field Kit with Electrode for General Purposes, Hach, UK), taking measurements in triplicate. For pH and temperature measurements of water in the field, a 500ml beaker of water was taken from the source, the mixture was vigorously mixed and the probe inserted for *in situ* measurements in triplicate.
The pH of the soil is best measured *in situ* to avoid any potential changes in sediment which may occur during transport to the laboratory (Gumbley *et al*., 2005). To measure the pH of soil samples, a slurry was created at the field site by mixing the soil sample with distilled water using a ratio of 2.5:1, soil sample: distilled water (Rowell, 2014). Temperature was not taken for soil samples.

### 3.3.2 Determining soil moisture content and organic content

To measure the moisture content of the soil samples, 30g of sample from each site was dried in an oven at 105˚C, for 4 hours. The samples were weighed before and after drying and the difference in weight was equal to the % moisture of the sample (BS1377 PART 2:3.2) (British Standards Institute, 1990).
The same process was then repeated in a 550˚C furnace, to determine the loss on ignition %/organic matter % of the soil samples. The following calculations were used to measure the moisture and organic matter content of the samples:
Moisture content:
% of fresh weight = ([mass fresh soil- mass dry soil]/mass fresh soil) x 100

Organic matter content:

% of dry weight = ([mass dry soil- mass of ash after ignition]/ mass dry soil) x 100

### 3.3.3 Soil particle size analysis and classification from field samples

After drying, subsamples from each site were pooled and passed through a set of 10 standard geotechnical sieves, from 5mm to 63 micron (Appendix 1). The initial weight of the dry sample was measured, and the stack of sieves shaken on a sieve shaker for 20 minutes (British Standards Institute, 1990). After this, the amount of sample retained on each sieve was weighed and recorded. The results of dry sieving were then plotted on a particle size distribution chart according to British Standard 1377 Part 2, and compared to a soil sizing chart to determine soil type (Appendices 2 and 3)(British Standards Institute, 1990).

### 3.3.4 Scanning Electron Microscopy (SEM) to characterise soil microbiome

Soil samples were taken on a day without rainfall at the 'Endcliffe Park' site and visualisations created using SEM. This was to give a detailed, and magnified view of the typical soil structure- including any microbes that are present under 'usual' (non-flood) conditions. This process was carried out at the University of Sheffield's Electron Microscopy Service. The following steps were carried out:

Specimens (each specimen=0.5g of soil) were fixed in 2.5% Glutaraldehyde (Merck, Germany) in 0.1M phosphate buffer (Merck, Germany) for 3 hours at 4˚C. Specimens were then washed in 0.1M Sodium Cacodylate buffer (Merck, Germany, three times, with 30 minute intervals at 4˚C. Next, secondary fixation in 2% aqueous osmium tetroxide (Merck, Germany) was carried out for 1 hour at room temperature. The specimens were then dehydrated through a graded series of ethanol (Fisher Scientific, UK):

- 75% ethanol for 15 minutes
- 95% ethanol for 15 minutes
- 100% ethanol for 15 minutes
- 100% ethanol again for 15 minutes
- 100% ethanol dried over anhydrous Copper sulphate (Merck, Germany) for 15 minutes
- 50/50 mixture of 100% ethanol and 100% Hexamethyldisilazane (Merck, Germany) for 30 minutes
- 100% Hexamethyldisilazane for 30 minutes.

Then, specimens were air-dried overnight. When dry, specimens were mounted on 12.5mm diameter stubs, attached with Sticky Tabs, and coated with approximately 25nm of gold in an Edwards S150B sputter coater (Edwards, UK). Specimens were finally, imaged using a Tescan Vega3 LMU SEM (Tescan, Czech Republic) at an accelerating voltage of 15Kv.

### 3.4 Laboratory based - Soil Column investigation

In order to analyse the movement of bacteria originating from floodwater across the soil-water interface and into the soil profile, a column experiment was designed. This was designed to replicate a small section of soil from the urban environment and subject it to 'urban flood' conditions, by running contaminated water through the column, taking water samples temporally and soil samples at different depths once the floodwater had run through the column. This allowed changes in behaviour and distribution of present pathogens to be tracked.

Prior to the experiments using floodwater, preliminary salt tracer tests were also carried out, with data fitted to an advection-diffusion model (ADE model). This allowed the appropriate flow rates to use to be determined, as well as finding the diffusivity value of the soil, and also determine when would be best to take samples from the column to track potential movement of pathogens (via water) through the column. Full details of the preliminary tests can be found in Chapter 6.

#### 3.4.1 Column Design

It was not only important that the column depth represented field conditions and spatial/ temporal scales relevant to urban flood events, but it was also important that the depth of the column was not too large, as practicalities, such as taking samples, and filling and emptying the column may have become too difficult. Although many soil column studies have been carried out, there is no standardised method for designing or running a soil column experiment (Lewis, J.and Sjöstrom, J., 2010). A column made of inert acrylic was designed with a depth of 1.28 metres and an inner diameter of 24cm (Figure 3).In depth details of column design, soil choice, and set-up can be found in Chapter 6.

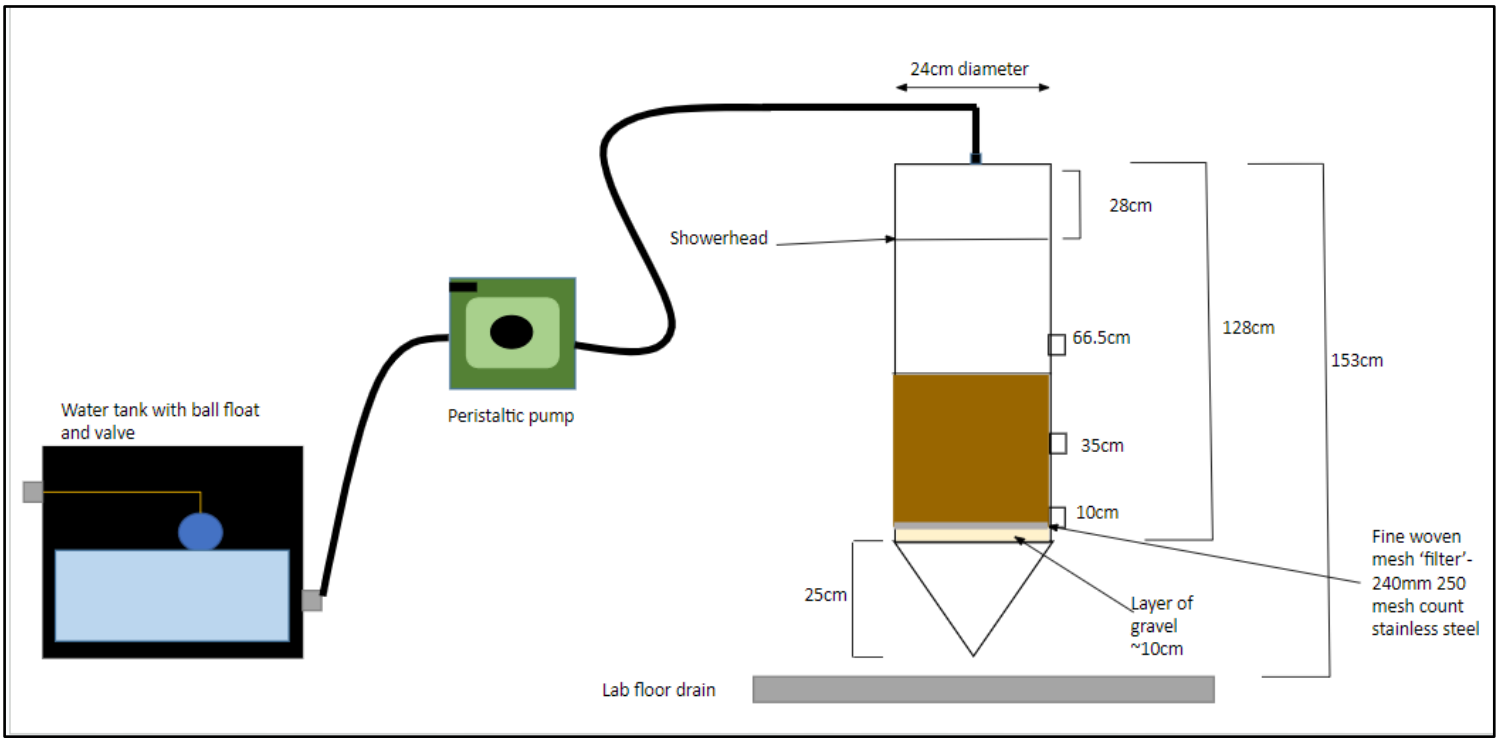


**Figure 3**. Illustrative diagram of column set up, showing sampling ports at different depths

Instead of creating synthetic floodwater, floodwater was taken from the field to use in the column experiments. Flood water was taken from the site 'Tongue Gutter', this was due to the fact the above mentioned pipe at this site spilled almost constantly and so it was guaranteed there would be enough water to collect. This site was also known to be contaminated with faecal bacteria, thus providing a diverse bacterial community to study. On the morning of each soil trial, approximately 100 L of 'floodwater'- water originating from the pipe and forming a pool (Figure 1)- was collected in sterile containers and transported to the laboratory within 30 minutes. Water was filtered through a 240mm 250 mesh count stainless steel sheet to remove any large pieces of debris to prevent any blockages, before being used in the column experiments.

### 3.4.2 Flood Water Trials

3 soil column trials were carried out using flood water collected from the field. Each trial was set up in the same way, and conditions for each trial replicated as closely as possible. Soil was hand-packed into the column prior to each trial. This generally involved: sieving the soil into the column in 10cm increments, which were then compacted with a weighted 'tamper' device (further details can be found in 'Chapter 6'). Care was taken to ensure the same packing process was carried out for each trial for repeatability, however due to re-packing the column, it was expected there would be some differences in results between the trials, as it was not possible for each 'new' column to be precisely the same as the others. Re-packing of the soil results in a loss of the natural soil structure, which in turn can affect the distribution of soil properties and therefore the transport of the floodwater and thus bacteria.

For each trial, after saturation of column (and leaving to settle overnight), clean water was pumped into the column, using the peristaltic pump until a head depth of between 7cm and 13cm was reached with the outflow tap open. The head was maintained by a constant inflow of water, and the open port. Once the correct flow rate was selected on the pump to ensure a steady head depth was maintained, the peristaltic pump began pumping the flood water collected from the field site into the column. In terms of 'trial duration', the trial began once the pump switched to pumping flood water.

Water samples were then taken every 30 minutes or one hour ,from the outflow at the bottom of the column. Some trials were shorter in duration than others due to practicalities such as the lab closing for the day, or health and safety risks i.e working alone out of hours. Water was collected into a sterile 15ml falcon tube, with at least 10ml collected per sample and immediately transferred to a refrigerator (4℃).

Soil samples were taken from the column ~12 hours after the trial ended. This allowed the floodwater to naturally drain from the column, and so soil samples could be taken without disturbing the soil column, or creating 'holes' which would affect the flow path of the water while the trial was running. Soil samples were taken from each trial at the following depths: soil surface (top 10 cm), 40cm (from bottom of column), 33cm, 25cm, 17cm, mesh (10cm sitting on top of mesh). Each soil sample was extracted using the corer (figure 6) and transferred to a sterile 50ml falcon tube and placed into a refrigerator immediately (4℃). Each soil sample weighed at least 10g.

For each trial the following pre-trial samples were also taken: clean water (used to saturate soil), floodwater, and a pre-trial soil sample (before being packed into the column).

**All analysis methods outlined from this point were utilised in all 3 studies- Chapters 4, 5, and 6.**

### 3.5 Bacterial abundance: Heterotrophic and coliform plate counts

Microbial abundance can be investigated by culturing on selective nutrient agar plates. This process – heterotrophic plate counts (HPC)- allows any cultivable, heterotrophic bacteria (bacteria that require organic nutrients for growth) present in the samples to be determined and was performed for all soil and water samples. It should be noted that HPC only gives information on a small part of the microbial community. However, this method is widely accepted and used in water industries (Allen et al., 2004) , due to the low cost, and ease of use, and so is used for this study.
Before plating, seral dilutions were performed for soil and water samples. However, a slurry was made for soil samples using 1g of soil and 10ml of PBS- Phosphate buffered saline (Invitrogen™ PBS (Phosphate-Buffered Saline) Tablets, ThermoFisher Scientific, United Kingdom).

After dilution, 0.1ml from each sample tube was pipetted onto plates containing LB Agar (Invitrogen™ LB Agar, powder (Lennox L agar) ThermoFisher Scientific, United Kingdom), as well as plates containing Chromocult Coliform Agar (Chromocult® Coliform Agar, Merck, Germany). Preliminary tests determined that a dilution factor of $10^{-3}$ would be used for samples plated on LB Agar, and a dilution factor of 10-1 would be used for samples plated on Chromocult Coliform Agar.

This was carried out in a laminar flow (SafeFAST Classic class II A1/A2 -212, Microbiological Safety Cabinet, FASTER SRL, Ferrara, Italy), in order to avoid contamination of samples from airborne bacteria, and to protect lab users from any potentially infectious bacteria. The plates were then inverted and left to incubate at 37˚C for 24 hours.
Chromocult coliform media is selective, and so it signifies the presence of just two bacteria- *Escherichia coli* (*E. coli*) and *Citrobacter freundii* (*C. freundii*) - both faecal indicators. On Chromocult, *E. coli* colonies appear purple and *C. freundii* appear as pink colonies (Figure 4).

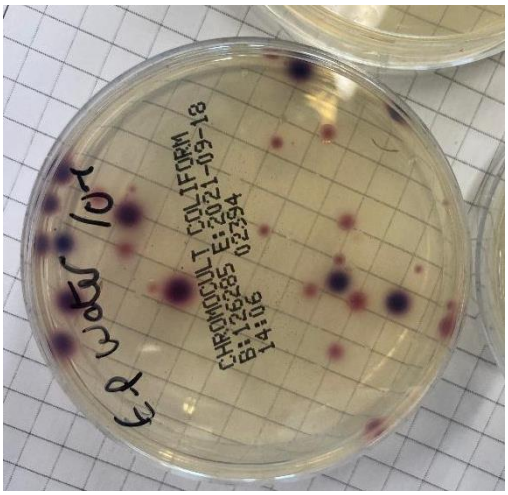

**Figure 4.** Colonies of *E. coli* (purple) and *C. freundii* (pink) on a chromocult agar plate.

After incubation, the colonies on each plate were counted. Once the colony forming units (CFU) have been counted, the number of CFUs per millilitre (CFU/ml) of the original sample was calculated by multiplying the number of CFU's counted, by the dilution factor. This method gives a general estimation of the concentration of heterotrophic bacteria in a sample.

**3.6 Molecular genomic methods**

The use of these molecular techniques allows the characterisation of those components of the soil microbial community that currently cannot be cultured (Hosokawa et al., 2022).

3.6.1 DNA extraction- Soil samples

DNA was extracted from soil samples using the PowerSoil® DNA Isolation Kit (Mo Bio Labs Inc.). The kit utilises the bead beating method. The 'Experienced User' protocol included in this kit was followed (Mo Bio, 2016).

3.6.2 DNA Extraction- Water

*3.6.2.1 Filtration and cell concentration of water samples*

To extract DNA from water samples, the samples first require filtration. This concentrates any microbial cells in the water onto the filter. To do this a three-way vacuum filtration system was used (Sartorius 3 Branch Microsart™ Manifolds with Funnels, Sartorius, Germany) with nitrocellulose filters with a pore size of 0.2µm and 47mm diameter (MF-Millipore® Membrane Filter, 0.22 µm pore

size, 47 mm diameter, mixed cellulose esters (MCE) membrane, Merck, Germany). This pore size is required as most bacteria and pathogens are much larger, at 0.5-5µm (Levin and Angert, 2015) and so will not pass through the filter, but remain on the filter for DNA extraction.
This method prepares water samples for DNA extraction by trapping the microbial cells on the filter. The extracted DNA will be used for PCR, as well as for sequencing. After 500ml of each water sample has been filtered, the filters were stored in the freezer at -20˚C until DNA is extracted.

For the water samples, the Cetyltrimethylammonium Bromide (CTAB) method was used as described by Karunakaran *et al.,* 2016, and is outlined here. Filters were cut into strips using a sterile scalpel, and half of each filter was placed into 2ml Eppendorf tubes. Then, 740 µL of SET lysis buffer (SET lysis buffer consists of 40 mM EDTA (Ambion, Warrington, UK), 50mM Tris-HCl pH 9 and 0.75 M sucrose (Sigma Aldrich Co., UK.) ) and 90 ml of lysozyme 10 mg/mL (Sigma Aldrich Co.,UK) were added. Tubes were incubated at 37 °C for 30 minutes with rotation in a hybridisation oven (Thermo Scientific, UK). After incubation 90 mL of 10 % sodium dodecyl sulphate (SDS) (Sigma Aldrich Co., UK.) and 25 mL of proteinase K 20 mg/mL (Applied Biosystems, Life Technologies Ltd., UK) were added. Tubes were then incubated at 55 °C in the hybridisation oven for 2 hours with rotation.After the 55 ˚C incubation, any supernatant was withdrawn and transferred into 2 mL sterile Eppendorf tube. 137 µL 5M NaCl 1 % and 115 µL Hexadecyltmethyl ammonium bromide (CTAB) / NaCl solution, were added to the tubes containing the supernatant (Sigma Aldrich Co., UK). Tubes were then incubated at 65 °C for 30 minutes with rotation.

After incubation, 838 µL of chloroform:isoamyl alcohol (24:1) (Sigma-Aldrich, UK) was added in a fume cupboard and the tubes then centrifuged at 13,000 RPM for 5 minutes. After centrifugation, the supernatant was extracted into a 2 mL tube and steps 9 and 10 were repeated.
DNA was then precipitated, by adding 815 µL of 100 % isopropanol (Sigma-Aldrich, UK). The tubes were then left overnight at -20˚C.

The next day, samples were centrifuged for ten minutes at maximum RPM. After centrifugation, the supernatant was removed carefully, leaving the DNA pellet in the tube. The pellet was then washed twice with 1 mL of 70 % ethanol, centrifuged for 5 minutes and the ethanol decanted.
Tubes containing the pellet were then air dried in a fume cupboard for approximately 2 hours (until ethanol had dissipated) and then suspended in 50 µL DEP-treated sterile water (Thermo Scientific, UK).

#### *3.6.2.2 Quantification of extracted DNA*

The amount of DNA in each sample was quantified using a QuBit™ 4 Fluorometer (Thermo Fisher Scientific, UK), following manufacturer's instructions. The 'DNA Broad Sensitivity Kit' as well as the 'DNA High sensitivity Kit' were used to ensure all ranges of concentrations of DNA were measured from 0.2ng/µl -1000ng/µl (Thermo Fisher Scientific, 2021).

### 3.7 Sequencing the bacterial 16s rRNA gene

Next generation sequencing was performed at the NEOF molecular facility at the University of Liverpool, targeting the 16s rRNA gene, fourth hypervariable (V4) region. Hypervariable regions

show different sequence diversity between various species of bacteria- and thus can be used for species identification. The V4 region, specifically has high functionality in the ribosome and is highly conserved (Morosyuk et al., 2017; Van de Peer, et al., 1996).
The Illumina Platform was used - for sequencing a of the 16s rRNA gene to estimate bacterial community structure. The samples in this study were sequenced using Illumina MiSeq® pair-end technology (Illumina, United States). Illumina MiSeq® has been chosen as it has a sequencing platform that provides a maximum of 250 million reads of lengths up to 2 × 300 base pairs (Caporaso *et al.,* 2012). To target the 16S rRNA gene V4 region (16S v4), the following forward and reverse primers were used. Portion of the sequence in **bold** indicates the recognition sequence, allowing for a secondary nested PCR process to be performed.

F: 5'**ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNN**GTGCCAGCMGCCGCGGTAA3'
R: 5'**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**GGACTACHVGGGTWTCTAAT3'

A successful amplification product results in a 254bp insert as described by Caporaso et al (2011 PNAS 108, 4516-4522).

3.7.1 Test Polymerase Chain Reaction (PCR) of extracted DNA

After DNA extraction and quantification, a test PCR was performed on extracted samples in the NEOF visitor lab in Sheffield, prior to send the DNA for sequencing at the NEOF lab in Liverpool. This test was performed to check that the quantity and quality of the extracted DNA would allow for the subsequent amplifications needed for sequencing. For this PCR the following controls were included. An extraction kit negative control- this was a mock nucleic acid extraction, using the reagents provided in the extraction kit, but without any actual sample. An extraction kit positive control- this was an extraction from a community of cultured *E. coli* to ensure extraction kit accuracy. A PCR negative control was also used. This was a no-template control, which omits any DNA or RNA template from the PCR reaction. This serves as a general control for nucleic acid contamination. A PCR positive control was also used, again using *E. coli.*

The test PCR used the following conditions: Initial denaturation was at 98 °C for 2 mins. Then 10 cycles of: denaturation 95 ˚C 20 secs, annealing 65 ˚C 15 secs and extension 70 ˚C 30 secs. With a final extension step of 72 ˚C for 5 min. The PCR mixture contained 5 µL of template DNA, 10 µL 2x Kapa HiFi HotStart Master Mix (Roche, Switzerland), 0.25 µL of 10 µM of each primer, and molecular biology grade water to a final volume of 20 µL.
The size of the PCR amplicons was then checked using gel electrophoresis.

3.7.2 Gel electrophoresis to check DNA size after test PCR

To check the size of the tests PCR amplicons the agarose gel electrophoresis method was used. A 0.8% agarose gel was made with agarose powder (Sigma Aldrich, United Kingdom) and Tris/Borate/EDTA (TBE) buffer, diluted X10 (Thermo Scientific, United Kingdom). 5ul Ethidium bromide (Thermo Fisher Scientific, 2018) was used to stain the gel and to allow for DNA visualisation. Orange G (Thermo Fisher Scientific, United Kingdom) at X6 dilution was used as a loading/tracking dye, and 1X TBE (Merck, Germany) buffer was used in the agarose gel box. In this case a 1KB DNA

ladder was used (1kb Hyperladder™, Meridian Bioscience, United States) to check for amplicon size, with gels ran at a voltage of 100V for approximately one hour. A trans-illuminator was then used to visualise the gel and DNA bands using UV.

### 3.7.3 Illumina sequencing at NEOF molecular facility in Liverpool

The NEOF Liverpool facility uses a 16s rRNA dual index nested PCR for Illumina's sequencing (NEOF, 2024). The 1st PCR amplification uses the primers and conditions described previously in Section 4.6.1 (Caporaso et al 2011,). This primer design incorporates a recognition sequence to allow a secondary nested PCR process. The PCR products were then purified with AMpure beads using a 1:1 ratio (Beckman Coulter, USA). The second PCR is then performed to incorporate Illumina adapter sequences to allow for the sequencing of samples on the Illumina MiSeq platform. Barcodes for sample identification are also incorporated at this point.

The general sequences of the forward and reverse primers for the second PCR are illustrated below. The 8bp barcode for sample identification after sequencing is underlined.

**N501 f:**

5'AATGATACGGCGACCACCGAGATCTACACTAGATCGCACACTCTTTCCCTACACGACGCTC 3'

**N701 r:**

5'CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTGACTGGAGTTCAGACGTGTGCTC 3'

The secondary PCR was performed under the same cycling conditions as described for the first PCR, but for 15 cycles. PCR mixture contained 10 µL of amplicon from 1st PCR, 10 µL 2x Kapa Master Mix, 0.5 µL of 10 µM of each primer, and molecular biology grade water to a final volume of 20µL. Secondary amplicons were purified with AMpure beads (Beckman Coulter, USA ), as before.

### 3.7.4 Bioinformatics: DNA sequences analysis

Firstly, the raw Fastq files were trimmed for the presence of Illumina adapter sequences using Cutadapt version 1.2.1 (Martin, 2011). The option -O 3 was used, so the 3' end of any reads which match the adapter sequence for 3 bp. or more were trimmed. The reads were then further trimmed using Sickle version 1.200 with a minimum window quality score of 20 (Joshi and Fass, 2011). Reads shorter than 15 bp. after trimming were removed. These trimming steps were performed at the NEOF Liverpool molecular facility (Figure 5).

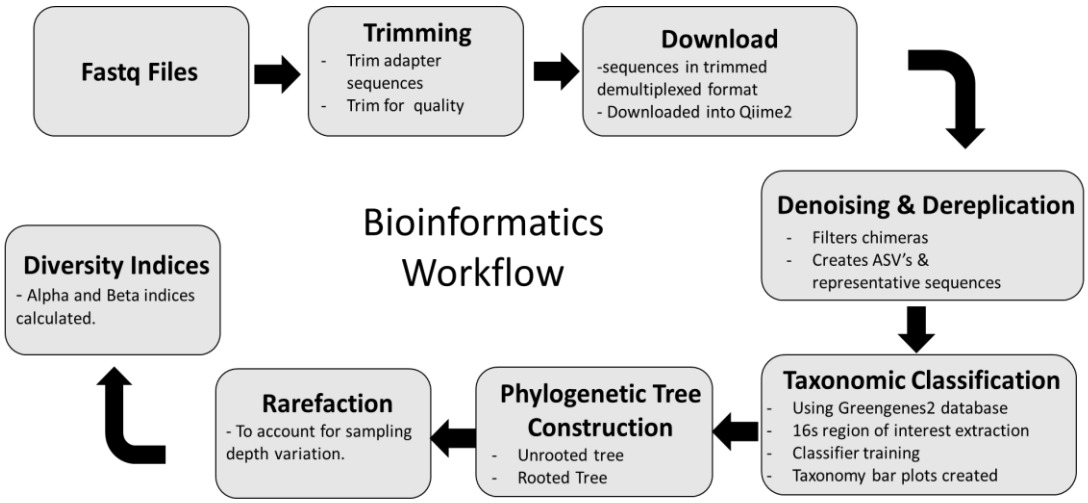


**Figure 5**. Pipeline of workflow used for bioinformatics analysis.

After downloading the demultiplexed trimmed sequences, the data was imported into the Quantitative Insights into Microbial Ecology 2 program in the Paired End Fastq Manifest Phred33 format (QIIME2, version 2019.7) (Bolyen *et al.,* 2019). This means the sequence files were imported with a tab delimited 'Manifest' file which maps sample identifiers to absolute filepaths that contain sequence and quality data for the sample. After sequences were imported, they could be visualised in Qiime2 View (Bolyen *et al.,* 2019). This gave information on the quality of forward and reverse reads using a random sampling of 10000 reads, as well as showing a summary of sequence lengths. This information was important for the next step.

Next, De-novo amplicon sequence variants (ASVs) were identified using the 'DADA2' command. This method is based on sequencing error correction algorithms and generates ASVs, rather than using clustering methods, which typically generate individual operational taxonomic units (OTUs).
This step denoises paired-end sequences, dereplicates them, and filters chimeras. For this step chosen truncation length was specific to each study, and is outlined in the relevant chapter. This step provided an ASV abundance table, along with representative sequences.

With the sequence variants defined, the next step was to carry out taxonomic classification using Greegenes2 version 2022.10 (Mcdonald et al., 2023). The full length 16S rRNA sequences that are used within the Greenegenes2 2022.10 phylogenetic tree were downloaded into QIIME2. After the database was downloaded the 16s rRNA region of interest was extracted. This step is required to reach optimum assignment performance. Next the classifier was trained on the set of extracted reads using the Naive Bayes algorithm. Then taxonomic assignments were assigned to the ASV representative sequences using the trained classifier at a confidence level of 0.7. This meant each taxonomic level must reach a confidence level of <= 0.7 to be classified. At this point bar plots of taxonomy could be visualised in Qiime2, and exported for each sample.

After sequences were assigned to taxonomic levels, a phylogenetic tree was constructed. This allowed the computation of phylogenetically aware beta-diversity metrics. This involved creating first an un-rooted tree, and then a rooted tree in Qiime2. The detected diversity partly depends on

the sampling depths. An incompletely sampled community will appear less diverse than a fully sampled community. Therefore, to compare the diversity between samples different sampling depths needed to be accounted for. This was carried out with rarefaction. Rarefaction is performed by reducing all the samples to the same sequence count, by random sub-sampling of the sequences. The minimum and maximum sampling depths were set again, and selected to ensure sequencing depth within samples were retained and no entire sample was lost.

Finally diversity indices were calculated for alpha and beta diversity. This was done by using the diversity 'wrapper' script' in QIIME2. The script co-investigates alpha- and beta-diversity at the same time. This gave multiple metrics for both alpha and beta diversity.

### 3.8 Diversity Indices

To fully investigate the bacterial diversity within the samples both alpha and beta diversity metrics were calculated in Qiime2.

#### 3.8.1 Alpha Diversity

Alpha Diversity is used to measure diversity on a 'local' scale- 'within' sample diversity (Whittaker, 1960) .In this study, 3 different measures were used to determine alpha diversity: Faith's phylogenetic diversity (PD) Index, Abundance based coverage estimator (ACE) Richness index, and Simpson's evenness index (Chao and Lee, 1992; Faith, 1992; Simpson, 1949). Faith's PD index uses phylogenetic distance to calculate the diversity of a given sample (Faith, 1992), while the ACE Richness is a nonparametric richness index that measures the number of unique species, using sample coverage, defined based on the sum of the probabilities of the observed species (Chao and Lee, 1992). Simpson's evenness index is a measure of diversity which considers the number of species present, as well as the relative abundance of each species (Simpson, 1949). Each diversity index is based on different assumptions about the species diversity and thus, by using a combination of several indices, a more robust view of the biodiversity can be gained.

#### 3.8.2 Beta Diversity

Beta Diversity is described as the ratio between regional and local species diversity (Whittaker, 1960), and is useful to compare differences in diversity between different groups, i.e. different sample sites or types. To calculate beta diversity in this study, weighted Unifrac distance analysis was used. There are two types of Unifrac distances: an unweighted Unifrac distance that considers only species presence and absence and counts the fraction of branch length unique to each community. The weighted method, used here, uses species abundance and also weights the branch length with abundance difference (Chen *et al.,* 2012). This phylogenetically aware metric takes into account presence and absence of species, as well as abundance- and is thus quantitative.

### 3.9 Statistical Analysis

Different analyses were used, tailored to the findings of each study and are outlined in detail in the relevant chapter. However, a brief description can be found below.

To test for differences generally the independent t-test was used. This test determines whether there is a statistically significant difference between the means in two groups (Mcdonald, 2014). To test for correlations, the Pearson's correlation coefficient test was used. This method calculates the ratio between the covariance of two variables, and takes into account the standard deviations (Kirch, 2008).

For Beta Diversity analysis, Principal Coordinate Analysis (PCoA) was used. This method converts data on Euclidean distances between items into a 3 dimensional scatter graph plot. The plot has 3 axis, one for each 'dimension'. The value displayed on each axis is called the 'eigenvalue'. Axes are ranked by their eigenvalues, and so the first axis has the highest eigenvalue and thus explains the most variance. Each eigenvalue represents the amount of variance in each dimension as a percentage of the total variation in the distance matrix, and so the sum of the eigenvalues will equal the sum of the variance of all variables in the data set (Bakker, 2024).

The plot can be viewed, and shows information on the 'closeness' of items, allowing groups and clusters to be found. In Qiime2, this utilised a distance matrix of beta diversity values for each sample. This allowed data to be separated based on metadata groups, i.e. sample type, location, date, rainfall amount etc. This was mainly used in this study to visualise beta diversity data, and to determine if separation of samples had statistical significance or if it might have originated by chance, Permutational Multivariate Analysis of Variance (PERMANOVA) analysis was carried out. PERMANOVA is a non-parametric multivariate statistical permutation test. It uses a distance matrix constructed from a dissimilarity measure, and tests the null hypothesis that there are no differences in the relative magnitude (presence/absence) of a set of variables among objects from different treatments or groups (Anderson, 2017). In this circumstance the dissimilarity measure is that computed using the beta diversity distance matrix to determine differences in beta diversity between the metadata groups. This test was chosen here as its assumptions are negligible, and it works well with complex data/ many groups of data.

# Chapter 4
# Investigating the bacterial communities of urban flood water, and the impacts on soils at urban flood sites

## 4.1 Abstract

Sewage overflows and surface run off contributing to urban floods can act as vectors for the distribution of pathogens, known to cause disease among human populations. Most previous studies concerning waterborne health risk, have focused on using faecal indicators due to this being affordable and easier to perform than other methods (Odonkor and Ampofo, 2013, Yard et al., 2014). However, traditional indicators do not accurately reflect the true risk that urban flooding poses this is because indicators only act as an estimate for faecal contamination and thus potential - pathogens presence, and do not provide information on the specific microbial communities that might be present. There is a current lack of understanding regarding specific pathogenic species present during and after flood events, - which species are present, how many are present, and whether they travel between flood water and urban surfaces and soils. A more specific and sensitive approach is required and so this study utilised routine sampling in the field to track seasonal changes at flood prone sites and cover all weather conditions, along with heterotrophic plate counts to estimate the numbers of bacteria present, and next generation sequencing of the 16s rRNA gene to determine the bacterial community structure up to species level. Physico -chemical factors were also measured to determine whether they had any influence on the bacterial community structure. This study found that urban floodwater contained species known to be associated with the human intestinal tract, along with pathogenic species, and those that can cause disease- including *Arcobacter cryaerophilus*, and bacteria of the genus *Bacteroides*. The composition of the bacterial community structure was dependant on the origin of the floodwater. At the study sites it was also found that the presence of pathogenic bacterial species was mostly independent of any measured physico-chemical factors- including rainfall , however at one sample site bacterial diversity of floodwater did seem to increase when a higher rainfall depth was observed in the preceding days. Many bacterial species were shared by both the floodwater and soil at flood sites, indicating that certain species may move from the floodwater into the soil, via infiltration of water into soil, transporting the bacteria. This research has significantly helped to expand the data ranges available on not only microbial diversity in urban floodwater, but also on the affected soils. This data also provides information on how the movement of bacteria between the soil and water can have consequences for the microbial risk posed by an urban flood area in terms of the risk of disease, either from the floodwater or contaminated soils. This study also suggests that thought should be given to other environmental factors, including the source of the flood, as this can determine whether pathogenic or disease causing bacteria are likely to be present.

**Keywords**: climate change, pathogens, urban floods, sequencing.

## 4.2 Introduction

In recent years, the frequency and occurrence of extreme hydrological events – such as intense precipitation, has been increasing (Pall et al., 2011). Precipitation events have increased in intensity - volume, as well as frequency in some cases, consequently leading to an increase in flood events (Waters et al., 2010). Many sewer systems in densely populated urban areas are becoming unable to cope with sudden influxes of rainfall runoff, resulting from intense precipitation. (National Infrastructure Commission, 2022). Combined sewer systems - the main sewer type used in most UK towns and cities - are designed to discharge diluted wastewater, storm water, and sewage usually, via a 'release valve'. This is a combined sewer outflow (CSO). The discharge flows out into a river or water body, when the sewer becomes 'full' to prevent damage to the sewer infrastructure, as well as to prevent uncontrolled flooding of homes and buildings (Phillips et al., 2012).

It should be noted that CSO's are not the only system that discharges water via outlets. Surface water outflow pipes also discharge rainwater that runs off roads, gutters, fields, etc. into rivers, water bodies, and into groundwater via soakaways. Surface water systems discharge regularly, as the discharge is rainwater runoff, and so it is not usually thought of as harmful, in terms of the risk it poses to public health (Webber et al., 2018). However, in some cases, misconnection of pipes can also lead to wastewater being discharged, along with surface runoff rainwater (Chandler and Lerner, 2015).

Many studies have investigated how CSO spills effect the ecology of rivers (Veronesi et al., 2013; Wang, 2014; Wu et al., 2019), as well as the immediate risk posed to the public using the river recreationally (de Man et al., 2013; Gray, 2008; Ten Veldhuis et al., 2010). However contaminated water can also end up discharging into the urban environment – before it reaches the 'release valve'- due to misconnections, old/deteriorating infrastructure, blockages or from the lack of local hydraulic capacity in the pipe network under intense rainfall, leading to an urban flood or exceedance event. This may result in contaminated water surcharging from manholes and gulleys onto streets, roads, and recreational areas (Abdellatif et al., 2013; NASEM, 2019; Olds et al., 2018). Although more infrequent, flood water can also reach people's homes – presenting a more immediate risk of contact with contaminated water. As urban areas tend to comprise a variety of impermeable surfaces, as well as permeable but easily saturated areas, such as parkland, the excess water cannot be easily absorbed into soil. On the one hand, the permeable properties of such areas can work favourably as a designated area to saturate and store flood water, preventing damage to homes and buildings (Förster et al., 2008; Schultz, 2006). However, when this is not the intention for the area, the water can sit in the area for days or even several weeks, potentially posing a microbiological risk (Gholami et al., 2010; Serinaldi et al., 2018).

Further, an urban flood event cannot only be physically, economically, and environmentally damaging in the short term– but can also pose a more long term risk due to the potentially dangerous contaminants that use the flood water as a vehicle: metals, chemicals, and disease causing pathogens (Gray, 2008; Kundzewicz et al., 2010). Toxic metals and organic chemicals, originating from industrial, domestic, and agricultural practices are often found in sediments and soils that act as a 'sink' – accumulating and storing the chemicals, and often 're-suspending' or 're-releasing' them during heavy rainfall (Gaw et al., 2014 ; Lara-Martin et al., 2014; Liang et al., 2013 ). Therefore, it is not surprising that during a flood event it has been shown that the typical chemical load (nitrates, zinc, arsenic, and iron) in surface water increases temporarily, but significantly (Yard

et al., 2014). Contact with such metals and chemicals can cause dermatological, digestive, and respiratory irritation, as well as in more severe cases, poisoning (Euripidou and Murray, 2004). Microbial contaminants – pathogens- are also important to consider. Such pathogens, usually include those that are transmitted via the faecal-oral route, causing gastrointestinal illness, as well as more severe, life- threatening diseases (Goodfellow and Taube, 2016). The literature has shown that some common pathogens found in floodwater can survive from days up to several weeks following a flood event- thus lengthening the time in which they could infect someone, and also increasing the length of time the 'flood' poses a risk to public health (Gowrisankar et al., 2017; Jiang et al., 2005; Mawdsley et al., 1996; Yard et al., 2014).

Naturally, the floodwater, along with the contaminants are only a risk if members of the public physically come into contact with it. Scenarios in which this is likely to happen may include the following: children playing in floodwater, vehicles driving through floodwater and splashing pedestrians, or the floodwater being widespread in areas that may be the sole access route- and so there is no other option other than to walk through (Schets et al.,2008; Ten Veldhuis et al., 2010) . Of course, if the flood is more severe, there is also the possibility the floodwater may enter homes and gardens, making contact with residents unavoidable. in addition to this, after the water recedes, contaminants may be 'left behind' in the area, with metals and chemicals being stored in soils, and pathogens surviving on surfaces and in soils for days or even weeks (Scoullos, 2020; Suleyman et al., 2018). Without the floodwater, there is no visible indicator to the public of possible contamination in the area, and so members of the public may unknowingly be frequenting areas in which pathogens still remain.

The risks urban floods pose to public health, is yet to be fully understood. There are a number of factors to consider, all of which merit investigation. Some factors are far reaching- i.e. the root cause, or contributor, of flooding- climate change, and extreme weather events. Other factors are more specific, such as the processes involved in wastewater treatment, hydraulic performance of infrastructure and movement of water through these systems. To fully consider this problem, social and psychological studies could even be carried out on human behaviour around floods. This topic is multi-faceted with lots of uncertainties and many paths to investigate.

Understanding the presence and behaviour of pathogens in floodwater itself is also essential to better understand public health risk. Although the potential for other contaminants to cause illness is of importance, it is contact with a wide variety of pathogens that can potentially cause severe, infectious, and also potentially epidemic illness among the population. Not only do we need to investigate the pathogenic and disease-causing species present, - but also, how many are present, how they move during a flood, and how they travel between the flood water and urban surfaces. Bacteria can live in soil, so at sites that flood regularly it is important to study the bacterial community even when the obvious risk (floodwater) has departed. I.e we need more evidence as to whether there is a persistent health risk at flood sites, if risk is higher at certain times, or if the health risks are short term only and disappear after a few hours/days. Past studies tend to focus on Faecal Indicator Bacteria (FIB) such as *E. coli*- due to ease of detection and levels being regulated in legislations for microbial surveillance of rivers and waterways, such as the Water Framework Directive (DEFRA, 2014). FIB are traditionally used as an indicator for the presence of faeces and thus disease (Odonkor and Ampofo, 2013; Price and Wildeboer, 2017; Shah et al., 2016; Ten Veldhuis et

al., 2010; Yard et al., 2014).However, it has been shown that some FIB are naturally present in the environment as 'naturalised' FIB rather than enteric FIB, meaning using on FIB counts alone to detect pollution from faeces may not be reliable (Devane et al., 2020; Evans et al., 2019). This study will focus, not only on FIB , but will characterise the broad range of bacterial species, including pathogens, that are commonly found in urban floodwater.By conducting repeated sampling at flood prone sites ,this study aims to develop a new understanding of the bacterial composition of soil and water in urban flood areas and how this varies over seasonal timescales.

To build a better understanding of these processes this study will focus on the specific aspects of : identifying the bacterial composition of urban flood water, and evaluating whether frequent flooding has an effect on the soil bacterial community, understanding whether the bacteria from the floodwater move between the soil and water, and assessing how the composition of the bacterial communities at a flood site change over seasonal timescales. A better understanding of these unknowns, would allow the health of the public to be safeguarded following an urban flood event- using data to determine the safety of a recently flooded area.

**4.3 Materials and methods**

4.3.1 Experimental rationale and overview

In order to investigate the range of bacterial pathogens often found at urban floodwater sites, an extensive/detailed field study was carried out. This allows the microbiome of urban flooding to be determined *in-situ*, whilst accounting for seasonal variation. Many studies have previously conducted urban flood studies in the field, but most are carried out over relatively short time periods- days, or weeks, or some investigate 'one-off', flood events (Castro-Ibáñez et al., 2015; de Man et al., 2014; Shah et al., 2021) . This study will instead, focus on sites that experience frequent small scale surface water flooding, over a longer time-scale. . As well as this, previous studies also tend to focus on the analysis of a small number of pathogenic species-usually those used as faecal indicators, such as *E.coli*, and coliform species (Edberg et al., 2000). Thus, the longer term changes in viability, mobility, and behaviour of the true range of the microbiological community present, is often overlooked.

Naturally, due to the nature of flood studies allowing only for 'opportunistic' sampling, collecting long term data is certainly more challenging. However, such data will be extremely valuable in helping to expand the data ranges available on not only microbial diversity, but also how this can change seasonally and temporally, and how this varies over the soil-water interfaces influenced by urban flooding.

4.3.2 Sample site selection

To determine appropriate sites the Environment Agency's surface flood risk map was used in conjunction with The Rivers Trust's 'Is my river fit to play in?' map tool, and advice from Sheffield City Council to identify sites at annual risk of surface flooding, that are also nearby to CSO outlets that spill regularly (Environment Agency, 2019; The Rivers Trust, 2021). Word of mouth in the local area, as well as advice from volunteer groups such as 'Rivelin Valley Conservation Group', and

'Shiregreen Litter Pickers' were also sources of information for the selection of potential sample sites. After preliminary visits and tests during the first year of the study, it was decided a maximum of two sample sites was feasible, in order to ensure that sampling and analysis were conducted in an accurate and reproducible way, without the deterioration of samples. As a consequence of this, and after initial visits, two sampling sites were chosen in Sheffield, that had different sources of 'contamination', for routine sampling.

#### *4.3.2.1 Site 1: Tongue Gutter*

This site is located by a river – Hartley Brook Dike - (53°25'43.3"N 1°28'08.1"W) , (also known as Sheffield Lane Dike and Tongue Gutter), which runs parallel to residential housing areas, and feeds into Blackburn Brook, a stream that eventually meets the River Don. Alongside the river and sample point, a public footpath is also frequented by dog walkers and cyclists. At the sample site there is an outfall pipe discharging into a creek/brook. On every visit to the site the outfall pipe was consistently discharging and a strong smell was present, independently to the presence of rainfall The site had a lot of aesthetic pollution, some of which appeared to flow out of the outfall pipe, while another source of aesthetic pollution was via fly tipping from members of the public (Figure 6).

**Figure 6.** Sampling site at Tongue Gutter on a visit in May 2022.

At Tongue Gutter, separate sewer systems are used in the homes nearby. Misconnections to systems like this have previously been found as the source of contamination (Chandler and Lerner, 2015), and could be the case at this site. This would mean foul water was connected to a surface water sewer system, and thus wastewater may be discharged via the surface water outlet. On several site visits the discharging water appeared frothy and bubbly, which could have been an indicator that the water discharging was mixed with a soap or detergent- supporting the theory of misconnections. This site has a surface flood risk categorised as 'high', and so has an annual risk of surface floods of <3.3% (Environment Agency, 2019) (Figure 7).

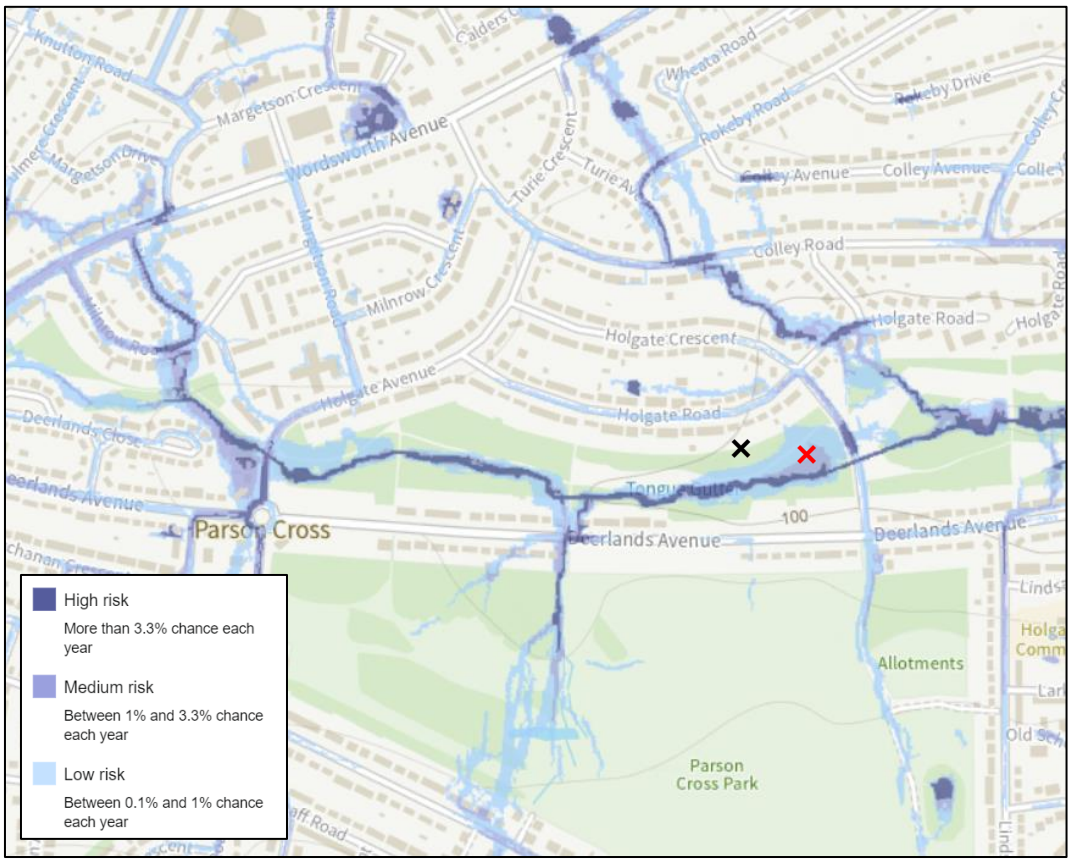


**Figure 7**. Surface water flood risk map of 'Tongue Gutter' site (Environment Agency, 2019). Red cross indicates area samples were taken, black cross indicates the area background soil samples were taken.

*4.3.2.2 Site 2: Endcliffe Park.*

The second sample site is a large city park, Endcliffe Park, which lies along the course of the Porter Brook River – (53°22'06.4"N 1°30'21.0"W). The area consists of grassy recreational parkland, along with woodland- with the chosen sample site being located in the former. The sample location is an area that frequently has a large pool of standing water- especially visible after rain. The site is also close to two gullies-which are thought to contribute to the water pool, with water being seen flowing out of one several times after rainfall.
This site is particularly interesting due to the proximity to recreational ground as well as a children's playground (Figure 8).

**Figure 8.** Sampling site at Endcliffe Park, with visibly surcharging gully during a visit after heavy rain in February 2022.

During particularly heavy rain, the pool of potentially contaminated water has been known to reach the playground itself. Similarly to site 1, this site has a surface flood risk categorised as 'high', and so has an annual risk of surface floods of <3.3% (Environment Agency, 2019) (Figure 9).

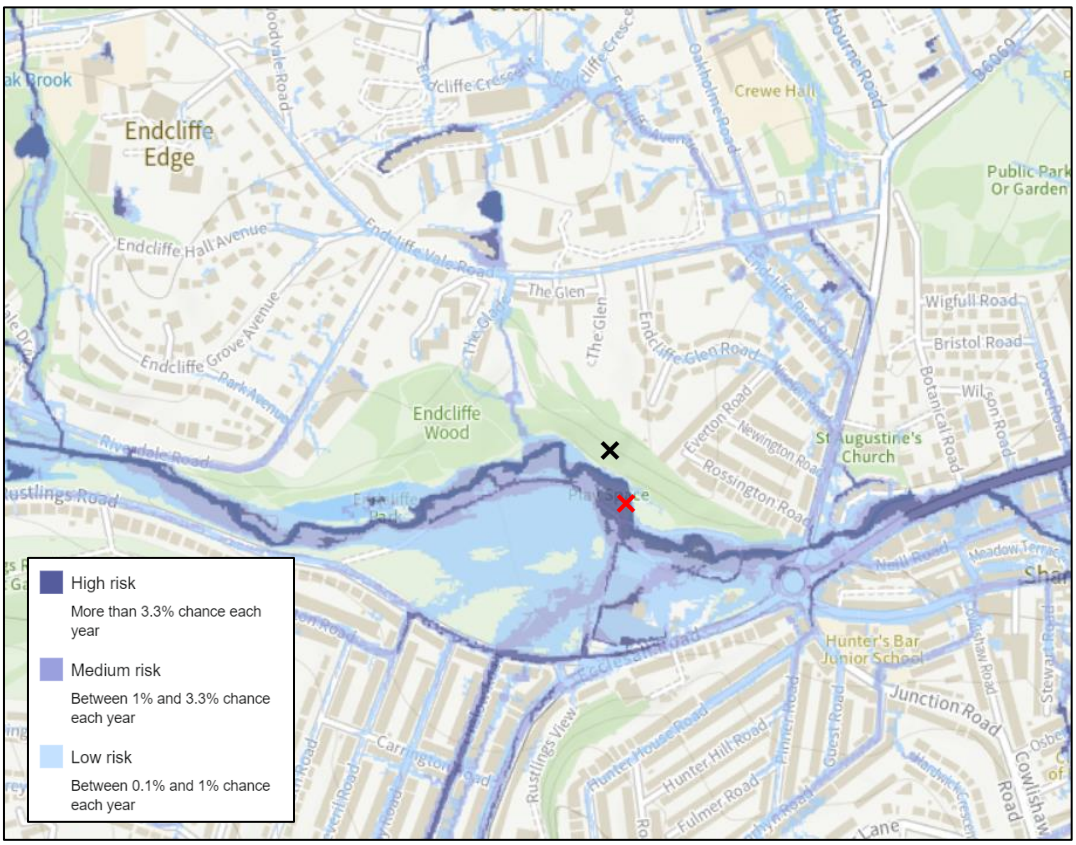


**Figure 9**. Surface water flood risk map of 'Endcliffe Park' site (Environment Agency, 2019). Red cross indicates area samples were taken, black cross indicates the area background soil samples were taken.

4.3.3 Sampling regime

Both sites were visited 12 times, over a 10-month period, under different weather conditions. Visits to each site were carried out roughly every 3-6 weeks, with 'opportunistic' sampling also taking place following any intense rainfall events. On each visit, triplicate soil samples were taken, from 0-10cm depth, at both sites using a basic handheld soil auger, as well as control (background) soil samples located close by to the sample site but in an area not known to flood (Figures 7 and 9). Samples were placed in Ziploc plastic bags and sealed to prevent drying-out. The samples were stored in the dark and transported within 4 hours of collection to the laboratory and stored at 4 ºC (again in the dark) until analysis was undertaken. This minimised any potential changes in the microbial communities in the soil samples. For the molecular genomic work, 50 g sub-samples of soil were sealed in small plastic bags and were preserved at –20 ºC for DNA analysis.

Flood water samples were also obtained when possible, at each visit. For Site 1: Tongue Gutter, samples of flood water were taken at every visit. This was due to the fact the outflow pipe was always spilling, regardless of rainfall amount. Whereas, at Site 2: Endcliffe Park, water samples were only able to be taken on 5 out of 14 visits. This was due to the origin of the flood at Endcliffe Park being an overflowing gully combined with surface water-, creating a 'flood'- when heavy rainfall occurred in the preceding days. Only then, could a water sample be taken.

4.3.4 Rainfall Monitoring

To understand the influence of precipitation on microbial dynamics, rainfall was monitored using a remote GPRS rain gauge located less than 5 miles from both sites (Detectronic, CIVICA, United Kingdom) (Figure 10). The gauge automatically logs rainfall data every hour. This gauge uses a 'tipping bucket' design accompanied by a GPRS data logger (Detectronic, CIVICA, United Kingdom). This rain gauge complies with the 'Guide to Meteorological Instruments No. 8 of the WMO', has an unlimited capacity and has an accuracy of +/- 1% at 26mm/hour (Detectronic, 2020; WMO, 2020).

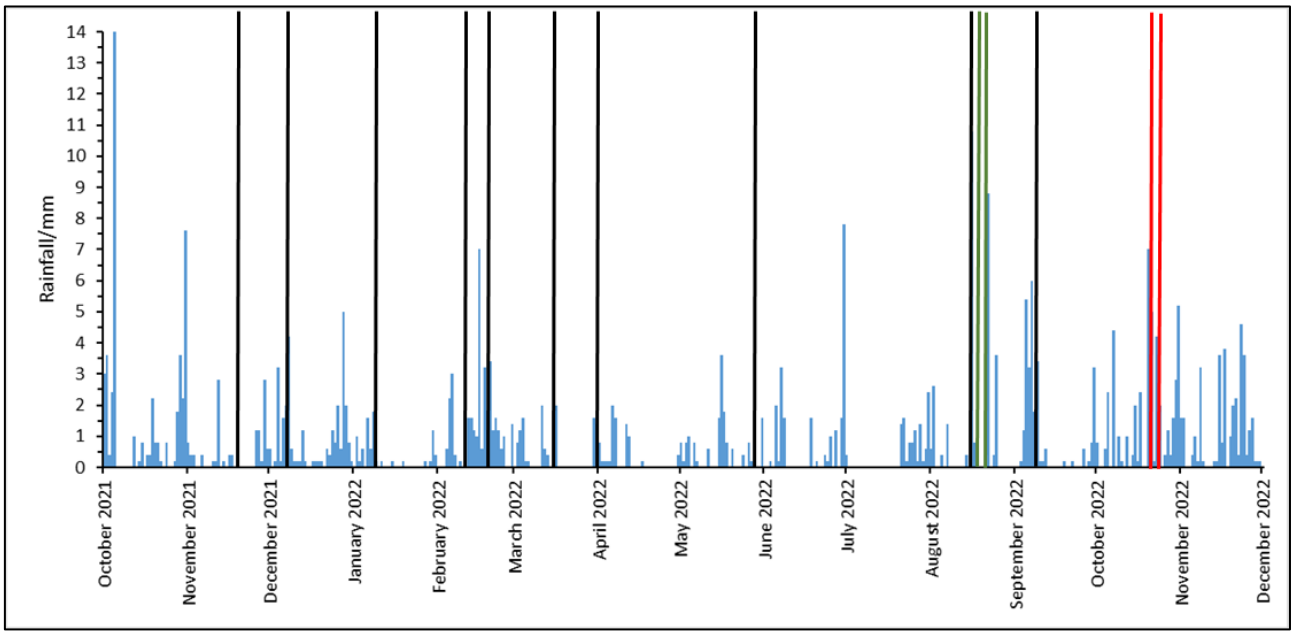


**Figure 10**. Daily rainfall over the sampling period, as measured by rain gauge. Black vertical lines indicate the day both sites were visited. Red line indicates only Tongue Gutter visited. Green line indicates only Endcliffe Park visited.

**4.4 Sample analysis**

4.4.1 pH and Temperature

The first parameters measured for every soil and water sample were pH and temperature. This was done in the field using a portable probe (Sension+ PH1 Portable pH Meter, Field Kit with Electrode for General Purposes, Hach, UK, taking measurements in triplicate. For pH and temperature measurements of water in the field, a 500ml beaker of water was taken from the source, the mixture was vigorously mixed and the probe inserted for *in situ* measurements in triplicate.

The pH of the soil is best measured *in situ* in order to avoid any potential changes in sediment which may occur during transport to the laboratory (Gumbley *et al*., 2005). To measure the pH of soil samples, a slurry was created at the field site by mixing the soil sample with distilled water using a ratio of 2.5:1, soil sample: distilled water (Rowell, 2014). Temperature was not taken for soil samples.

4.4.2 Determining soil moisture content and organic content

To measure the moisture content of the soil samples, 30g of sample from each site was dried in an oven at 105˚C, for 4 hours. The samples are weighed before and after drying and the difference in weight is equal to the % moisture of the sample (BS1377 PART 2:3.2) (British Standards Institute, 1990). The same process is then repeated in a 550˚C furnace, to determine the loss on ignition %/organic matter % of the soil samples. The following calculations were used to measure the moisture and organic matter content of the samples:

Moisture content:
% of fresh weight = ([mass fresh soil- mass dry soil]/mass fresh soil) x 100

Organic matter content:

% of dry weight = ([mass dry soil- mass of ash after ignition]/ mass dry soil) x 100

4.4.3 Soil particle size analysis and classification from field samples

After drying, subsamples from each site were pooled and passed through a set of 10 standard geotechnical sieves, from 5mm to 63 micron (Appendix 1). The initial weight of the dry sample was measured, and the stack of sieves shaken on a sieve shaker for 20 minutes (British Standards Institute, 1990). After this, the amount of sample retained on each sieve was weighed and recorded. The results of dry sieving were then plotted on a particle size distribution chart according to British Standard 1377 Part 2, and compared to a soil sizing chart to determine soil type (Appendices 2 and 3)(British Standards Institute, 1990).

4.4.4 Bacterial abundance: Heterotrophic and coliform plate counts

Microbial abundance can be investigated by culturing on selective nutrient agar plates. This process – heterotrophic plate counts (HPC)- allows any cultivable, heterotrophic bacteria (bacteria that

require organic nutrients for growth) present in the samples to be determined and was performed for all soil and water samples. It should be noted that HPC only gives information on a small part of the microbial community. However, this method is widely accepted and used in water industries (Allen et al., 2004) , due to the low cost, and ease of use, and so is used for this study.

Before plating, serial dilutions were performed for soil and water samples. However, a slurry was made for soil samples using 1g of soil and 10ml of PBS- Phosphate buffered saline (Invitrogen™ PBS (Phosphate-Buffered Saline) Tablets, ThermoFisher Scientific, United Kingdom). After dilution, 0.1ml from each sample tube was pipetted onto plates containing LB Agar (Invitrogen™ LB Agar, powder (Lennox L agar) ThermoFisher Scientific, United Kingdom), as well as plates containing Chromocult Coliform Agar (Chromocult® Coliform Agar, Merck, Germany). Preliminary tests determined that a dilution factor of 10-3 would be used for samples plated on LB Agar, and a dilution factor of 10-1 would be used for samples plated on Chromocult Coliform Agar.

This was carried out in a laminar flow (SafeFAST Classic class II A1/A2 -212, Microbiological Safety Cabinet, FASTER SRL, Ferrara, Italy, in order to avoid contamination of samples from airborne bacteria, and also protect lab users from any potentially infectious bacteria. The plates were then inverted and left to incubate at approximately 37˚C for 24 hours.

When using LB broth as the nutrient media, bacteria grow as usual, however Chromocult media is selective, and so it signifies the presence of just two bacteria- *Escherichia coli* (*E. coli*) and *Citrobacter freundii* (*C. freundii*) - both faecal indicators. On Chromocult, *E. coli* colonies appear purple and *C. freundii* appear as pink colonies (Figure 11).

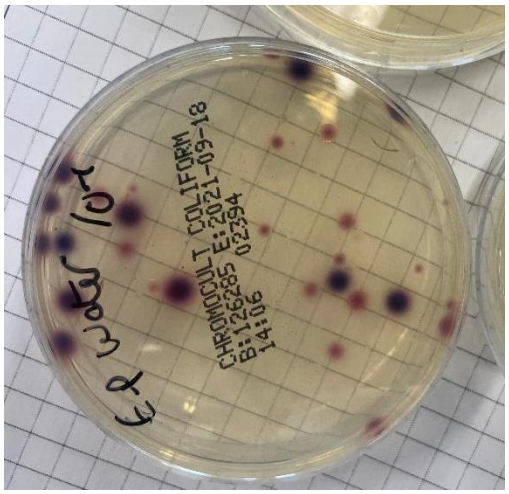


**Figure 11.** Colonies of *E. coli* (purple) and *C. freundii* (pink) on a chromocult agar plate.

After incubation, the colonies on each plate were counted. Once the colony forming units have been counted, the number of colony forming units per millilitre (CFU/ml) of the original sample was calculated by multiplying the number of CFU's counted, by the dilution factor. This method gives a general estimation of the concentration of heterotrophic bacteria in a sample.

4.4.5 Scanning Electron Microscopy (SEM) to characterise soil microbiome

Soil samples were taken on a day without rainfall at the 'Endcliffe Park' site and visualisations created using SEM. This was to give a detailed, and magnified view of the typical soil structure- including any microbes that are present under 'usual' (non-flood) conditions. This process was carried out at the University of Sheffield's Electron Microscopy Service. The following steps were carried out:
Specimens (each specimen=0.5g of soil) were fixed in 2.5% Glutaraldehyde (Merck, Germany) in 0.1M phosphate buffer (Merck, Germany) for 3 hours at 4˚C. Specimens were then washed in 0.1M Sodium Cacodylate buffer (Merck, Germany, three times, with 30 minute intervals at 4˚C. Next, secondary fixation in 2% aqueous osmium tetroxide (Merck, Germany) was carried out for 1 hour at room temperature. The specimens were then dehydrated through a graded series of ethanol (Fisher Scientific, UK):

- 75% ethanol for 15 minutes
- 95% ethanol for 15 minutes
- 100% ethanol for 15 minutes
- 100% ethanol again for 15 minutes
- 100% ethanol dried over anhydrous Copper sulphate (Merck, Germany) for 15 minutes
- 50/50 mixture of 100% ethanol and 100% Hexamethyldisilazane (Merck, Germany) for 30 minutes
- 100% Hexamethyldisilazane for 30 minutes.

Then, specimens were air-dried overnight. When dry, specimens were mounted on 12.5mm diameter stubs, attached with Sticky Tabs, and coated with approximately 25nm of gold in an Edwards S150B sputter coater (Edwards, UK). Specimens were finally, imaged using a Tescan Vega3 LMU SEM (Tescan, Czech Republic) at an accelerating voltage of 15Kv.

### 4.4.6 Molecular genomic methods

The use of these molecular techniques allows the characterisation of those components of the soil microbial community that currently cannot be cultured (Hosokawa et al., 2022).

### 4.4.7 DNA extraction- Soil samples

DNA was extracted from soil samples using the PowerSoil® DNA Isolation Kit (Mo Bio Labs Inc.). The kit utilises the bead beating method. The 'Experienced User' protocol included in this kit was followed (Mo Bio, 2016).

### 4.4.8 DNA Extraction- Water

#### *4.4.8.1 Filtration and cell concentration of water samples*

To extract DNA from water samples, the samples first require filtration. This concentrates any microbial cells in the water onto the filter. To do this a three way vacuum filtration system was used (Sartorius 3 Branch Microsart™ Manifolds with Funnels, Sartorius, Germany) with nitrocellulose filters with a pore size of 0.2µm and 47mm diameter (MF-Millipore® Membrane Filter, 0.22 µm pore size, 47 mm diameter, mixed cellulose esters (MCE) membrane, Merck, Germany). This pore size is required as most bacteria and pathogens are much larger, at 0.5-5µm (Levin and Angert, 2015) and so will not pass through the filter, but remain on the filter for DNA extraction.

This method prepares water samples for DNA extraction by trapping the microbial cells on the filter. The extracted DNA will be used for PCR, as well as for sequencing. After 500ml of each water sample has been filtered, the filters were stored in the freezer at -20˚C until DNA is extracted.
For the water samples, the Cetyltrimethylammonium Bromide (CTAB) method was used as described by Karunakaran *et al.,* 2016, and is outlined here. Filters were cut into strips using a sterile scalpel, and half of each filter was placed into 2ml Eppendorf tubes. Then, 740 µL of SET lysis buffer (SET lysis buffer consists of 40 mM EDTA (Ambion, Warrington, UK), 50mM Tris-HCl pH 9 and 0.75 M sucrose (Sigma Aldrich Co., UK.) ) and 90 ml of lysozyme 10 mg/mL (Sigma Aldrich Co.,UK) were added. Tubes were incubated at 37 °C for 30 minutes with rotation in a hybridisation oven (Thermo Scientific, UK). After incubation 90 mL of 10 % sodium dodecyl sulphate (SDS) (Sigma Aldrich Co., UK.) and 25 mL of proteinase K 20 mg/mL (Applied Biosystems, Life Technologies Ltd., UK) were added. Tubes were then incubated at 55 °C in the hybridisation oven for 2 hours with rotation.After the 55 ˚C incubation, any supernatant was withdrawn and transferred into 2 mL sterile Eppendorf tube. 137 µL 5M NaCl 1 % and 115 µL Hexadecyltmethyl ammonium bromide (CTAB) / NaCl solution, were added to the tubes containing the supernatant (Sigma Aldrich Co., UK). Tubes were then incubated at 65 °C for 30 minutes with rotation.

After incubation, 838 µL of chloroform:isoamyl alcohol (24:1) (Sigma-Aldrich, UK) was added in a fume cupboard and the tubes then centrifuged at 13,000 RPM for 5 minutes. After centrifugation, the supernatant was extracted into a 2 mL tube and steps 9 and 10 were repeated.
DNA was then precipitated, by adding 815 µL of 100 % isopropanol (Sigma-Aldrich, UK). The tubes were then left overnight at -20˚C.

The next day, samples were centrifuged for ten minutes at maximum RPM. After centrifugation, the supernatant was removed carefully, leaving the DNA pellet in the tube. The pellet was then washed twice with 1 mL of 70 % ethanol, centrifuged for 5 minutes and the ethanol decanted.
Tubes containing the pellet were then air dried in a fume cupboard for approximately 2 hours (until ethanol had dissipated) and then suspended in 50 µL DEP-treated sterile water (Thermo Scientific, UK).

#### *4.4.8.2 Quantification of extracted DNA*

The amount of DNA in each sample was quantified using a QuBit™ 4 Fluorometer (Thermo Fisher Scientific, UK), following manufacturer's instructions. The 'DNA Broad Sensitivity Kit' as well as the 'DNA High sensitivity Kit' were used in order to ensure all ranges of concentrations of DNA were measured from 0.2ng/µl -1000ng/µl (Thermo Fisher Scientific, 2021).

### 4.4.9 Sequencing the 16s rRNA gene

Next generation sequencing was performed at the NEOF molecular facility at the University of Liverpool, targeting the 16s rRNA gene, fourth hypervariable (V4) region. Hypervariable regions show different sequence diversity between various species of bacteria- and thus can be used for species identification. The V4 region, specifically has high functionality in the ribosome and is highly conserved (Morosyuk et al., 2017; Van de Peer, et al., 1996).

The Illumina Platform was used - for sequencing a of the 16s rRNA gene to estimate bacterial community structure. The samples in this study were sequenced using Illumina MiSeq® pair-end technology (Illumina, United States). Illumina MiSeq® has been chosen as it has a sequencing platform that provides a maximum of 250 million reads of lengths up to 2 × 300 base pairs (Caporaso *et al.,* 2012). To target the 16S rRNA gene V4 region (16S v4), the following forward and reverse primers were used. Portion of the sequence in **bold** indicates the recognition sequence, allowing for a secondary nested PCR process to be performed.

F: 5'**ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNN**GTGCCAGCMGCCGCGGTAA3'
R: 5'**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**GGACTACHVGGGTWTCTAAT3'

A successful amplification product results in a 254bp insert as described by Caporaso et al (2011 PNAS 108, 4516-4522).

#### *4.4.9.1 Test Polymerase chain reaction (PCR) of extracted DNA*

After DNA extraction and quantification, a test PCR was performed on extracted samples in the NEOF visitor lab in Sheffield, prior to send the DNA for sequencing at the NEOF lab in Liverpool. This test was performed to check that the quantity and quality of the extracted DNA would allow for the subsequent amplifications needed for sequencing. For this PCR the following controls were included. An extraction kit negative control- this was a mock nucleic acid extraction, using the reagents provided in the extraction kit, but without any actual sample. An extraction kit positive control- this was an extraction from a community of cultured *E. coli* to ensure extraction kit accuracy. A PCR negative control was also used. This was a no-template control, which omits any DNA or RNA template from the PCR reaction. This serves as a general control for nucleic acid contamination. A PCR positive control was also used, again using *E. coli.*

The test PCR used the following conditions: Initial denaturation was at 98 °C for 2 mins. Then 10 cycles of: denaturation 95 ˚C 20 secs, annealing 65 ˚C 15 secs and extension 70 ˚C 30 secs. With a final extension step of 72 ˚C for 5 min. The PCR mixture contained 5 µL of template DNA, 10 µL 2x Kapa HiFi HotStart Master Mix (Roche, Switzerland), 0.25 µL of 10 µM of each primer, and molecular biology grade water to a final volume of 20 µL.
The size of the PCR amplicons was then checked using gel electrophoresis.

#### *4.4.9.2 Gel electrophoresis to check DNA size after test PCR*

To check the size of the tests PCR amplicons the agarose gel electrophoresis method was used. A 0.8% agarose gel was made with agarose powder (Sigma Aldrich, United Kingdom) and Tris/Borate/EDTA (TBE) buffer, diluted X10 (Thermo Scientific, United Kingdom). 5ul Ethidium bromide (Thermo Fisher Scientific, 2018) was used to stain the gel and to allow for DNA visualisation. Orange G (Thermo Fisher Scientific, United Kingdom) at X6 dilution was used as a loading/tracking dye, and 1X TBE (Merck, Germany) buffer was used in the agarose gel box. In this case a 1KB DNA ladder was used (1kb Hyperladder™, Meridian Bioscience, United States) to check for amplicon size, with gels ran at a voltage of 100V for approximately one hour. A trans-illuminator was then used to visualise the gel and DNA bands using UV.

*4.4.9.3 Illumina sequencing at NEOF molecular facility in Liverpool*

The NEOF Liverpool facility uses a 16s dual index nested PCR for Illumina's sequencing (NEOF, 2024). The 1st PCR amplification uses the primers and conditions described previously in Section 4.6.1 (Caporaso et al 2011, PNAS 108, 4516-4522). This primer design incorporates a recognition sequence to allow a secondary nested PCR process. The PCR products were then purified with AMpure beads using a 1:1 ratio (Beckman Coulter, USA). The second PCR is then performed to incorporate Illumina adapter sequences to allow for the sequencing of samples on the Illumina MiSeq platform. Barcodes for sample identification are also incorporated at this point.

The general sequences of the forward and reverse primers for the second PCR are illustrated below. The 8bp barcode for sample identification after sequencing is underlined.
**N501 f:**
5'AATGATACGGCGACCACCGAGATCTACACTAGATCGCACACTCTTTCCCTACACGACGCTC 3'
**N701 r:**
5'CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTGACTGGAGTTCAGACGTGTGCTC 3'

The secondary PCR was performed under the same cycling conditions as described for the first PCR, but for 15 cycles. PCR mixture contained 10 µL of amplicon from 1st PCR, 10 µL 2x Kapa Master Mix, 0.5 µL of 10 µM of each primer, and molecular biology grade water to a final volume of 20µL. Secondary amplicons were purified with AMpure beads (Beckman Coulter, USA ), as before.

*4.4.9.4 Bioinformatics: DNA sequences analysis*

Firstly, the raw Fastq files were trimmed for the presence of Illumina adapter sequences using Cutadapt version 1.2.1 (Martin, 2011). The option -O 3 was used, so the 3' end of any reads which match the adapter sequence for 3 bp. or more were trimmed. The reads were then further trimmed using Sickle version 1.200 with a minimum window quality score of 20 (Joshi and Fass, 2011). Reads shorter than 15 bp. after trimming were removed. These trimming steps were performed at the NEOF Liverpool molecular facility (Figure 12).

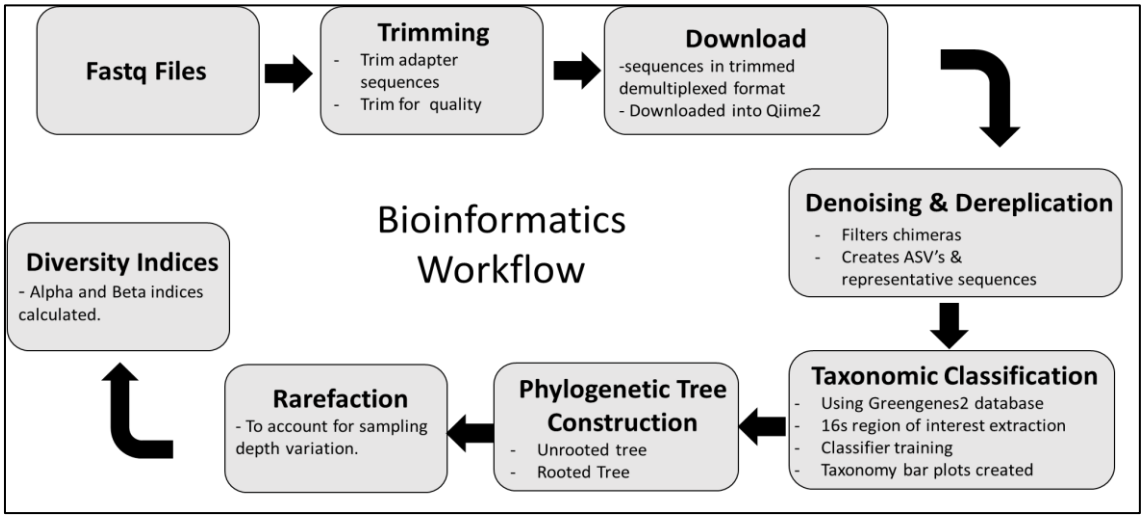


**Figure 12**. Pipeline of workflow used for bioinformatics analysis.

After downloading the demultiplexed trimmed sequences, the data was imported into the Quantitative Insights into Microbial Ecology 2 program in the Paired End Fastq Manifest Phred33 format (QIIME2, version 2019.7) (Bolyen *et al.,* 2019). This means the sequence files were imported with a tab delimited 'Manifest' file which maps sample identifiers to absolute filepaths that contain sequence and quality data for the sample. After sequences were imported they could be visualised in Qiime2 View (Bolyen *et al.,* 2019). This gave information on the quality of forward and reverse reads using a random sampling of 10000 reads, as well as showing a summary of sequence lengths. This information was important for the next step.

Next, De-novo amplicon sequence variants (ASVs) were identified using the 'DADA2' command. This method is based on sequencing error correction algorithms and generates ASVs, rather than using clustering methods, which typically generate individual operational taxonomic units (OTUs).
This step denoises paired-end sequences, dereplicates them, and filters chimeras. For this step a truncation length of 230bps was chosen for forward and reverse reads, as when viewing the aforementioned interactive quality plots, this is the base position at which quality decreased for both the forward and reverse reads. This step provided an ASV abundance table, along with representative sequences.

With the sequence variants defined, the next step was to carry out taxonomic classification using Greegenes2 version 2022.10 (Mcdonald et al., 2023). The full length 16S rRNA sequences that are used within the Greenegenes2 2022.10 phylogenetic tree were downloaded into QIIME2. After the database was downloaded the 16s region of interest was extracted. This step is required to reach optimum assignment performance. Next the classifier was trained on the set of extracted reads using the Naive Bayes algorithm. Then taxonomic assignments were assigned to the ASV representative sequences using the trained classifier at a confidence level of 0.7. This meant each taxonomic level must reach a confidence level of <= 0.7 to be classified. At this point bar plots of taxonomy could be visualised in Qiime2, and exported for each sample. This data also allowed for the creation of 'Heatmaps' based on relative abundance, this was created using exported data, and the online 'Heatmapper' tool (Babicki et al., 2016).

After sequences were assigned to taxonomic levels, a phylogenetic tree was constructed. This allowed the computation of phylogenetically aware alpha- and beta-diversity metrics. This involved creating first an un-rooted tree, and then a rooted tree in Qiime2. The detected diversity partly depends on the sampling depths. An incompletely sampled community will appear less diverse than a fully sampled community. Therefore, to compare the diversity between samples different sampling depths needed to be accounted for. This was carried out with rarefaction. Rarefaction is performed by reducing all the samples to the same sequence count, by random sub-sampling of the sequences. The minimum and maximum sampling depths were set at 2000 and 130000, again, according to the quality plots created during the 'DADA2' step, and also selected to ensure sequencing depth within samples were retained and no entire sample was lost.

Finally diversity indices were calculated for alpha and beta diversity. This was done by using the diversity 'wrapper' script' in QIIME2. The script co-investigates alpha- and beta-diversity at the same time. This gave multiple metrics for both alpha and beta diversity.

## 4.5 Diversity Indices

To fully investigate the bacterial diversity within the samples both alpha and beta diversity metrics were calculated in Qiime2.

### 4.5.1 Alpha Diversity

Alpha Diversity is used to measure diversity on a 'local' scale- 'within' sample diversity (Whittaker, 1960) .In this study, 3 different measures were used to determine alpha diversity: Faith's phylogenetic diversity (PD) Index, Abundance based coverage estimator (ACE) Richness index, and Simpson's evenness index (Chao and Lee, 1992; Faith, 1992; Simpson, 1949). Faith's PD index uses phylogenetic distance to calculate the diversity of a given sample (Faith, 1992), while the ACE Richness is a nonparametric richness index that measures the number of unique species, using sample coverage, defined based on the sum of the probabilities of the observed species (Chao and Lee, 1992). Simpson's evenness index is a measure of diversity which takes into account the number of species present, as well as the relative abundance of each species (Simpson, 1949). Each diversity index is based on different assumptions about the species diversity and thus, by using a combination of several indices, a more robust view of the biodiversity can be gained.

### 4.5.2 Beta Diversity

Beta Diversity is described as the ratio between regional and local species diversity (Whittaker, 1960), and is useful to compare differences in diversity between different groups, i.e. different sample sites or types. To calculate beta diversity in this study, weighted Unifrac distance analysis was used. There are two types of Unifrac distances: an unweighted Unifrac distance that considers only species presence and absence and counts the fraction of branch length unique to each community. The weighted method, used here, uses species abundance and also weights the branch length with abundance difference (Chen *et al.,* 2012).This phylogenetically aware metric takes into account presence and absence of species, as well as abundance- and is thus quantitative.

## 4.6 Statistical Analysis

Before performing statistical tests, the Shapiro-Wilks test was performed for all data to determine the normality of the data (Shapiro and Wilk, 1965). For differences between: physico-chemical factors in monitored soil, and water and background soil samples, along with differences in alpha diversity measures, the independent t-test was used. This test determines whether there is a statistically significant difference between the means in two groups (Mcdonald, 2014). To test for correlations, the Pearson's correlation coefficient test was used. This included testing for correlations between physicochemical factors, between different sample types, in relative abundance at different taxonomic levels between sample types, correlations between different alpha diversity measures, and between samples taken on different days. This method calculates the ratio between the covariance of two variables, and takes into account the standard deviations (Kirch, 2008).

For Beta Diversity analysis, Principal Coordinate Analysis (PCoA) was used. This method converts data on Euclidean distances between items into a 3 dimensional scatter graph plot. The plot has 3

axis, one for each ‘dimension’. The value displayed on each axis is called the ‘eigenvalue’. Axes are ranked by their eigenvalues, and so the first axis has the highest eigenvalue and thus explains the most variance. Each eigenvalue represents the amount of variance in each dimension as a percentage of the total variation in the distance matrix, and so the sum of the eigenvalues will equal the sum of the variance of all variables in the data set (Bakker, 2024).

The plot can be viewed, and shows information on the ‘closeness’ of items, allowing groups and clusters to be found. In Qiime2, this utilised a distance matrix of beta diversity values for each sample. This allowed data to be separated based on metadata groups, i.e. sample type, location, date, rainfall amount etc. This was mainly used in this study to visualise beta diversity data, and to determine if separation of samples had statistical significance or if it might have originated by chance, Permutational Multivariate Analysis of Variance (PERMANOVA) analysis was carried out. PERMANOVA is a non-parametric multivariate statistical permutation test. It uses a distance matrix constructed from a dissimilarity measure, and tests the null hypothesis that there are no differences in the relative magnitude (presence/absence) of a set of variables among objects from different treatments or groups (Anderson, 2017). In this circumstance the dissimilarity measure is that computes using the beta diversity distance matrix to determine differences in beta diversity between the metadata groups. This test was chosen here as its assumptions are negligible, and it works well with complex data/ many groups of data.

**4.7 Results**

This section describes in detail the results of analysis of physico-chemical factors, soil analysis, taxonomic analysis, and diversity indices measured. Any differences or correlations between factors are outlined. The section is split by sample site.

Soil characterisation: Soil particle size and class/type

After soil sizing using the sieve method, and classification according to British Standard 1377 Part 2 it was found that the soil found at ‘Tongue Gutter’ was predominantly coarse sand, with fine gravel. The soil at ‘Endcliffe Park’ was similar, also being classified as a coarse sand with fine gravel. However the soil at Tongue Gutter did generally contain more material of a larger particle size, than that of the soil at Endcliffe Park (Figure 13).

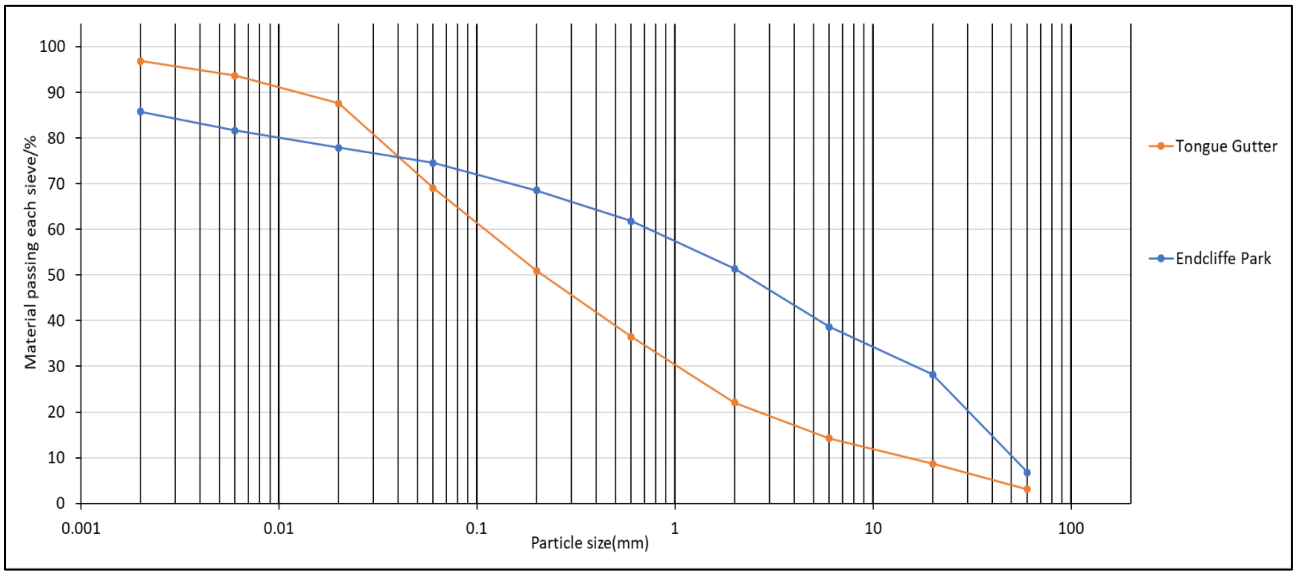


**Figure 13**. Soil gradation curve from soil samples collected at the 2 sampling sites.

Scanning Electron Microscopy (SEM) image analysis.

The images created using SEM give a visualisation of how the general soil structure looks, supporting the soil gradation carried out (Figure 13). It also allows visualisation of part of the soil microbial community under dry weather. In figure 14, SEM micrographs showed an irregular rough surface with high concentration of granular, inorganic material covering the analysed surface. Filamentous microorganisms and biofilm-like structures tended to accumulate on small/specific areas of the analysed surface, without covering the entire analysed surface.

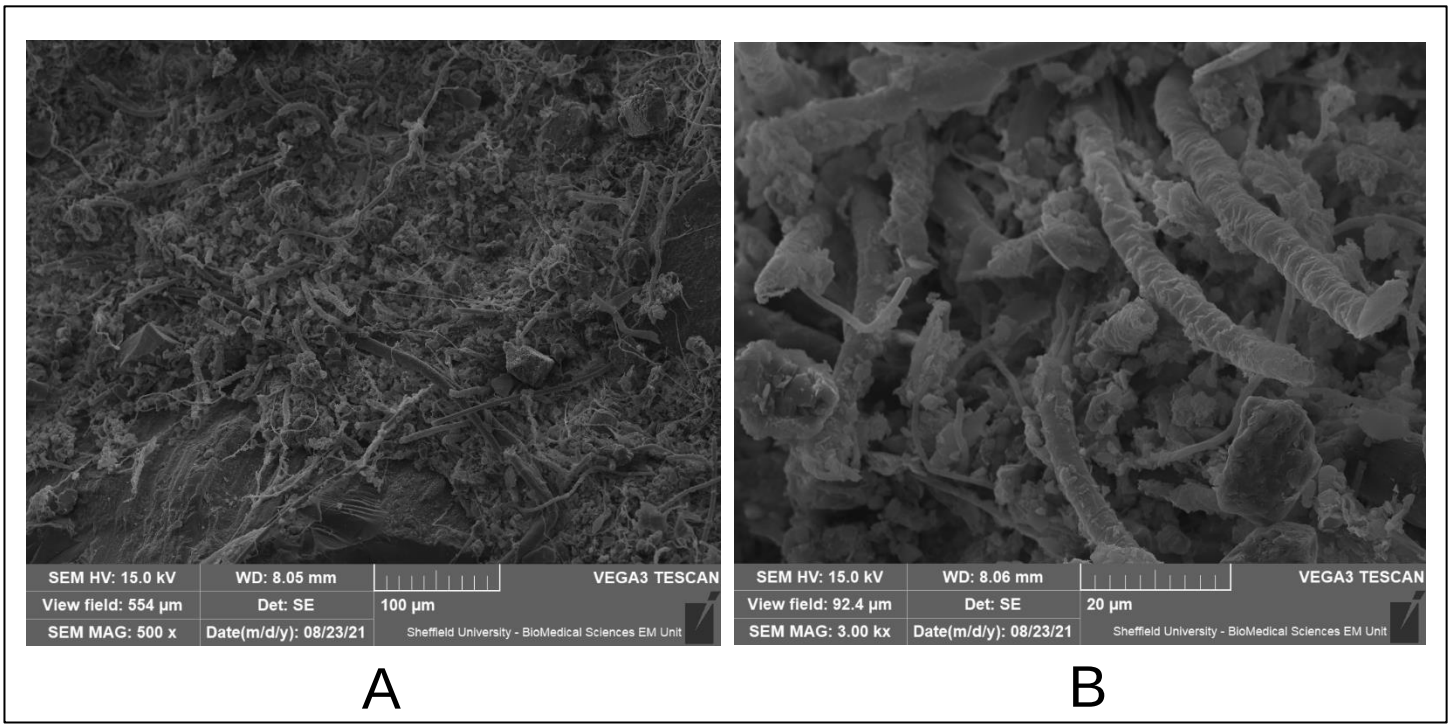


**Figure 14**.SEM images of soil samples taken from Endcliffe Park during a visit with no rainfall

4.7.1 Tongue Gutter- Analysis of monitored and background soil samples

Physicochemical analysis and estimation of bacterial abundance of soil

Table 1 includes physicochemical data obtained from soil samples, along with coliform colony counts. Over the sampling period, from November 2021 to October 2022, the pH of the soil remains fairly stable, with the lowest pH recorded being 7.09, and the highest 8.76. A higher pH range, 8.15-8.76 is evident over Winter 2021 and into Spring 2022, with a lower range, 7.09-7.88, recorded over Summer and Autumn 2022. Background samples showed similar values for the Winter 2021 and Spring 2022 period (7.79-8.78). In contrast, over Summer and Autumn 2022 background pH increased ranging from 7.38-8.50, significantly different to the Monitored Soil pH range for this period (t= 2.49, P=0.017).

As expected, with increased rainfall, increased moisture in the soil is also observed. Soil moisture content shows some variation over the sampling period but is not significantly different to background samples, despite being higher on average (t=-0.039, p=0.485) .The same is true for the organic content of the soil, with background data showing a higher average organic content than sampled data. Despite this- the organic matter content of the Monitored Soil over the monitoring period is not significantly different from background samples (t=-1.71, p=0.052).

Total coliform counts are much higher over the whole sampling period in all soil samples, when compared to background levels, with coliform counts ranging from 10000- 1040000 CFUs per gram of soil. The highest recorded coliform count for background samples is only 70000 CFUs per gram of soil. There was no significant correlation between coliform counts and rainfall amount.

**Table 1.** Results of site 1- Tongue Gutter physico-chemical analysis, coliform counts, and rainfall depth (n=12).

| Date of visit | Average pH of soil | Cumulative Rainfall 5 days (120 hours) before sample /mm | Total number of coliforms / CFU per g | Moisture content/% | Organic content by Loss on Ignition/% |
|---|---|---|---|---|---|
| 22/11/21 | 8.39 | 0 | 10000 | 28.12 | 6.74 |
| 13/12/21 | 8.15 | 10.4 | 650000 | 28.82 | 4.06 |
| 22/1/22 | 8.39 | 0.2 | 600000 | 28.31 | 4.09 |
| 16/2/22 | 8.19 | 21.6 | 0 | 29.04 | 3.66 |
| 23/2/22 | 8.76 | 54.6 | 170000 | 20.89 | 17.89 |
| 16/3/22 | 8.47 | 9.0 | 0 | 25.35 | 11.24 |
| 31/3/22 | 8.15 | 4.2 | 190000 | 31.08 | 12.63 |
| 30/5/22 | 7.53 | 1.8 | 350000 | 18.75 | 15.75 |
| 17/8/22 | 7.1 | 21.2 | 1040000 | 32.16 | 6.29 |

| | | | | | |
|---|---|---|---|---|---|
| 7/9/22 | 7.88 | 21.0 | 780000 | 28.12 | 6.59 |
| 24/10/22 | 7.09 | 63.0 | 50000 | 26.97 | 23.55 |
| 25/10/22 | 7.3 | 61.2 | N/A | 40.92 | 8.16 |

coliforms*= *E.coli* and *C .freundii* colony forming units

Bacterial community composition

Figure 7, shows the percentage of relative abundance at phylum level based on assigned ASV's. Taxonomic analysis showed a total of 23 assigned phyla for the monitored soil data, and 20 assigned phyla for background soil data. 20 of the phyla are the same for both datasets, this includes phyla such as, Acidobacteria, Planctomycetes, and Proteobacteria. However, the following phyla; Chlamydiae, Fusobacter, and Synergistetes, were not detected in the background data, but were present in the monitored soil data, but at relatively low abundance levels 0.01%-0.5% (Figure 15).

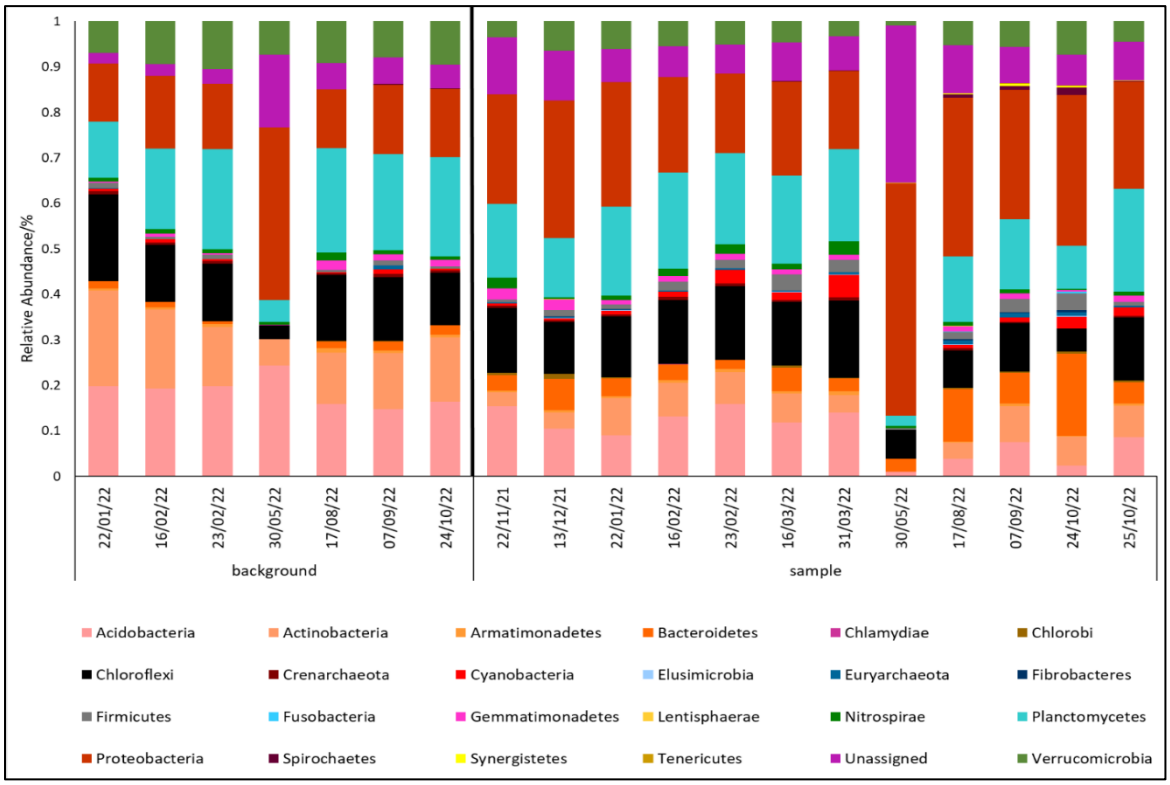


**Figure 15.** Comparison of bacterial communities present in monitored soil samples and background soil samples at Tongue Gutter, at phylum level (n=19). Unassigned taxa grouped into 'Unassigned' phyla category.

For both monitored soil and background soil samples, the 5 most abundant phyla across the sampling period were Proteobacteria, Chloroflexi, Planctomycetes, Verrucomicrobia, and Actinobacteria.

At genus level, after removing genera that have a relative abundance of <0.3%, 49 genera are present in Monitored Soil, with 35 genera present in background soil samples (Figure 16).
The most abundant genus for monitored soil over the sampling period is *Gemmata*-most abundant on 16th February 2022 at 2.15% relative abundance. Conversely, for background soil samples, *Rhodoplanes* was the most abundant genus, reaching 2.51% on 16th February 2022 When looking at the abundance of the genera that are shared by both datasets, very similar abundance levels of *Pirellula* were detected ~0.9% (average across the whole sampling period) for both monitored and background soil.

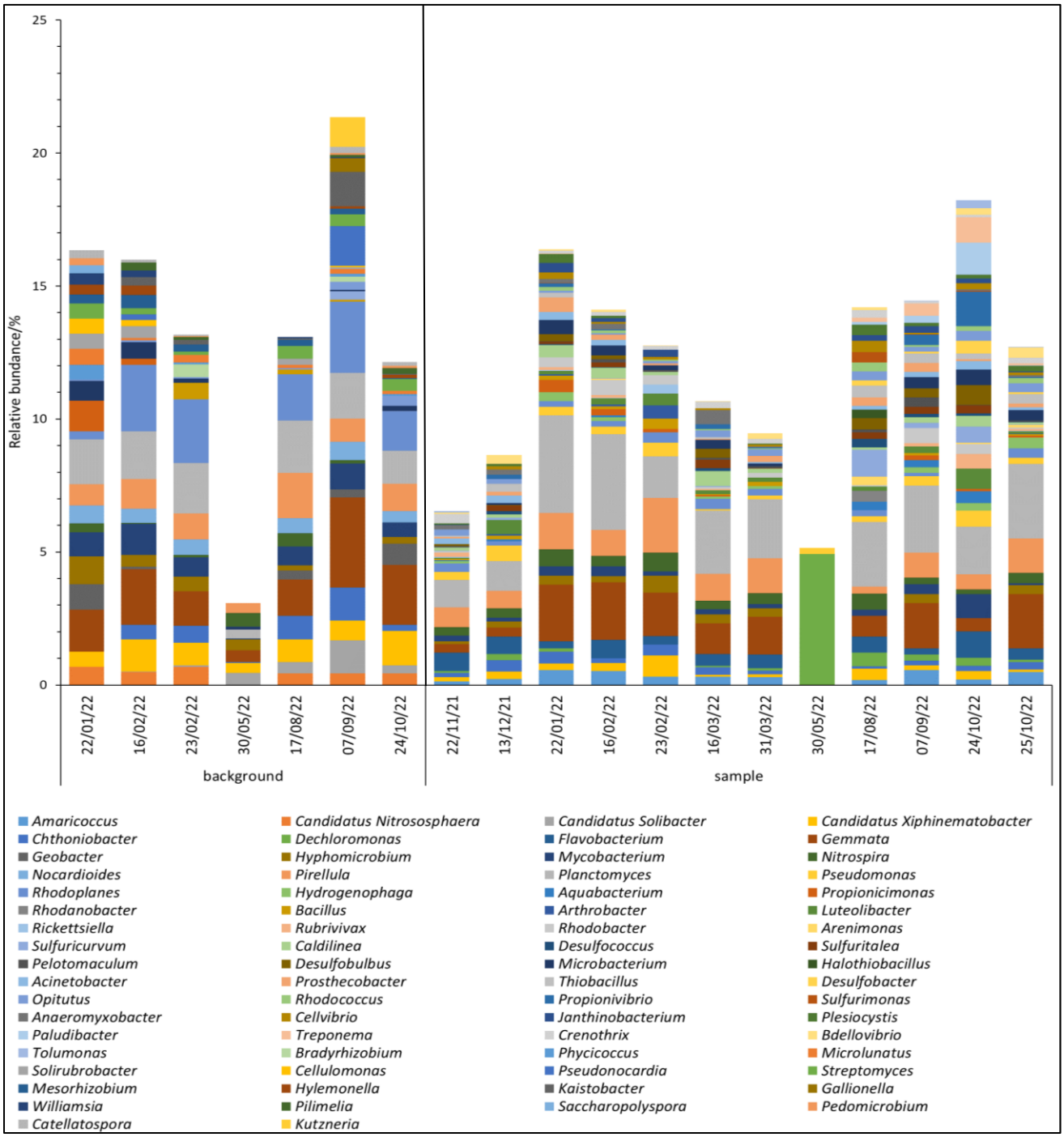


**Figure 16**. Relative abundance at Genera level of monitored soil and background soil (n=19). Sequences filtered to only include genera that have a relative abundance of >0.3%.

There were 38 genera detected in monitored soil, but not in background soil. 14 of those genera display highest relative abundance on the same date – 24th October 2022 . On this date, the highest 5 day cumulative rainfall depth was also recorded; 63mm, along with the highest organic content recorded in monitored soil - 23.55%.

Looking at changes in bacterial community structure over the whole sampling period, changes are seen seasonally in relation to abundance of *Pirellula* (Figure 17*)*- present in both background and monitored soil samples. In background soil, *Pirellula* abundance did not vary much over the sampling period (~0.8-1.7% abundance) However, Monitored soil, shows *Pirellula* abundance highest (average of 1.13%, compared to 0.73% for the rest of the sample period) when the rainfall amount was, lower, over January -March 2022. On the other hand, several genera found in monitored soil (not in background soil)- including *Thiobacillus* and *Treponema* had a higher abundance over the August-October 2022 period- when rainfall amount was higher (higher by 100mm in total over the period), compared to the previous time period (Figure 17). The same trend is also seen for *Bdellovibrio, Paludibacter*, and *Desulfobacter* (also not present in background soil).

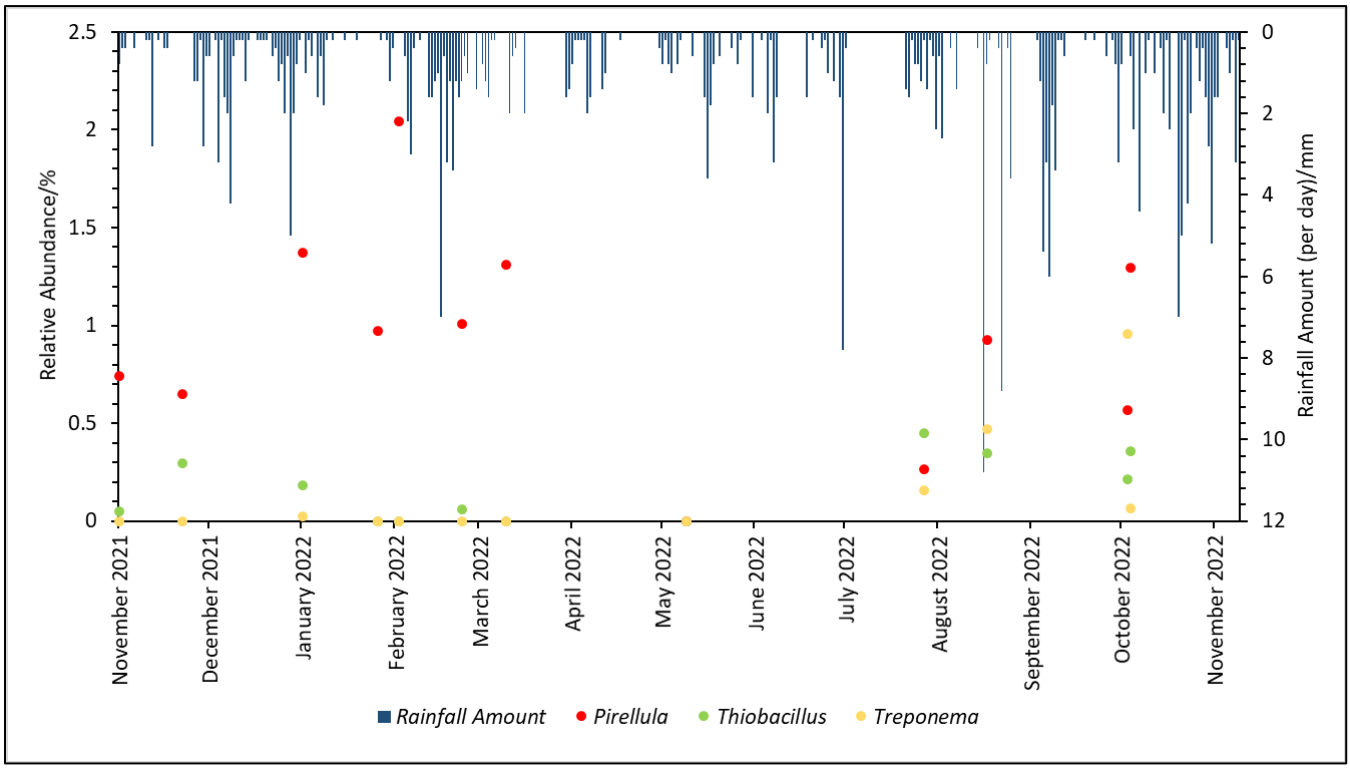


**Figure 17.** Abundance of *genera* in background and monitored soil at Tongue Gutter in relation to rainfall depth. Daily rainfall depth plotted on secondary axis

Further taxonomic analysis enabled categorisation of samples at a species level. Species detected at an average abundance of <0.03% for the whole sampling period were filtered out and the remaining species plotted on Figure 18. The monitored soil sample from 30th May 2022, was not able to be classified down to species level and so is not present in the species level heatmap for monitored soil.

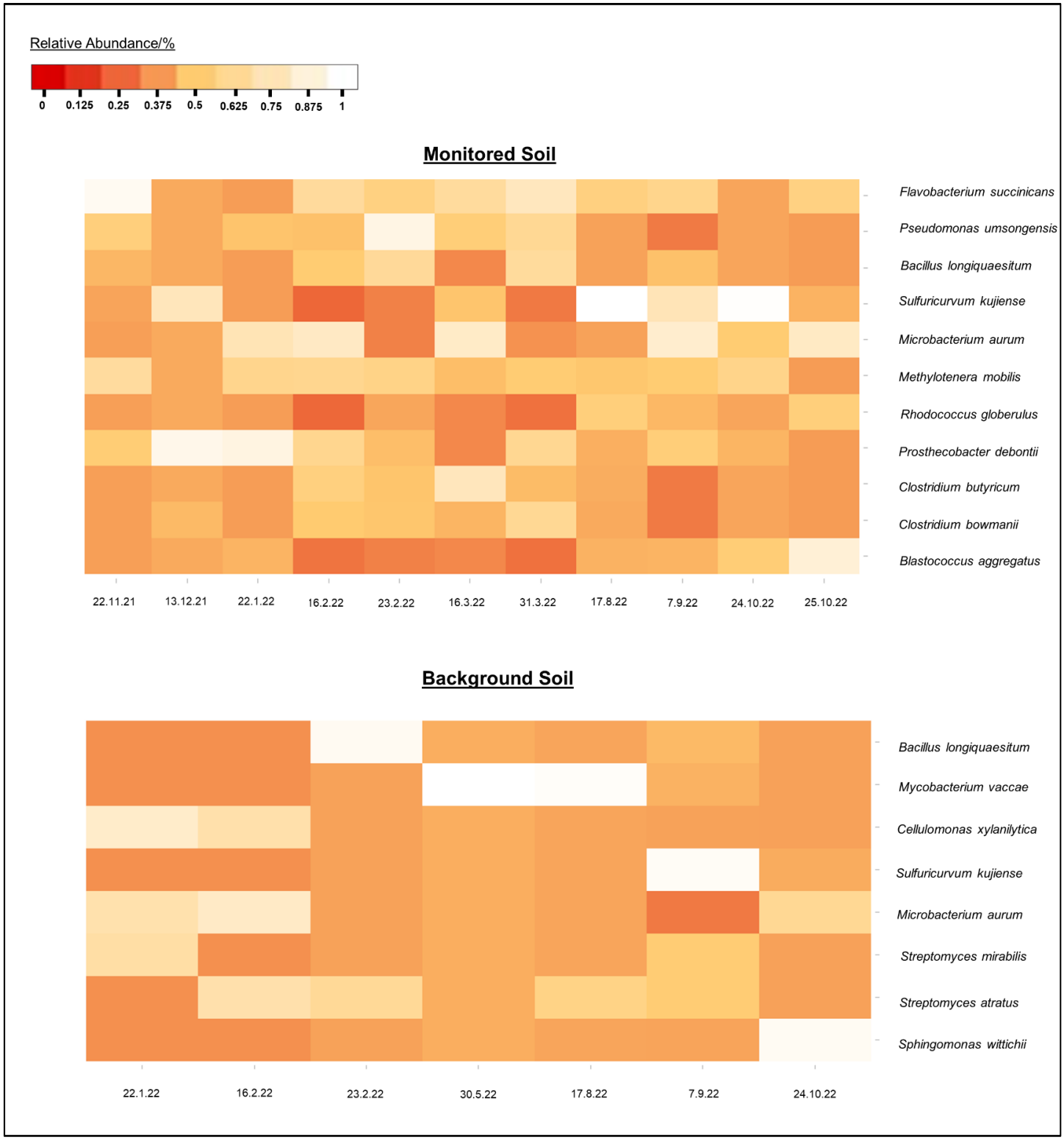


**Figure 18**. Heatmap based on the relative abundance of monitored soil samples and background soil at species level (n=18). Includes species that had an average relative abundance of >0.03% over the sampling period.

At species level there is a notable difference in community composition between monitored and background soil. The two datasets share only one species (relative abundance >0.03%). *Microbacterium aurum*. Across the sampling period, abundance of this species is generally higher in

background soil (average 0.12%), in comparison to monitored soil (0.08%). The most abundant species in monitored soil over the sampling period is *Sulfuricurvum kujiense,* highest on 17th August 2022 at 1% relative abundance. For background soil it is *Cellulomonas xylanilytica,* highest on 22nd January 2022, at 0.58% relative abundance.

Diversity indices of bacterial communities

Throughout section 2, comparisons have been made between the bacterial community composition of Monitored Soil as well as background soil using taxonomic assignments. By calculating several diversity indices, and using statistical analyses, true differences within and between the communities can be observed at species level.

Alpha Diversity

To determine alpha diversity, 3 unique measures were calculated at species level (Table 2). There is a significant correlation between Faith's PD index, ACE Richness index and Simpson's evenness index ($P < 0.01$). As one value increases, so do the others.

**Table 2.** Alpha diversity measures for sampled and background soil. Based on ASV's at species level (n=18).

| Sample type | Date | Faith's PD index | ACE Richness index | Simpson's Evenness index |
|---|---|---|---|---|
| Monitored Soil | 22/11/21 | 163.71 | 2024 | 0.67 |
| Monitored Soil | 13/12/21 | 97.18 | 1190 | 0.66 |
| Monitored Soil | 22/1/22 | 142.16 | 1831 | 0.59 |
| Background soil | 22/1/22 | 61.7 | 421 | 0.58 |
| Monitored Soil | 16/2/22 | 195.43 | 2321 | 0.61 |
| Background soil | 16/2/22 | 84.68 | 749 | 0.49 |
| Monitored Soil | 23/2/22 | 128.13 | 1507 | 0.49 |
| Background soil | 23/2/22 | 77.21 | 754 | 0.5 |
| Monitored Soil | 16/3/22 | 201.6 | 2320 | 0.66 |

| | | | | |
|---|---|---|---|---|
| Monitored Soil | 31/3/22 | 150.99 | 1923 | 0.58 |
| Monitored Soil | 30/5/22 | 16.47 | 37 | 0.36 |
| Background soil | 30/5/22 | 37.12 | 230 | 0.24 |
| Monitored Soil | 17/8/22 | 116.47 | 1365 | 0.5 |
| Background soil | 17/8/22 | 62.59 | 696 | 0.6 |
| Monitored Soil | 7/9/22 | 135.29 | 2020 | 0.66 |
| Background soil | 7/9/22 | 43.75 | 183 | 0.56 |
| Monitored Soil | 24/10/22 | 130.31 | 1495 | 0.54 |
| Background soil | 24/10/22 | 89.03 | 1252 | 0.66 |
| Monitored Soil | 25/10/22 | 182.47 | 2094 | 0.63 |

On average, alpha diversity (all indices measured) is higher in monitored soil, in comparison to background samples, meaning monitored soil has a higher diversity of bacteria present, in terms of alpha diversity. This difference in alpha diversity is significant, at $P < 0.01$ for ACE species richness, and for Faith's PD index, separating the communities in monitored soil from background soil. However, no significant difference is observed between the two datasets in relation to Simpson's evenness index, $P > 0.1$.

On average, Faith's PD for background soil was lowest from May-September 2022, and highest over January-February 2022 , and also October 2022. ACE richness index and simpson 's evenness were variable over the sampling period in background soil. For monitored soil, Faith's PD and ACE richness were variable over the sampling period, but Simpson's evenness was lowest over May-September 2022, and highest from December 2021-March 2022. A higher total rainfall depth was also seen over May-September 2022, compared to the December 2021-March 2022 period.

Beta Diversity

To calculate Beta diversity, weighted unifrac distance analysis was used.

Figure 19, shows PCoA analysis the separation in the two datasets indicating differences in the bacterial community composition between background and monitored soil samples. Testing this statistically using PERMANOVA statistical analysis of weighted unifrac distance values revealed

statistical significance at P <0.01, P=0.008. In terms of beta diversity, the bacterial composition at species level is different between background and monitored soil.

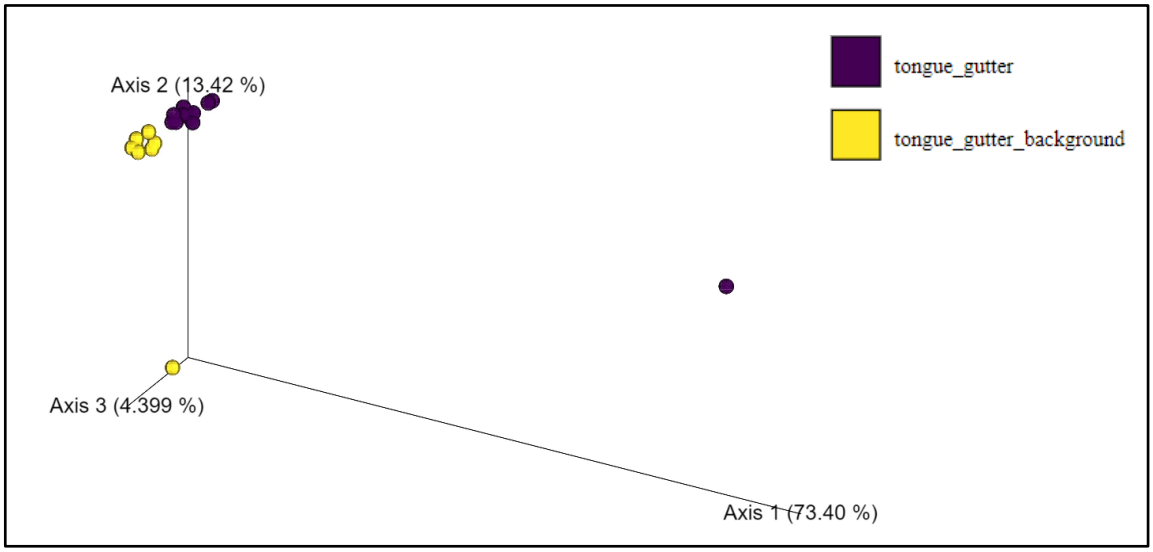


**Figure 19**. PCoA (Principle Coordinate Analysis) Emperor plot based on weighted unifrac distance analysis. n= 18

4.7.2 Tongue Gutter: Water and soil samples

Physicochemical analysis and estimation of bacterial abundance of soil and water

Table 3 shows the physicochemical data obtained from water samples, along with coliform counts. The pH of the water tends to be stable with the lowest pH recorded being 6.97, and the highest 8.32. Similarly, the soil pH, had a higher pH range, 7.4-8.32, evident over Winter 2021 and into Spring 2022, with a lower range, 6.97-7.62, recorded over Summer and Autumn 2022. The highest water pH was observed on 18th January 2022- this date also had the lowest rainfall recorded - 0mm. The lowest water pH was observed on, 17th August 2022.- this date also saw the highest recorded 5 day rainfall rainfall depth - 20.8mm. Despite this, when tested statistically the moderate correlation (R= -0.562) between pH and rainfall depth was not significant (P >0.1).

No statistically significant relationship was found between rainfall depth and coliform count (R= -0.15, P >0.1). In general, coliform counts for water samples had a - large and variable range- between 1200-9700 CFUs per ml of water (with the exception of the very high count seen on 18th January 2022). In comparison to soil samples, the coliform count was generally much lower in sampled water. When looking at the relationship between soil and water coliform counts, a positive correlation was found (R=0.29320. However, this is not statistically significant (P>0.1).

**Table 3.** Results of physiochemical analysis, and coliform counts of water samples (n=12).

| Date of visit | Average pH of water | Average Temperature/℃ | Cumulative Rainfall 5 days (120 hours) before sample /mm | Total number of coliforms / CFU per ml |
|---|---|---|---|---|
| 22/11/21 | 7.63 | 9.63 | 0 | 6700 |
| 13/12/21 | 7.43 | 11.6 | 10.4 | 6100 |
| 22/1/22 | 8.29 | N/A | 0.2 | 60000 |
| 16/2/22 | 8.24 | 8.7 | 21.6 | 6400 |
| 23/2/22 | 8.32 | 8 | 54.6 | 3200 |
| 16/3/22 | 7.87 | 9.67 | 9.0 | 3800 |
| 31/3/22 | 7.40 | 9 | 4.2 | 1200 |
| 30/5/22 | 7.62 | 10 | 1.8 | 9400 |
| 17/8/22 | 6.97 | 10 | 20.8 | 9700 |
| 7/9/22 | 7.35 | 10 | 21 | 6300 |
| 24/10/22 | 7.18 | 10 | 63 | 7000 |
| 25/10/22 | 7.31 | N/A | 61.2 | N/A |
| coliforms*= *E.coli* and *C .freundii* colony forming units | | | | |

Bacterial Community Composition

Taxonomic data is available for all water samples apart from those taken on the following dates: 22nd November 2021, 16th March 2022, and 25th October 2022 due to the extracted DNA not being of high enough quality/quantity to be sequenced (according to the threshold set by the NEOF sequencing facility in Liverpool).

Taxonomic analysis of water samples at phylum level, revealed 23 assigned phyla. The most abundant phyla include (from most abundant- on average, to least); Proteobacteria, Bacteroidetes, Firmicutes, Planctomycetes, and Acidobacteria. All 23 phyla were also shared by soil samples, with Proteobacteria being the most abundant phylum in both soil and water samples (Figure 20). Both soil and water also had Verrucomicrobia and Planctomycetes in relatively high abundance, Average Verrucomicrobia abundance in monitored soil was almost 4 times higher than in water, with Planctomycetes being almost 5 times higher.

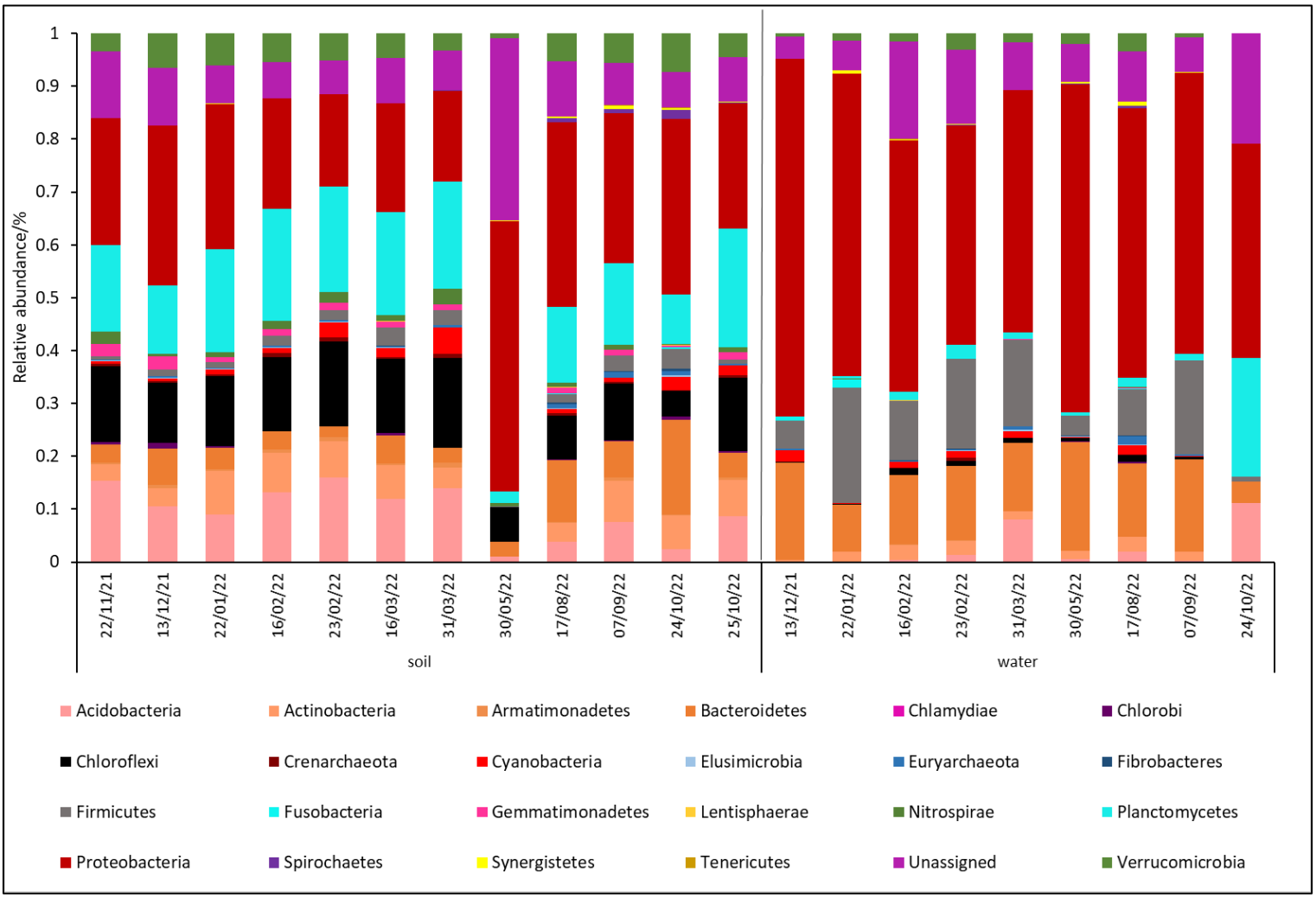


**Figure 20.** Comparison of bacterial communities present in water samples and soil samples, at phylum level (n=21). Unassigned taxa grouped into 'Unassigned' phyla category.

At Genus level, after removing genera that had a relative abundance of <0.3%, water samples contain 27 different genera. 10 genera found in sampled water, were also found in monitored soil. All 10 genera had an average relative abundance over the sampling period that was higher in water, compared to soil (Figure 21.)

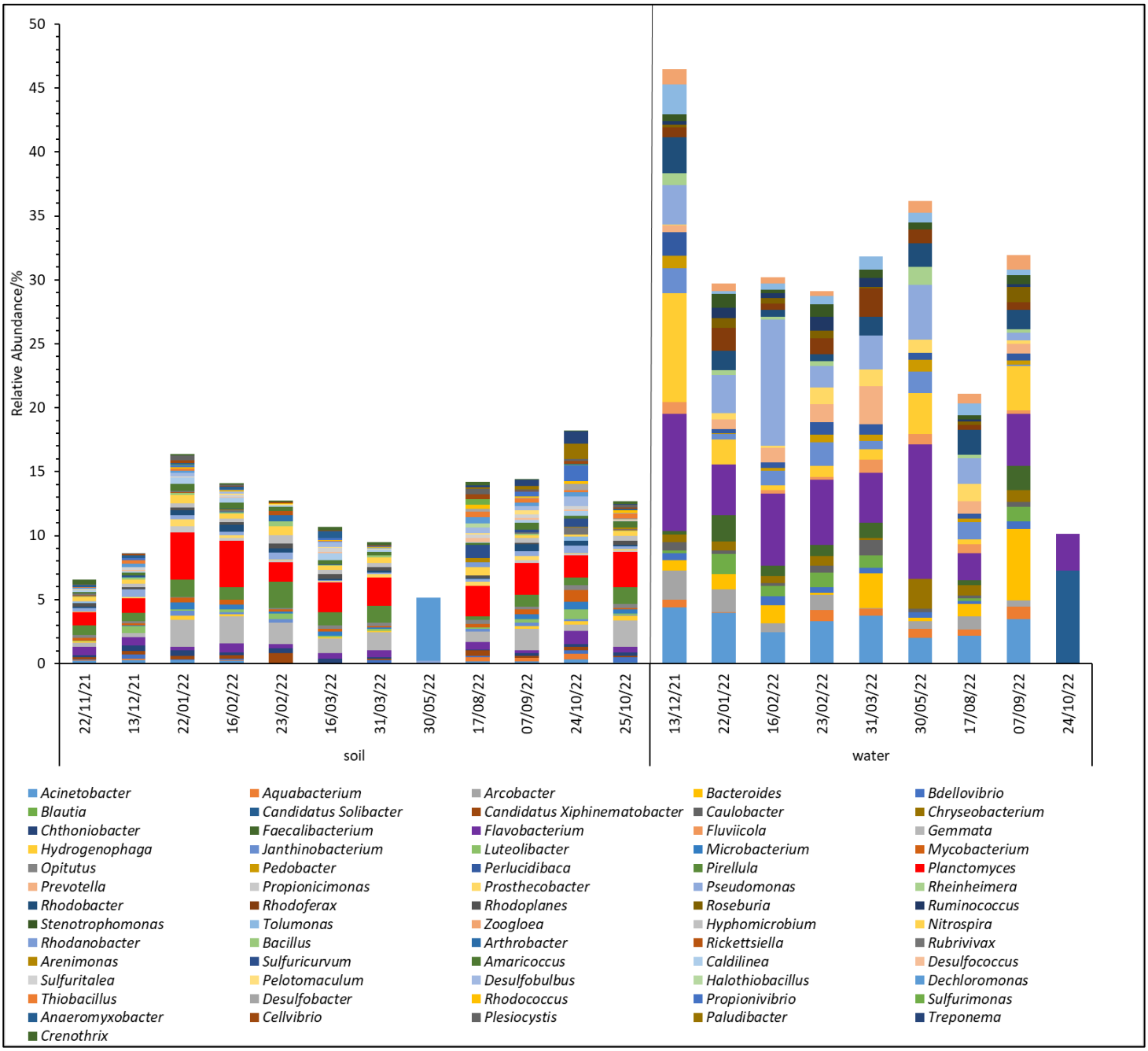


**Figure 21.**Relative abundance (> 0.3%) at genera level of water and soil samples (n=21).

On 13th December 2021, one third of detected genera displayed their highest measured relative abundance (2%-8%). This date also had the highest recorded water temperature (10˚C) (Table 3).When looking at seasonal trends in bacterial abundance in water, most genera had variable abundance, with no clear trends. However, *Prevotella, Rhodoferax and Ruminococcus* (not present in soil samples) had highest abundance from January -March 2022- when rainfall was on average lower, and water pH was higher compared to the subsequent period (Figure 22). Although rainfall was lower on average, a high rainfall depth was seen in February- after this period of increased rainfall, both *Prevotella* and *Rhodoferax* increased in abundance.

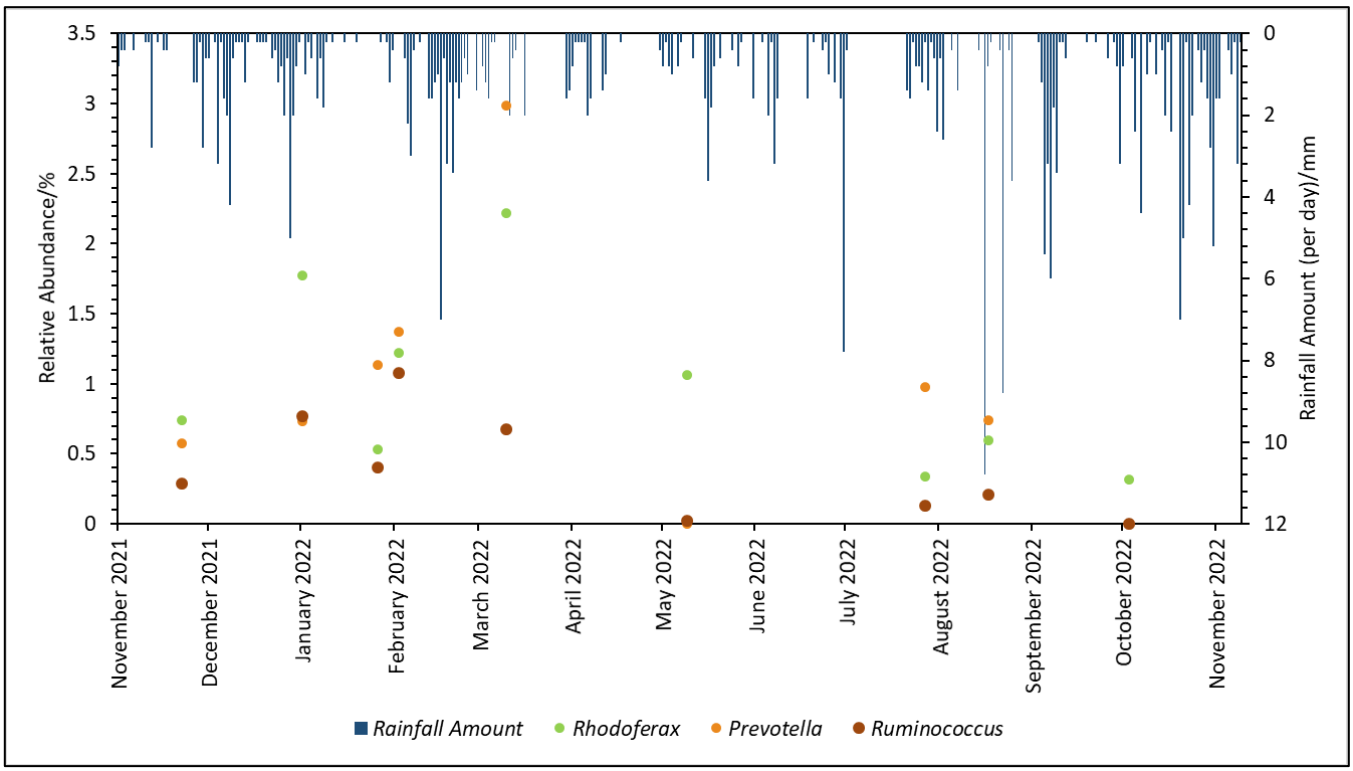


**Figure 22.** Abundance of genera in water samples at Tongue Gutter in relation to rainfall . Daily rainfall depth plotted on secondary axis.

Taxonomic analysis was further performed at a species level. Species detected at an average abundance of <0.1% for the whole sampling period were filtered out and the remaining species analysed and plotted. The threshold (0.1%) is higher for water than that of soil, due to the higher relative abundance and diversity of species generally displayed by water in comparison to soil samples. Sequences from all water samples were able to be assigned at species level, apart from the water sample taken on 24$^{th}$ October 2022.

With the less abundant species filtered out, for water samples, *Flavobacterium succinicans*, had the highest relative abundance, with a range of 0.1%-2.8% over the sampling period (Figure 23.) This was the only species present in all water samples and was also prominent in soil samples. Several other *Flavobacterium* species were also present in water samples, this includes: *Flavobacterium frigidarium, Flavobacterium gelidilacus*, and *Flavobacterium columnare.* Other genera, with multiple species present in water samples included*: Pseudomonas, Bacteroides* and *Ruminococcus species. Faecalibacterium prausnitzii* and *Prevotella copri*, also had relatively high abundance in water samples. Ranging from 0.27%-1.83%, and 0.44%-3% respectively. However, these species were not present in soil samples (at the levels tested for- >0.03%).

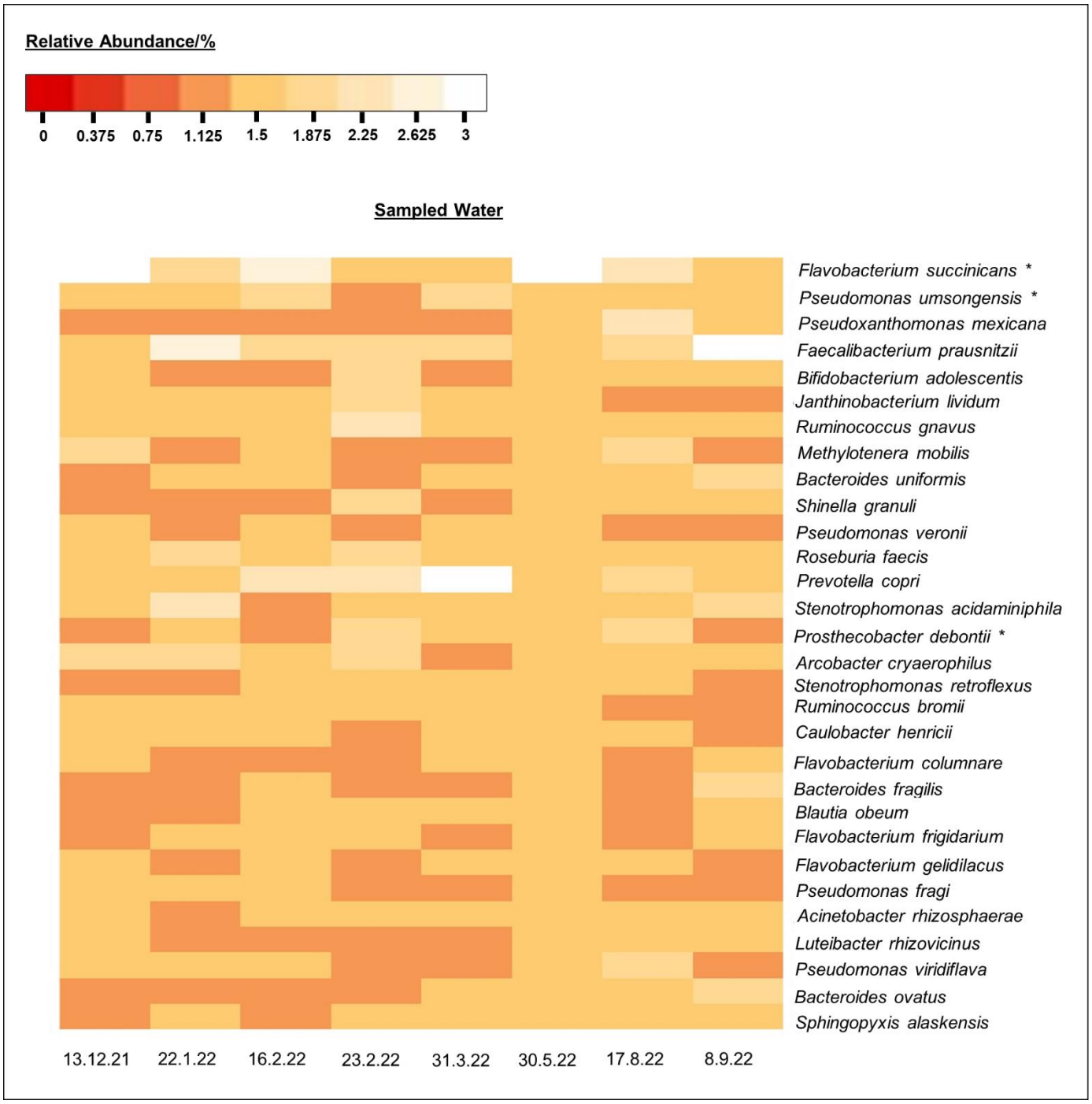


**Figure 23.** Heatmap of relative abundance of water samples at species level. Includes species that had an average relative abundance of >0.1% over the sampling period (n=8). Species labelled with a '*' are species that were also detected in monitored soil at >0.03% seen in Figure 18.

Interestingly, when looking at the species graphs for monitored soil (Figure 10) and water (Figure 18), in soil, more than in water, the community composition is due to rare species in lower abundance. Most of the abundance in soil samples is due to ASVs in very low abundance that are not present in the graphs (the most abundant bacteria only cover a small % of the total bacteria community). While in water, the most abundant species represent between 30-50% of the composition of the total community.

When looking at changes in the bacterial structure of water at species level over the whole sampling period, changes are seen in abundance of two faecal associated species*; Faecalibacterium prausnitzii*, and *Roseburia faecis,* (Figure 24). *F. prausnitzii* abundance appears to decrease at the start of February , when rainfall was lower, but shows increases in abundance after periods of increased rainfall – January, March, and August. *R. faecis*, similarly saw an increase in abundance when rainfall increased in January and August, . 2 species, *Flavobacterium gelidilacus* and *Bacteroides ovatus*, had higher abundance from May-September 2022, when rainfall was higher, compared to the rest of the sampling period. This period also had a lower water pH and higher temperature (Table 3).

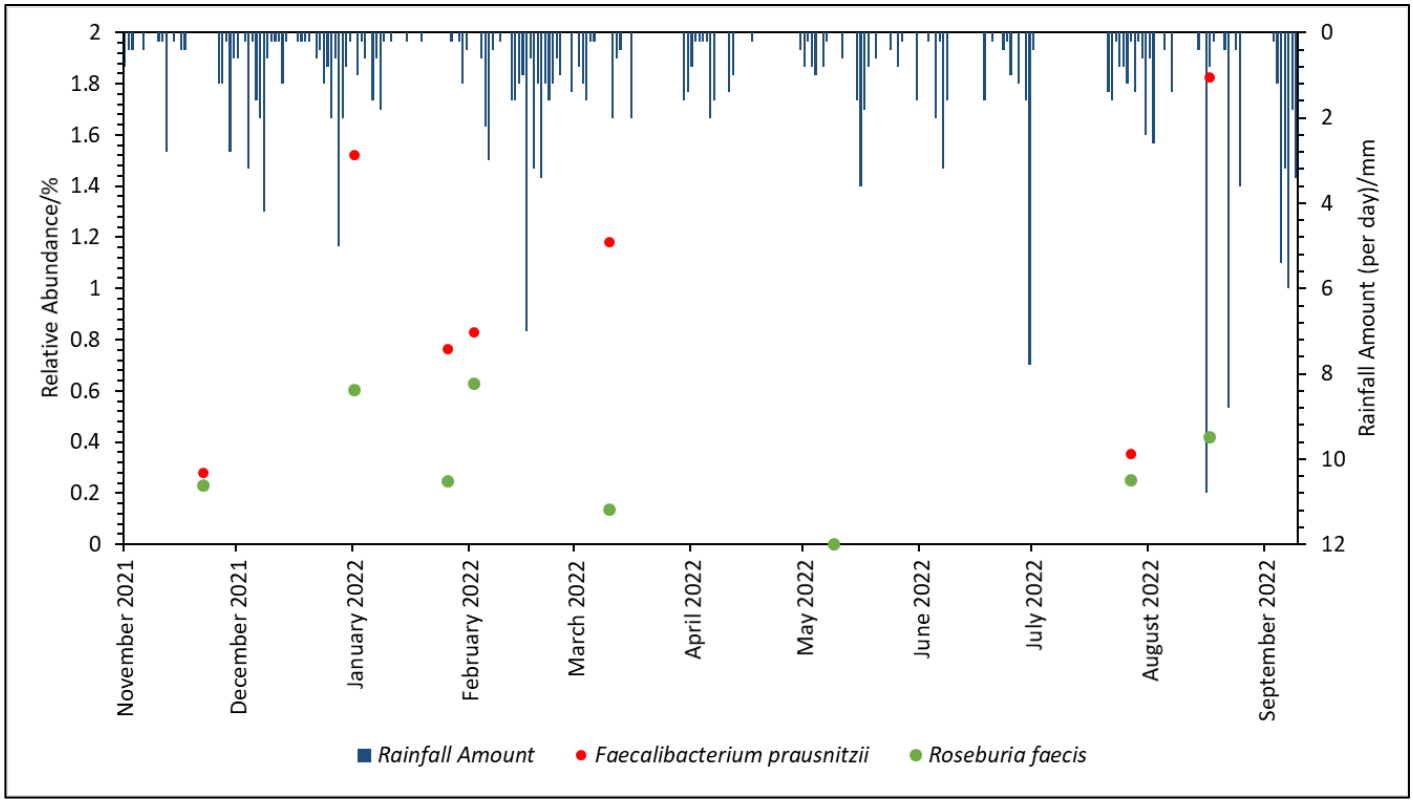


**Figure 24.** Abundance of species in water samples at Tongue Gutter in relation to rainfall . Daily rainfall depth plotted on secondary axis.

Diversity of observed bacterial communities

Throughout section 2, comparisons have been made between the bacterial community composition of water and soil samples by looking at taxonomy. Further, by calculating several diversity measures, and using statistical analyses, differences within and between the communities in the analysed samples can be elucidated.

Alpha Diversity

To determine alpha diversity, 3 unique measures were calculated at species level (Table 4) using the water samples.

**Table 4.** Alpha diversity indices calculated for water samples (n=8).

| Date of visit | Faith's PD index | ACE Richness index | Simpson's Evenness index |
|---|---|---|---|
| 13/12/21 | 42.19 | 408 | 0.44 |
| 22/1/22 | 106.04 | 1550 | 0.52 |
| 16/2/22 | 88.72 | 1089 | 0.51 |
| 23/2/22 | 57.84 | 510 | 0.47 |
| 31/3/22 | 36.99 | 417 | 0.61 |
| 30/5/22 | 65.96 | 689 | 0.38 |
| 17/8/22 | 95.34 | 1130 | 0.58 |
| 7/9/22 | 64.65 | 803 | 0.57 |

On 22nd January 2022 the highest Faith's PD and ACE richness metric were observed for water samples. Interestingly, the highest number of coliform counts were also recorded on this day (Table 3.) Evenness was highest on 31st March 2022. Over the whole sampling period Faith's PD and ACE richness was (on average) highest from May-September 2022 (when a higher rainfall depth was also seen) compared to the rest of the sampling period. Simpson's Evenness was variable over the sampling period. All alpha diversity measures were generally higher for soil samples, compared to water samples (Table 2 and 4.)

Beta Diversity

To calculate Beta diversity, weighted unifrac distance analysis was used. This allows differences to be tested between soil and water sample composition at a species level. Figure 28 shows the separation in the two datasets by the clustering of points- indicating differences in the bacterial microbiome composition between soil and water samples. Testing this statistically using PERMANOVA statistical analysis of weighted unifrac distance values revealed statistical significance , at $P = 0.011$. $P<0.05$ and so the significance is moderate- meaning that there is a meaningful difference between the bacterial communities of water and soil at species level. However, this observed difference is not as strong as the difference observed, in reference to beta diversity, between monitored soil and background soil (Figure 25).

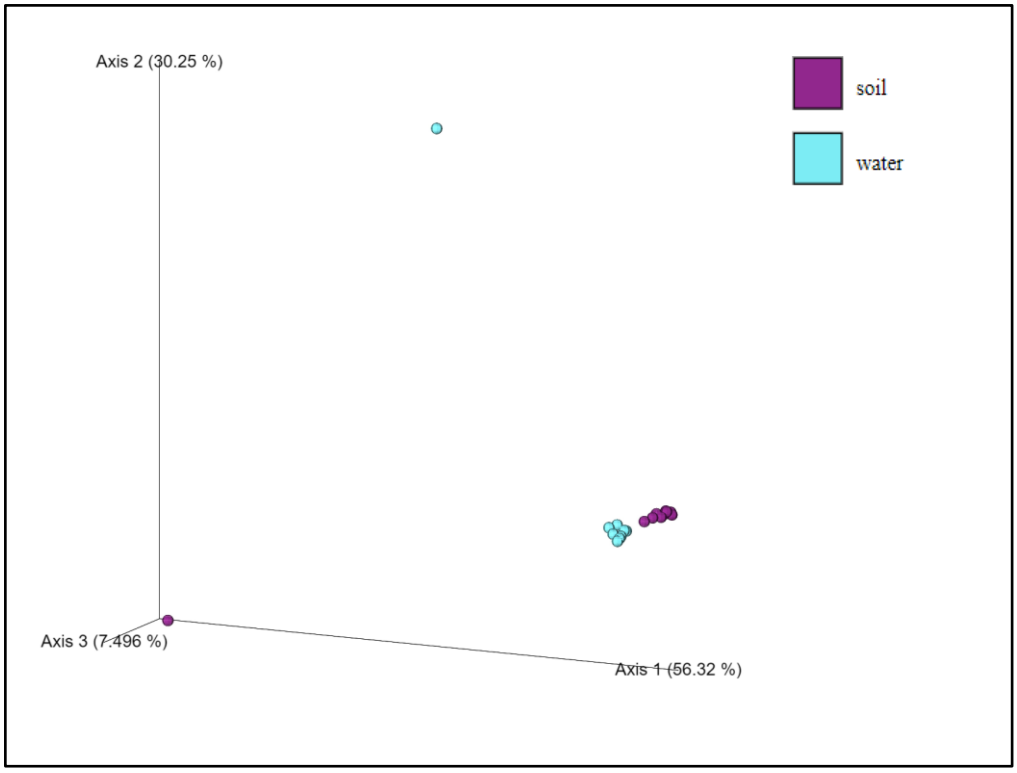


**Figure 25.** PCoA (Principle Coordinate Analysis) Emperor plot based on weighted unifrac distance analysis (n=22).

4.7.3 Endcliffe Park soil and background soil

Physicochemical analysis and estimation of bacterial abundance of soil

Table 5 shows the physicochemical data obtained from soil samples, along with coliform counts, and rainfall depth (cumulative 5 day). Over the sampling period, from November 2021 to September 2022, much like with the soil pH at Tongue Gutter, the pH showed a higher range over Winter and Spring 2022, 8.11-8.66, with a lower range over Summer and Autumn 2022, 7.04-7.64. Background samples at Endcliffe Park did not have any notable seasonal pH trends, however the pH, on average, was higher than that of Monitored Soil- 8.12, compared to 7.95 for background soil. Despite this, statistically the difference in pH between the two datasets was not significant, P >0.1.

Soil moisture ranged from 6.32%-50.01%, correlating positively with rainfall amount, although the correlation was not statistically significant. Background soil moisture was significantly lower than that of monitored soil. Average moisture content over the sample period for monitored soil was 28.8%, whereas background soil was 22.1%. The highest proportion of organic content was observed on 17$^{th}$ August 2022- this date also had the highest moisture content recorded for monitored soil. However the correlation between moisture content and organic content was not statistically significant.

Coliform counts ranged from 10000 to 1090000 per gram in monitored soil. In background soil, this range was lower, 10000 to just 340000. There was a positive correlation between 5 day rainfall depth and coliform count, however this is not statistically significant.

**Table 5.** Results of physiochemical analysis, coliform count, and rainfall amount (n=12).

| Date of visit | Average pH of soil | Cumulative Rainfall 5 days (120 hours) before sample /mm | Total number of coliforms* / CFU per g | Moisture content/% | Organic content by Loss on Ignition/% |
|---|---|---|---|---|---|
| 22/11/21 | 8.29 | 0 | 10000 | 25.67 | 4.26 |
| 13/12/21 | 8.16 | 10.4 | 280000 | 30.79 | 8.26 |
| 22/1/22 | 8.12 | 0.2 | 150000 | 38.52 | 7.46 |
| 16/2/22 | 8.39 | 21.6 | 360000 | 33.93 | 7.68 |
| 23/2/22 | 8.66 | 54.6 | 90000 | 25.81 | 13.73 |
| 16/3/22 | 8.11 | 9.0 | 260000 | 23.57 | 9.32 |
| 31/3/22 | 8.15 | 4.2 | 50000 | 25.69 | 16.52 |
| 30/5/22 | 8.46 | 1.8 | 10000 | 6.32 | 12.36 |
| 17/8/22 | 7.23 | 21.2 | 700000 | 50.01 | 24.34 |
| 18/8/22 | 7.04 | 21.2 | 810000 | 33.58 | 11.38 |
| 19/8/22 | 7.16 | 21.4 | 1090000 | 26.73 | 8.53 |
| 7/9/22 | 7.64 | 21 | 80000 | 22.30 | 7.23 |

coliforms*= *e.coli* and *c.freundii* colonies

Bacterial Community Composition

After sequencing, taxonomic analysis showed that there were a total of 22 assigned phyla present in soil samples, and 18 assigned phyla present in background soil samples (Figure 26). The following phyla were not present in background soil samples: Spirochaetes, Fusobacteria, Synergistetes, and Lentisphaerae. These phyla also had very low abundance in monitored soil, <0.05%. Over the whole sampling period, Proteobacteria was the most abundant phylum in soil samples at 16.82% on average.

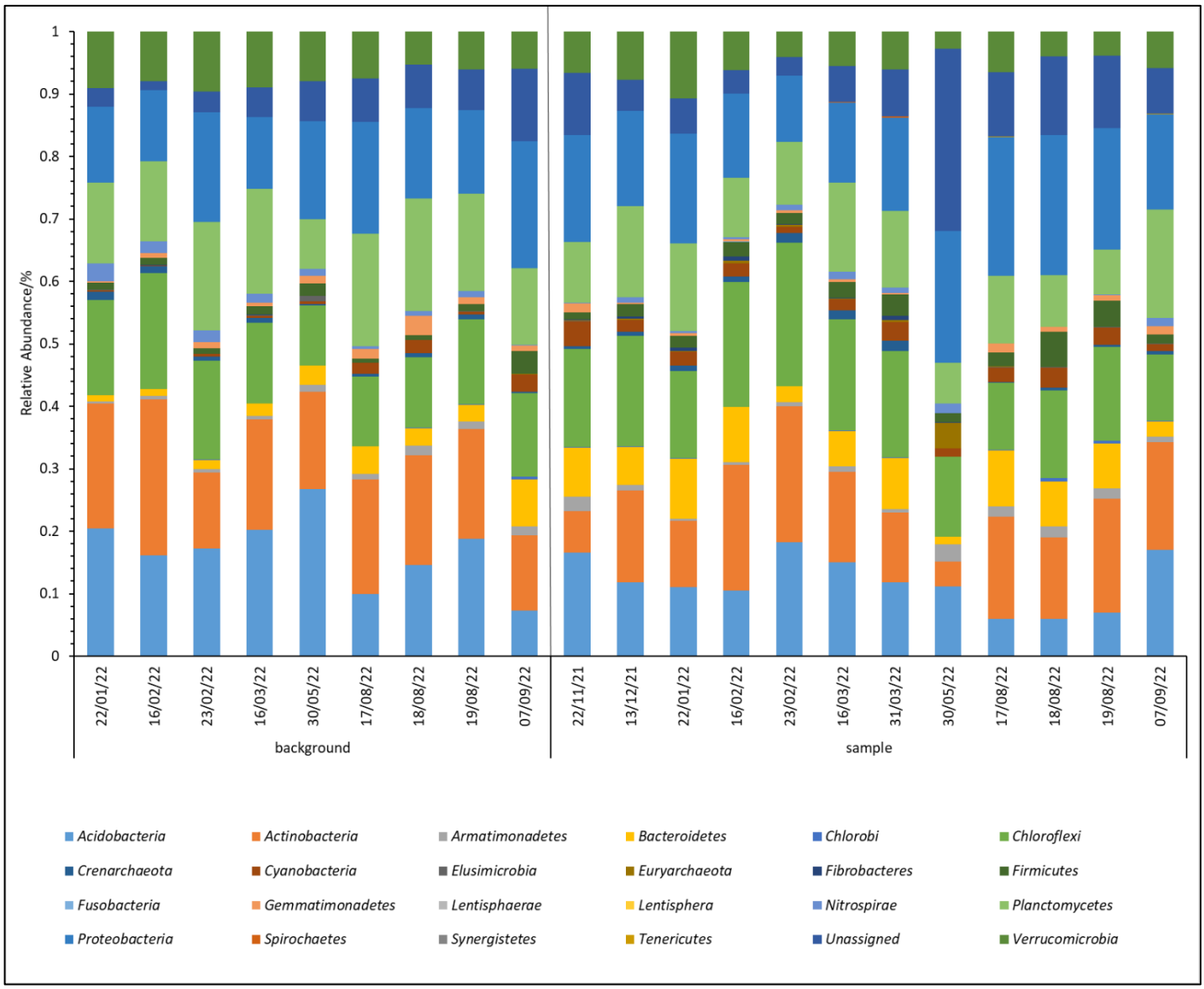


**Figure 26.** Comparison of bacterial communities present in soil samples and background soil samples, at phylum level. Unassigned taxa was grouped into 'Unassigned' phyla category (n=21)

After removing any unassigned genera, as well as genera at an average relative abundance of <0.3%, 25 genera were detected in samples in total (Figure 27). The most abundant genus for monitored soil and for background soil over the sampling period was *Gemmata, w*ith similar relative abundance levels at 1.92% and 2.15% respectively .

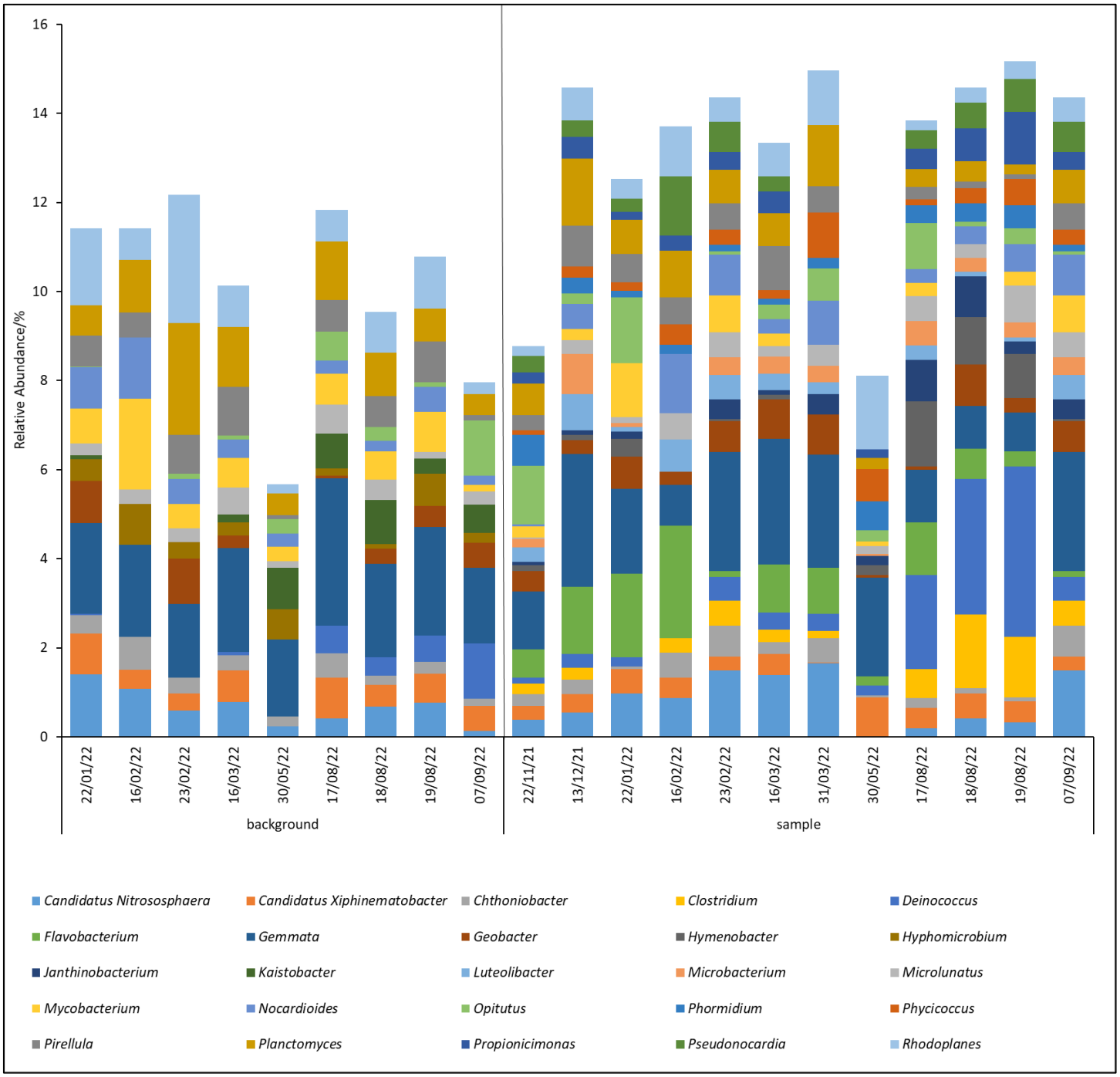


**Figure 27**. Relative abundance at Genera level of Monitored Soil and background soil. Includes only genera present at >0.3% in at least one sample (n=21).

Over the sampling period abundance of *Clostridium* (not in background soil), appears to be highest over August and September 2022- particularly high over August when rainfall amount was relatively high (half the months total rainfall (20mm) was seen 5 days before a sample visit). On the other hand, *Flavobacterium* (not in background soil), was highest in monitored soil over the Winter months- December 2021-February 2022, and actually increased in abundance each month (Figure 28). Particularly heavy rainfall (for the sampling period) was seen over February, with rainfall amount >140mm in total over 28 days.

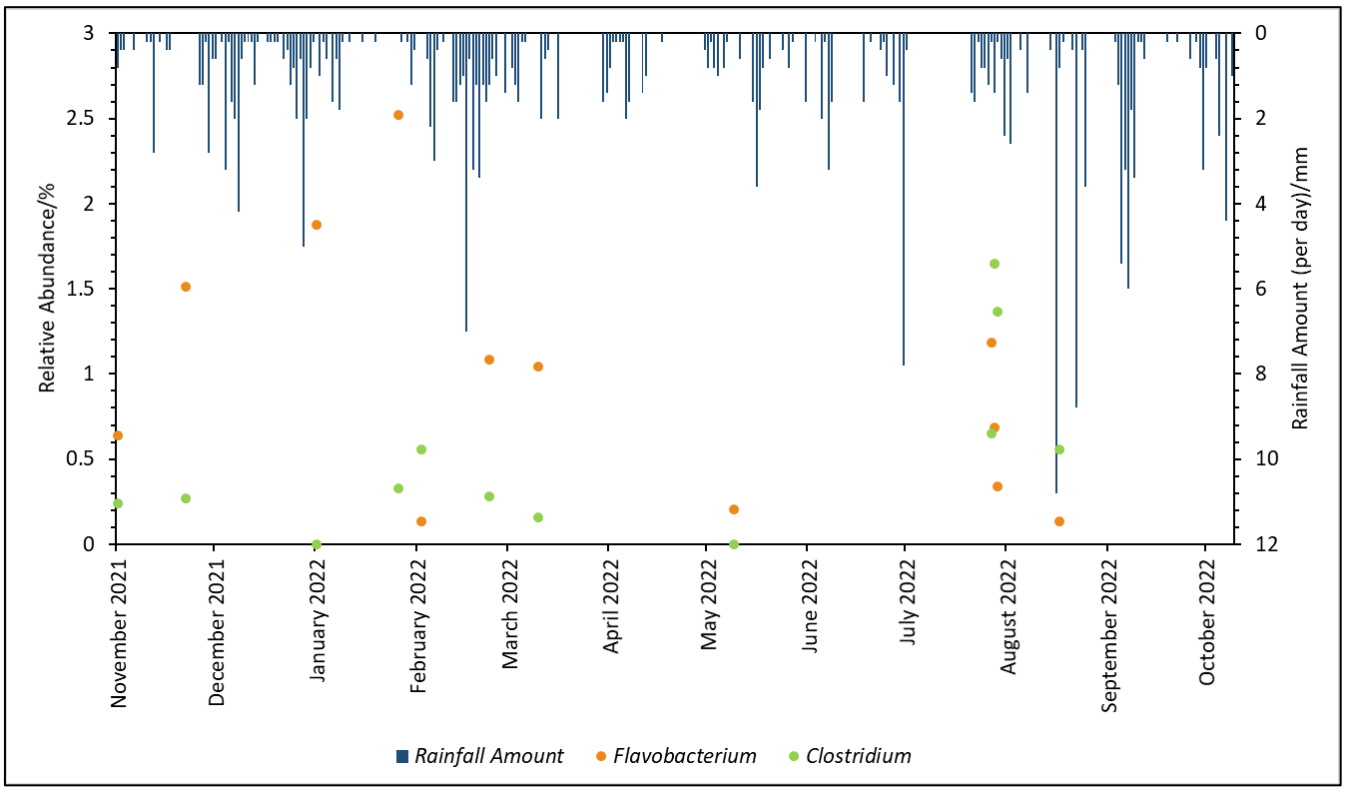


**Figure 28**. Abundance of genera in monitored soil samples at Endcliffe Park in relation to rainfall. Daily rainfall depth plotted on secondary axis

Species at an average abundance of <0.03% for the whole sampling period were filtered out and the remaining species analysed and plotted in Figure 29. 16 species were detected in monitored soil, and 8 detected in background soil. Two thirds of the species detected in background soil, are also present in monitored soil. For monitored soil, *Flavobacterium succinicans* was the most abundant species over the whole sampling period, highest on 16$^{th}$ February 2022 at 1.6%. *Clostridium* species were prevalent in Monitored Soil, with two species, *C. bowmanii* and *C. butyricum* present, however *Clostridium* species together only had a relative abundance of 0.12%.

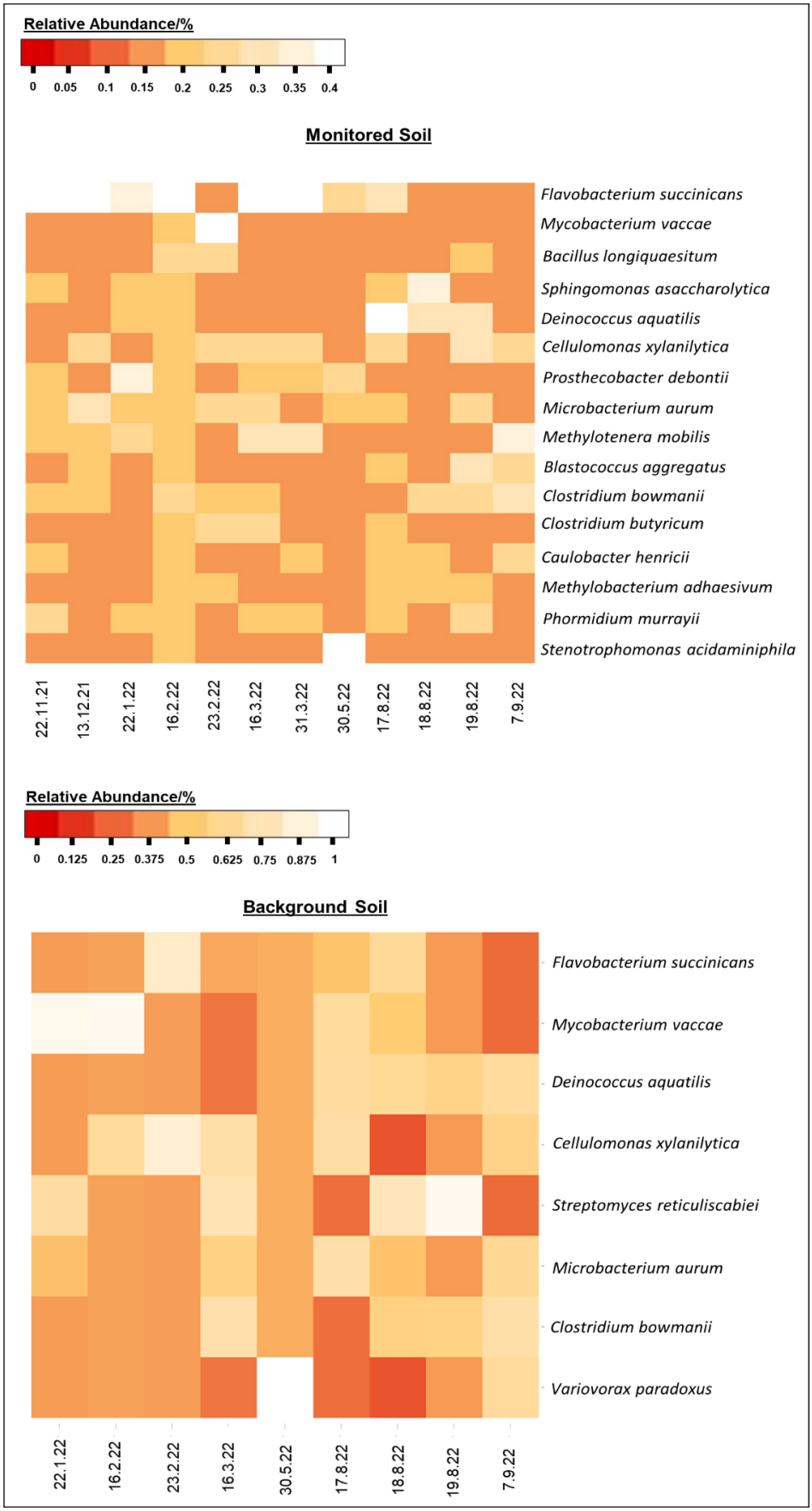


**Figure 29.** Heatmap of relative abundance of soil samples and background soil at species level. Includes species that had an average relative abundance of >0.03% over the sampling period (n=21).

Note the differences in key range for the two groups. This is due to the species present in background samples having much higher relative abundances than those in monitored soil. When looking at changes in the bacterial community of soil at species level over the whole sampling period, multiple species show elevated abundance over February – March 2022 in monitored soil. Such species include, *Bacillus longiquaesitum*, and *Methylotenera mobilis* (Figure 30*)*, as well as *Mycobacterium vaccae*, and *Flavobacterium succinicans*. Both February and August saw heavy rainfall, as well as increases in soil moisture content.

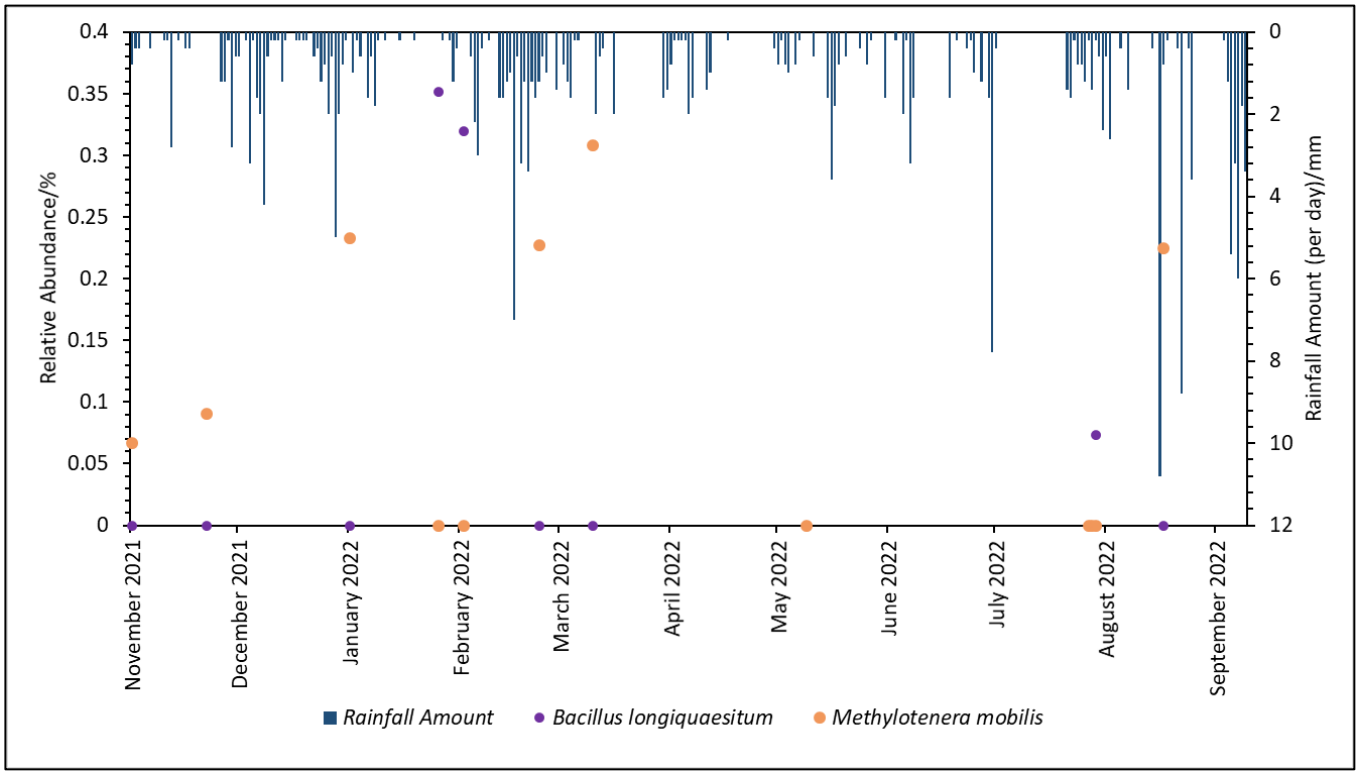


**Figure 30.** Abundance of species in monitored soil samples at Endcliffe Park in relation to rainfall. Daily rainfall depth plotted on secondary axis

Diversity of observed bacterial communities

Throughout section 2, comparisons have been made between the bacterial community composition of monitored soil as well as background soil by looking at taxonomy. By further calculating several diversity indices, and using statistical analyses, differences within and between the communities can be examined.

Alpha Diversity

To determine alpha diversity, 3 unique measures were calculated at species level (Table 6.)

**Table 6**. Alpha diversity measures for sampled and background soil (n=21).

| Sample type | Date | Faith's PD index | ACE Richness index | Simpson's Evenness index |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| Monitored Soil | 22/11/2021 | 72.01 | 1833 | 0.64 |
| Monitored Soil | 13/12/2021 | 41.72 | 1634 | 0.51 |
| Monitored Soil | 22/1/2022 | 148.75 | 854 | 0.60 |
| Background soil | 22/1/2022 | 41.08 | 675 | 0.51 |
| Monitored Soil | 16/2/2022 | 138.70 | 458 | 0.60 |
| Background Soil | 16/2/2022 | 65.43 | 420 | 0.62 |
| Monitored Soil | 23/2/2022 | 162.22 | 1109 | 0.56 |
| Monitored Soil | 16/3/2022 | 73.17 | 1171 | 0.56 |
| Background Soil | 16/3/2022 | 80.38 | 553 | 0.61 |
| Monitored Soil | 31/3/2022 | 76.18 | 797 | 0.46 |
| Monitored Soil | 30/5/2022 | 44.56 | 236 | 0.20 |
| Background Soil | 30/5/2022 | 35.89 | 1645 | 0.54 |
| Monitored Soil | 17/8/2022 | 215.28 | 1939 | 0.65 |
| Background Soil | 17/8/2022 | 167.24 | 255 | 0.58 |
| Monitored Soil | 18/8/2022 | 88.59 | 848 | 0.57 |
| Background Soil | 18/8/2022 | 20.28 | 1742 | 0.64 |
| Monitored Soil | 19/8/2022 | 52.61 | 523 | 0.57 |
| Background Soil | 19/8/2022 | 20.29 | 2100 | 0.68 |
| Monitored Soil | 7/9/2022 | 115.31 | 1679 | 0.70 |
| Background Soil | 7/9/2022 | 97.93 | 1250 | 0.58 |

Over the sampling period, soil samples had a significantly higher Faith's PD, and ACE Richness index metric, compared to background soil samples ($P < 0.01$). However, there was no significant difference between the evenness values for the two datasets. For soil samples, there was a positive correlation between Faith's PD, ACE Richness, and Simpson's Evenness ($P < 0.01$).For background samples, this was also true, however the correlation was only seen between Faith's PD and Richness at $P < 0.01$. All alpha diversity metrics for background samples were highest from August-September 2022, compared to the previous months. For monitored soil, simpson's evenness was variable over the sampling period, Faith's PD was also variable but had an increase over August-September 2022. ACE species richness was (on average) highest over November 2021-March 2022, but similarly showed an increase over August-September 2022.

Beta Diversity

To calculate Beta diversity, weighted unifrac distance analysis was used. This phylogenetically aware metric takes into account presence and absence of species, as well as abundance- and is thus quantitative. This allows differences to be tested between monitored soil and background soil composition at a species level. Figure 31 shows the plotted weighted unifrac values for the two sample groups. The two datasets are distinguishable from one another but do not seem to form discrete clusters. This indicates that the groups are similar in terms of beta diversity. Testing this using PERMANOVA analysis shows that there is a difference between the two groups, however this is only significant at P <0.05 (P=0.022).

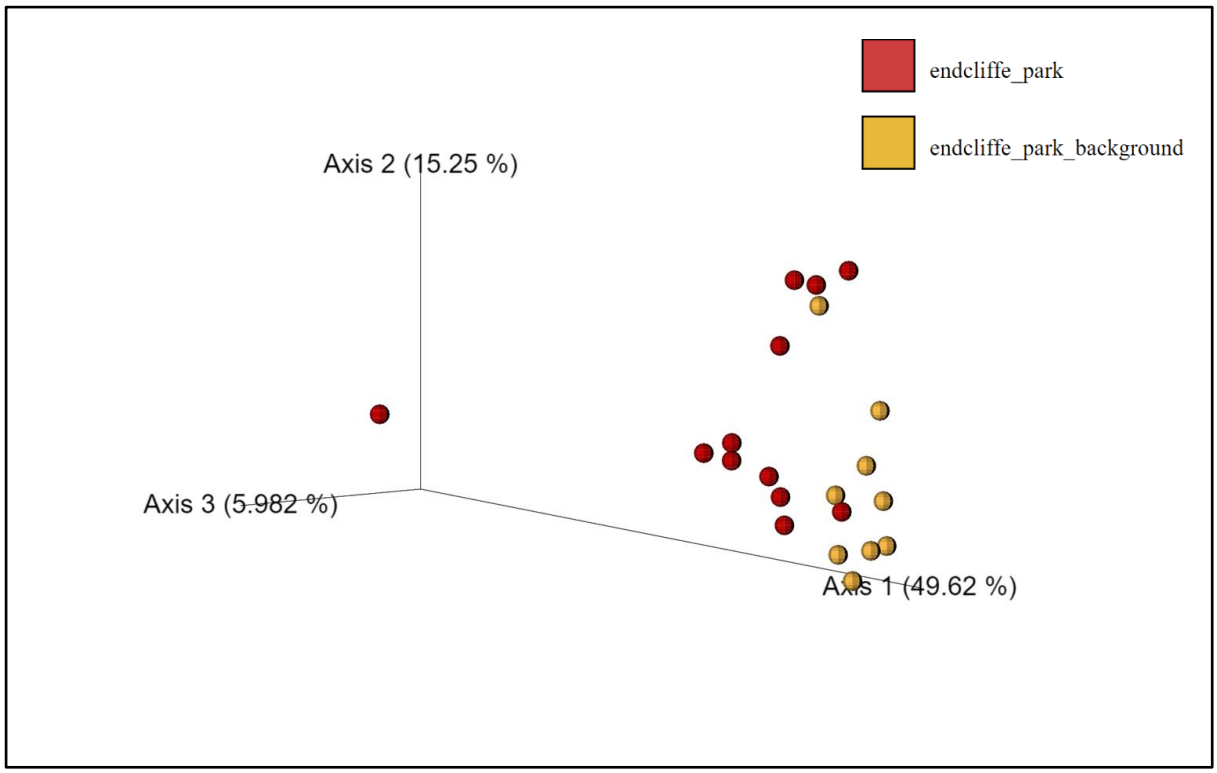


**Figure 31.** PCoA (Principal Coordinate Analysis) Emperor plot based on weighted unifrac distance analysis using soils samples (n=20)

4.7.4 Endcliffe Park water with comparison to Monitored Soil

Physicochemical analysis and estimation of bacterial abundance of soil

Table 7 shows the physicochemical data obtained from water samples, along with coliform counts and rainfall (5 day cumulative). Over the sampling period, fewer water samples were taken at this site, this is due to the nature of the site (water only available to sample on occasion of heavy rainfall for longer periods). Therefore, water samples cover the period between December 2021 and August 2022. Recorded pH ranges from 6.89 on 17th August 2022, to 8.48 on 16th March 2022. Seasonal trends cannot be determined, however average pH of water is slightly higher at 7.82 than that of monitored soil (7.95). Water temperature peaks at 9.3°C on 16th March 2022, with lower temperatures observed over the Winter months.

Coliform counts ranged from the lowest count of 10000 on 16th February 2022, to a high of 480000 on 17th August 2022. The highest 5 day rainfall depth was also recorded on 23rd February 2022 at 54.6mm. Too few data points were available to test for correlations. Water samples tend to have a lower coliform count compared to soil samples (Table 5), with the average coliform count for soil samples over the period being over double that of water.

**Table 7.** Results of physiochemical analysis, and coliform counts of soil samples (n=6).

| Date of visit | Average pH of water | Average Temperature of water /°C | Cumulative Rainfall 5 days (120 hours) before sample /mm | Total number of coliforms / CFU per ml |
|---|---|---|---|---|
| 13/12/21 | 7.12 | 7.6 | 10.4 | 250000 |
| 16/2/22 | 8.18 | 8.3 | 21.6 | 10000 |
| 23/2/22 | 8.21 | 7.7 | 54.6 | 0 |
| 16/3/22 | 8.48 | 9.3 | 9.0 | 50000 |
| 31/3/22 | 8.03 | 8.1 | 4.2 | 40000 |
| 17/8/22 | 6.89 | 9 | 21.2 | 480000 |
| coliforms*= *e.coli* and *c.freundii* colonies | | | | |

Bacterial Community Composition of water samples

After sequencing, taxonomic analysis showed that there were a total of 21 assigned phyla present in water samples (Figure 32). Only one phylum- Chlamydiae was present in water but not in soil. However, this phylum only appeared in one sample, at an abundance of 0.14%. In water and soil, the most abundant phylum was Proteobacteria, however in water samples the relative abundance on average across the whole sample period of this phylum was much higher at 34.33%, compared to just 16.82% in soil. On 16th March 2022 and 31st March 2022, only one phylum was present in water samples- Proteobacteria, at 67.7% and 33.5% (along with unassigned phyla).

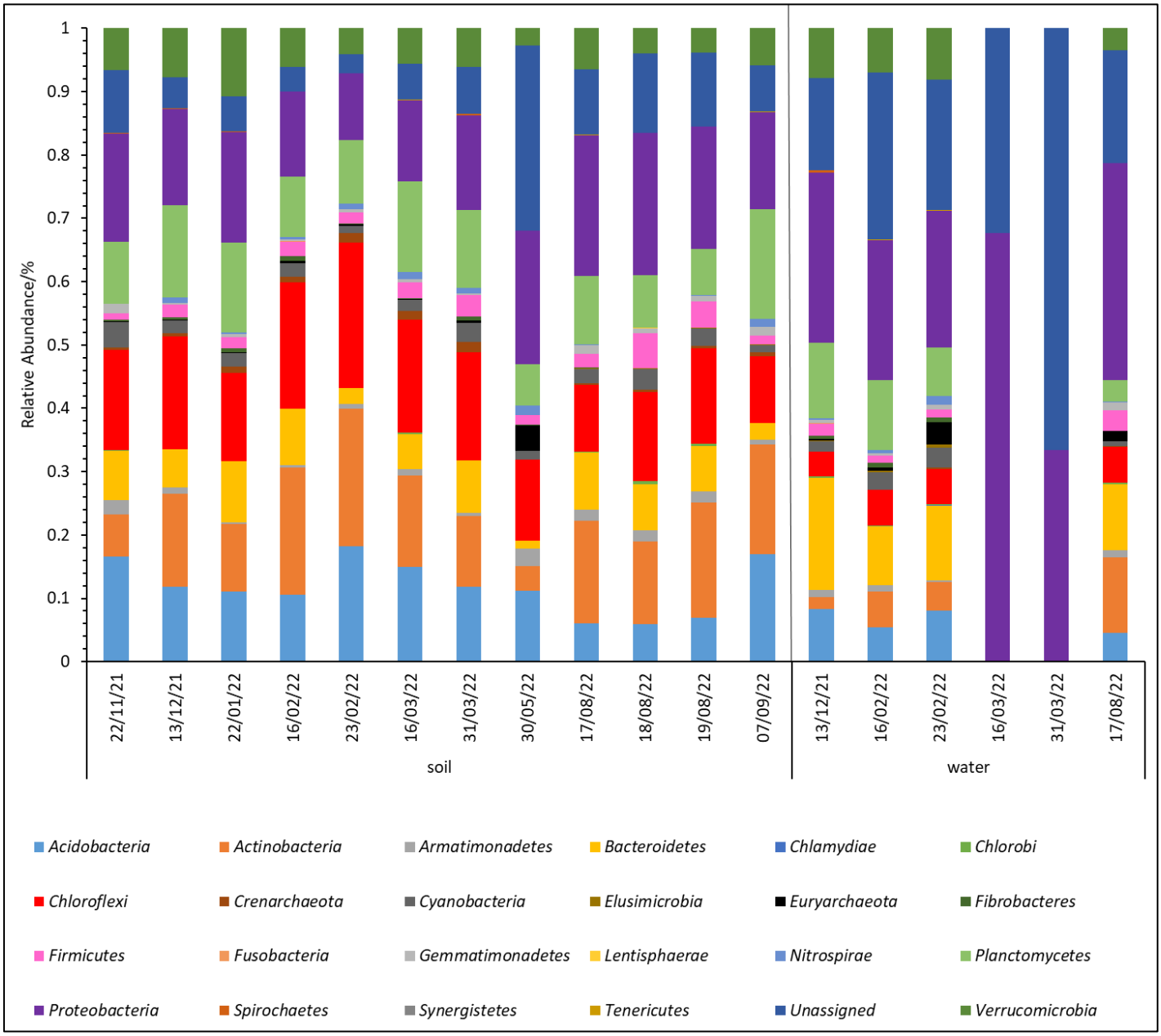


**Figure 32.** Comparison of bacterial communities present in water samples and soil samples, at phylum level (n=18). Unassigned taxa grouped into 'Unassigned' phyla category.

After removing any unassigned genera, as well as genera at an average relative abundance of <0.3%, 14 genera were detected in water samples (Figure 33). (DNA from samples from 16th March 2022 and 31st March 2022 were not able to be assigned at genus level and so are not included in the following data). Of the 14 assigned genera, one genus was unique to only water - *Pedobacter* . The most abundant genus in water on average for the whole sampling period was *Flavobacterium*- at 2.1% (Figure 37). This was more than double the average relative abundance of *Flavobacterium* found in soil samples, at only 0.9%. However, with only 6 water samples taken over 4 months at it is not possible to robustly interpret seasonal trends. Other genera with relatively high abundance in water included *Janthinobacterium,* and *Pseudomonas,* which had their highest abundance values on 17th August 2022 after a heavy rainfall event. This date also saw the lowest water pH and highest

number of coliform counts. *Janthinobacterium* and *Pseudomonas* were also found in soil samples, however the abundance of these genera were much higher in water than in soil.

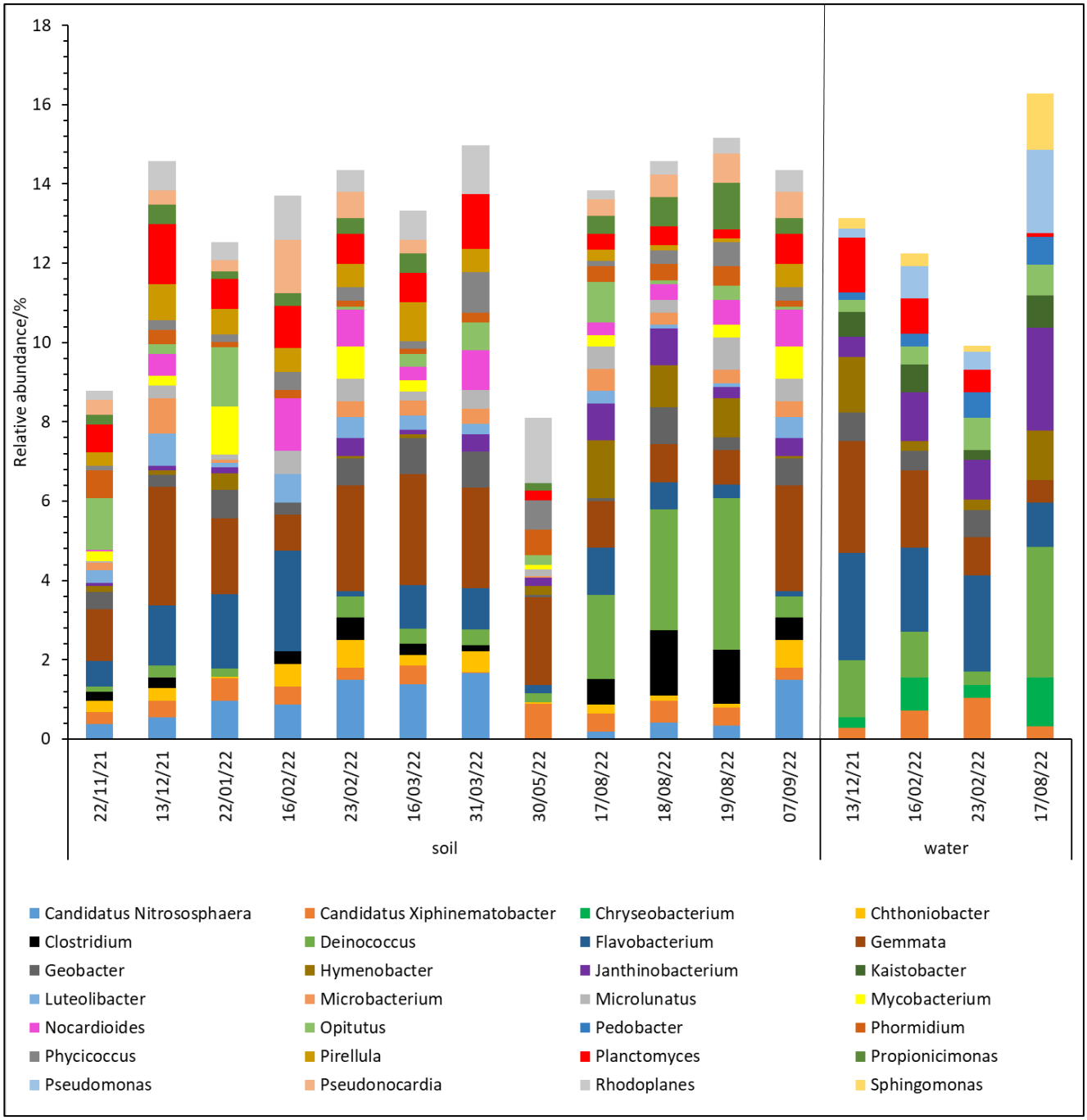


**Figure 33**. Relative abundance at genera level of water and soil samples (only genera at >0.3% abundance) (n=15).

Species detected at an average abundance of <0.05% for the whole sampling period were filtered out and the remaining species analysed and plotted in Figure 34. The threshold (0.05%) is lower for water than that of soil, due to the higher relative abundance of species generally displayed by water samples in comparison to soil samples.

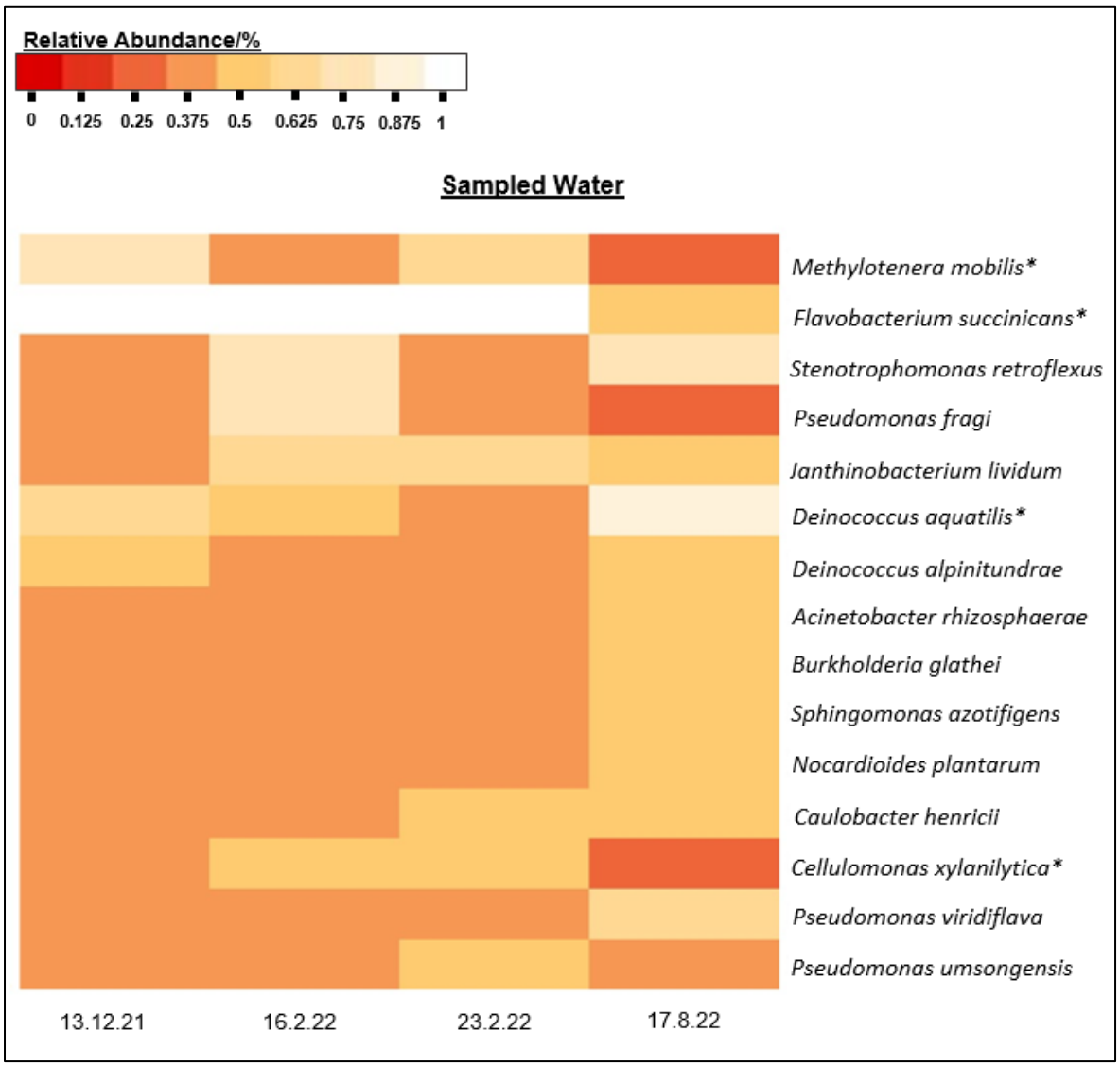


**Figure 34.** Heatmap of relative Abundance of water samples at species level. Includes species that had an average relative abundance of >0.05% over the sampling period (n=4). Species labelled with a '*' are species that are also detected in Monitored Soil at >0.03% seen in Figure 29.

Of the 15 species found at >0.05% abundance, one third were also detected in soil samples at >0.03% relative abundance (Figure 33). *Flavobacterium succinicans* was the most abundant species detected in both water and soil samples- 0.6% average abundance over the whole sampling period in water, and 0.47% in soil samples. 3 species of the *Pseudomonas* genus were also present in water samples, *Pseudomonas fragi, viridflava, umsongensis*, with *Pseudomonas frag*i. The three species together created an overall *Pseudomonas* species abundance of 0.21% on average over the sampling period. No *Pseudomonas* species are present in soil samples or background soil samples. On 17th August 2022 (after a heavy storm), a more diverse range of species were detected compared to the other sample dates. This date, saw the appearance of 5 species that did not appear in water samples taken on any other date. These species were: *Acinetobacter rhizosphaerae, Burkholderia glathei, Sphingomonas azotifigens, Nocardioides plantarum*, and *Pseudomonas viridiflava.* All of these species were present at between 0.22% and 0.25%.

Diversity of observed bacterial communities

Alpha Diversity

To determine alpha diversity, 3 unique measures were calculated at species level (Table 8). In water samples, Faith's PD and ACE Richness, were highest on 23rd February 2022, with Simpson's evenness highest on 17th September 2022.

**Table 8.**Alpha diversity measures for water samples (n=4).

| Date | Faith's PD index | ACE Richness index | Simpson's Evenness index |
|---|---|---|---|
| 13.12.21 | 94.25 | 1525 | 0.59 |
| 16.2.22 | 119.02 | 1464 | 0.44 |
| 23.2.22 | 140.84 | 1894 | 0.55 |
| 17.8.22 | 65.71 | 1027 | 0.68 |

Beta Diversity

Figure 35 shows the separation in the two datasets by the clustering of points- indicating differences in the microbiome composition between soil and water samples.

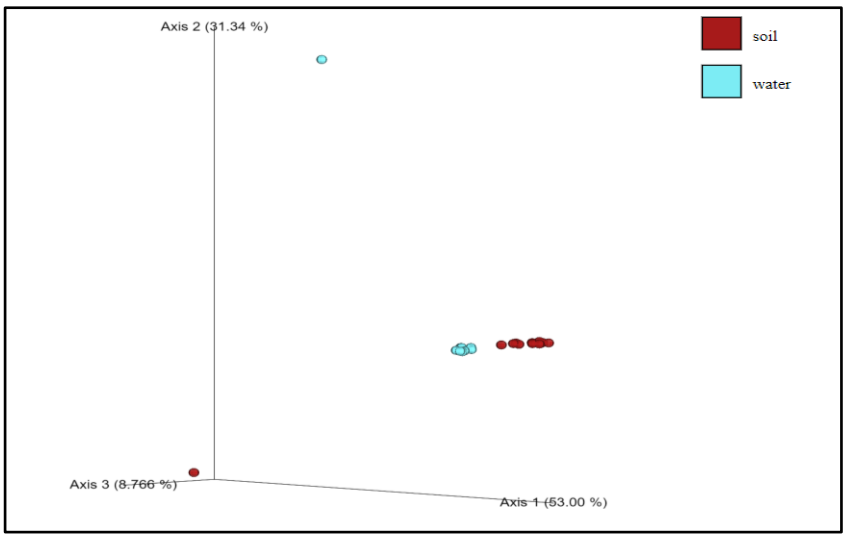


**Figure 35.** PCoA (Principle Coordinate Analysis) Emperor plot based on weighted unifrac distance analysis using soil and water samples (n=15).

Testing this statistically using PERMANOVA statistical analysis of weighted unifrac distance values revealed statistical significance, at P = 0.001. P<0.01 and so the significance was high- meaning that there was a meaningful difference between the microbial microbiomes of water and soil at species level. This difference between soil and water was more significant than the difference between background and Monitored Soil (Figure 35).

## 4.8 Discussion

### 4.8.1 Effect of frequent urban flooding on the soil bacterial microbiome

Through utilisation of ecological diversity measures, coliform counts, and next generation sequencing, a significant difference was found in the soil bacterial microbiome at locations subject to frequent urban floods, in comparison to locations that receive no flood water. It is recognised that flood waters can impact receiving soils, depleting nitrogen and oxygen content, altering physicochemical factors, and also changing the microbial structure (Francioli *et al.*, 2021, Furtak *et al*., 2020). However the floodwater ecosystem itself is variable, and many factors including; flood source, flow rate, and flood duration can influence the impact the water goes on to have. For example, on the one hand, flooding can lead to a reduction in bacterial activity, due to the destabilisation of the normal environment and creation of suboptimal conditions (González *et al,* 2016). Alternatively, the influence of flood water can actually increase bacterial diversity in soils and sediments in some instances, usually due to translocation of contaminants and pathogens carried in floodwater, and then deposited onto soils or surfaces (Mhuantong *et al.*, 2015). In this study, the latter has proven to be true in the monitored field sites. A significantly higher number of coliform counts were found in soil samples taken from both urban flood sites sampled, compared to background samples. The coliform counts here, were specifically coliform species- *C. freundii*, and *E. coli.* These species are traditionally used to indicate the presence of faeces (Edberg *et al.*, 2000; Odonkor and Ampofo, 2013; Guarino *et al.,* 1987), and so the higher numbers found in the monitored soil samples not only suggests that *E.coli* and *C.freundii* can be transferred from the urban flood water into the soil, but also that the urban floodwater itself is potentially contaminated with faeces, and thus other potentially pathogenic or dangerous bacteria.

At both sample sites, a significant difference was also found between the bacterial diversity of soils affected by urban flooding and background samples. At both sites, Faith's PD and ACE richness indices were significantly higher in monitored soil than background soil- indicating a higher species richness- a higher number of unique species present. It should be noted that alpha diversity alone is not enough to conclude bacterial community differences, with previous studies showing conflicting evidence on how flooding affects this diversity measure (Randle-Boggis, 2018; Shen et al., 2018). One laboratory based study found frequent floods increased alpha diversity of the bacterial soil community, however, a wider scale study looking at wetlands, found the opposite, with alpha diversity of the soil bacteria higher during a drought year, compared to a flood year (Randle-Boggis, 2018; Shen *et al*., 2018). With different studies showing contradictory results, there is no certain consensus on the effect of urban flooding on alpha diversity in soil. However, by also calculating beta diversity, one can be more confident in the influence flooding has on diversity as a whole.
In this study, Beta diversity, showed a significant difference when tested using the weighted unifrac distance metric (Lozupone and Knight, 2005). This metric shows that the bacterial composition at

species level is significantly different between background and monitored soil. The significant differences detected in both alpha and beta diversity suggest that urban floodwater is a driving factor in altering and even increasing bacterial diversity of soil at urban flood sites. This was also shown elsewhere in a lab based soil column study, with significant differences in soil bacterial diversity found using 16s rRna sequencing, however this study showed that the longer the flood duration the lower the alpha diversity, with diversity highest on the first day of the simulated flood (Sao et al., 2023). This could explain the higher alpha diversity values measured in monitored soil in our study, especially since the field sample sites selected were flooded frequently.

Diversity measures have shown there is a difference in bacterial community structure, but to see what that difference looks like taxonomically, sequencing was performed. Looking at sequencing data at a phylum level, revealed 22- 23 different phyla of bacteria present in soil at sites affected by urban flooding. Of the phyla found in soil samples, 82-87% were also present in background soil samples. The bacterial community found in both monitored and background soil are generally dominated by phyla commonly seen in the soil microbiome, such as: Proteobacteria, Acidobacteria, Actinobacteria, Verrucomicrobia, Chloroflexi, and Planctomycetes among others (Deng *et al.*, 2018). The interest here is in the phyla exclusively found in soil from the urban flood area, this includes: Spirochaetes,Fusobacter, Synergistetes, Lentisphaerae, and Chlamydiae. Species of these phyla include opportunistic pathogens, known to cause disease in humans, such as *Chlamydia trachomatis* in the Chlamydiae phylum (Becker, 1996), and *Fusobacterium nucleatum* in the Fusobacter phylum (Han, 2015). Species from the phyla Synergistetes are found in the human intestinal tract (Marchandin et al., 2010) and pathogens in the Spirochaetes phylum are known to cause Lyme disease via contaminated soil or water (Ditmar, 2011). The presence of these bacterial phyla indicate potential faecal contamination, but also indicate potential health risk.

Analysing further, and moving down to the taxonomic level of genus revealed further differences in the bacterial structure between background samples and samples from the flooded sites. *Pseudomonas* and *Flavobacterium* were present in affected soils at both of the sites sampled. It is not surprising that *Flavobacterium* is present- this genus is common in many environmental biomes, including river sediment, soils, and freshwater (Waśkiewicz and Irzykowska, 2014). This genus does not contain species known to be pathogenic to humans, however *Flavobacterium* has been (rarely) isolated from clinical samples in immuno-compromised patients with bacteremia, urine tract infections and soft tissue infections (Steinberg and Burd, et al., 2010). *Flavobacterium* more commonly poses a problem to aquatic environments- with many species being pathogenic to freshwater fish, causing devastating losses to populations (Loch and Faisal., 2015). From this study, it can be suggested that bacteria of this genus must have been transferred into the soil from the floodwater- due to it being absent in any background samples, and being present in the floodwater samples.

Much like *Flavobacterium*, *Pseudomonas* is a genus with bacterial species that are able to colonise a wide range of habitats (Madigan, 2005). Although at species level*, Pseudomonas* species pathogenic to humans was not detected in the soil in this study, *Pseudomonas viridiflava* was found at Tongue Gutter at ~0.2% (on average). An opportunistic pathogenic to plants. *P. viridiflava* can devastate crop yields, causing necrosis in plants, and like *P. aeroginus*, has shown antimicrobial resistance (Lipps and Samac, 2021). Additionally, it is thought that environmental Pseudomonas species, like *P.*

*viridiflava*, could even act as a vehicle for the transmission of resistance genes to human pathogens- such as *P. aeroginus (*Córdova et al., 2022).

4.8.2 Bacterial Microbiome of floodwater and bacteria of concern- Tongue Gutter

At the site- 'Tongue Gutter', water samples were collected on every visit, regardless of amount of rainfall. The outlet pipe at the site was spilling almost constantly, leading to the theory that this may be a misconnected outlet pipe.

When analysing the bacteria at genus and species level, two clear 'groups of interest' were seen. The first group is composed of bacteria associated with the human intestinal tract and indicate faecal pollution. Bacteria in this group include, *Roseburia, Faecalibacterium,* and *Ruminococcus* - these are found in the intestines and colons of humans, aiding in digestive processes and reducing inflammation in the bowels (La Reau and Suen, 2018; Lopez et al., 2017; Nie et al., 2021). At species level, several species belonging to the aforementioned genera were detected. This includes human gut bacteria such as - *Bifidobacterium adolescentis, Prevotella copri*, and Faeclibacteroum prausnitzii, along with two Ruminococcus species, *R.gnavus* and *R.bromii.* This is interesting, the presence of human gut bacteria, suggests that the floodwater is contaminated with human waste, This is expected given the source of the flood is a sewer outflow pipe, dilute sewage and thus FIB are likely to be detected. However if the system at this site was functioning as it should be, these types of bacteria should only (theoretically) be detected during heavier rainfall, when the sewer system becomes overloaded, however at this site *F.prasunitzii, R.gnavus*, and *R.faecis,* were detected on 7 out of 8 sample visits. This suggests human waste is discharged to this site frequently, and thus that the system is not functioning as expected, perhaps due to a misconnection. This raises the possibility for this to also be the case at other urban flooding sites, where an outflow pipe is the source of the flood.

Another group of bacteria found in the urban floodwater at this site contains genera known to cause disease amongst humans. Genera in this group include *Acinetobacter, Arcobacter, Pseudomonas* and *Bacteroidetes.* No pathogenic species belonging to the genera, *Acinetobacter* and *Pseudomonas* were detected. However, pathogenic species belonging to the *Arcobacter* and *Bacteroides* genera were found. 3 out of the 5 known *Arcobacter* species are pathogenic, causing mild disease in both humans and animals, with symptoms including diarrhoea, and stomach pain (Collado and Figueras, 2011). This study did find one such pathogenic species present in the sampled flood water.

*Bacteroides* is a genera of bacteria that are major colonisers of the human colon, and play an important role in human gut health- (Hooper et al., 2002). Species within this genera are not explicitly pathogenic, however, both *Bacteroides fragilis* and *Bacteroides ovatus,* are two species detected in the floodwater of this study that are known to cause clinically significant infection if spread into the bloodstream (Elsaghir and Reddivari, 2022). Disease is usually caused from the displacement of the bacteria already in the body, i.e. from intra-abdominal abscesses/ulcers etc, and so disease caused from this species via contact with floodwater is less likely.

After taxonomic analysis it is becoming apparent that the floodwater contains a range of bacterial species, originating from different ecological niches, but most pertinently ,species are present that may indeed pose a risk to human health, with potential pathogens and disease causing bacteria being present.

### 4.8.3 Bacterial Microbiome of Floodwater and bacteria of concern- Endcliffe Park

Given the origin of the urban flood at Endcliffe Park, it could be theorised that the flood water would only be made up of surface/rain water, and so bacteria associated with- sewage, and thus the human gut/faeces are unlikely to be found.

At this site, it is likely during lower rainfall that the floodwater being sampled is mostly made up of surface water, and as such the coliform counts are lower, whereas during higher rainfall, the gully at the site spills, meaning the floodwater now has another source of bacterial input, and thus increased coliform counts. However, a larger range of samples and thus data is required to draw any solid conclusions on the relationship between coliforms and rainfall amount in floodwater at this site.

Unlike the floodwater at Tongue Gutter, no bacterial species commonly associated with the human gastrointestinal tract were detected. This is positive in terms of the site functioning as it should- this indicates that the gully that surcharges at this site- - is not discharging water contaminated with sewage, but only discharging surface/rain water.

The only bacteria of concern found in floodwater at this site was *Pseudomonas,* with no *Pseudomonas* species detected in any monitored or background soil samples. 3 species of the *Pseudomonas* genus were present in water samples, *Pseudomonas fragi, viridflava* and *umsongensis*, with *Pseudomonas frag*i the most abundant of the three. However, much like the *Pseudomonas* species detected in floodwater at Tongue Gutter, none of these detected species are pathogenic or harmful to humans (Wu et al., 2015).

### 4.8.4 Transfer of bacteria between the soil and flood water interface

Having discussed the bacteria present in soil and water, it is apparent that some species are present in both sample types. Species usually found in water are present in soil, such as soil dweller *Methyltonera mobilis* found in flood water at Endcliffe Park, and vice versa, such as *Phormidium murrayi* , a freshwater bacteria, found in monitored soil at Endcliffe Park. This naturally leads to the theory that the bacteria are transferred between the soil and water interface (Abu-Ashour et al., 1994), using the floodwater as a vector (Caillon et al., 2021; Okaka et al., 2018). Transference from the water column, onto or into surfaces and soils is not uncommon, and it is widely known that many bacterial species do this- however many factors can affect and influence this including physico-chemical factors of both the soil and water, along with the structure and 'type' of the soil the bacteria is moving to, and of course the morphological features of the bacteria itself (Yang and van Elsas, 2018).

As mentioned, the floodwater can act as a vehicle, transporting bacteria during a flood event. Naturally when this flood water flows over soil, it has the ability to infiltrate into the soil. This means the bacteria may also do so. When discussing infiltration, the slope of the ground, as well as physical

properties of soil, including composition, particle density, and texture can significantly impact the water holding capacity of the soil, and also affect the infiltration rate of water into the soil (Patle *et al.*, 2019). Coarser soils usually have sand as a component, and tend to have higher infiltration rates, due to few, large, pore spaces. Soils with high clay contents, although reducing the infiltration rate, can hold more moisture- protecting the bacteria from desiccation, and also from UV (Adeniji *et al.,* 2013).

In this study, soil at both sites were found to be made up primarily of coarse sand, with some gravel. This would indicate that water, and thus bacteria suspended in the water, would be able to infiltrate into the soil at a relatively fast rate. The soil properties at the sample sites in this study, including structure and composition as well as moisture content, were suitable for facilitating the infiltration of floodwater and thus bacteria from the floodwater and into the soil.

During a flood event, the soil is likely already saturated- with the pore spaces filled with water , slowing the infiltration rate, and so other factors related to the bacteria itself, such as motility, hydrophobicity, and sorption ability will become more important in terms of bacterial movement. Despite this, of the bacteria found in soil and water at the sample sites, 3 are non-motile and so could rely on the floodwater as the sole means of transport. One such bacteria is *Prosthecobcater debontii*- found in soil and water at Tongue Gutter. *P.debontii* are non-motile and commonly found in freshwater (Hedlund et al., 1996). It is thought their presence in the soil is from deposition by the floodwater, especially since no *P.debontii* are detected in background soil samples.

Thus far, passive movement of bacteria has been discussed- using the floodwater as a means of translocation into the soil via infiltration. However some bacterial species are capable of active movement- using features such as flagella to aid motility. This means such bacteria could move independently of soil moisture/infiltration. Reasons for movement could be to reach more favourable conditions for growth and survival, (to reach an area with optimal pH, temperature, more nutrients)- known as chemotaxis- meaning bacteria could move from soil into water, or from water into soil, using the flagella to drive this movement (Abu-Ashour et al., 1994). Of the bacteria found in both soil and water samples in this study- indicating movement between the two- two species found are in fact flagellum driven. These species are *Flavobacterium succinicans* and *Pseudomonas umsongensis*.

When looking at *F.succinicans* specifically, there is always a higher relative abundance of this species in the water compared to the soil at Tongue gutter (Site 1). The average pH of water at this site - 7.7- is closer to the optimum pH for *F.succinincans* of 7, than the average pH of soil - 7.9. At Endcliffe Park, the relative abundance of this species varies over the sampling period. Interestingly a higher relative abundance of *F.succnincans* is found to again, potentially be dependent on pH of the soil and water. When the soil pH is closer to the optimum of 7, a higher relative abundance is observed in soil, and when the pH of water is closer to the optimum, then a higher relative abundance is seen in water. Thus, it could be theorised that *F.succinicans* moves between the soil and water interface to reach a more favourable environment.

Other factors directly related to the bacteria can also have an effect include sorption behaviour of the bacteria- their ability to attach to soil particles, and also whether the bacteria have the ability to

go on to form biofilms. Adsorption of bacterial cells from water to soil depends largely on the pore size of the soil and also the bacterial cell's motility. Water takes longer to infiltrate through a soil type with a small pore size, and so this means there is more time for bacteria carried in the water to form an attachment to soil particles (Jiang, 2005). Additionally, it has been found that motile cells have an increased adsorption rate, and decreased desorption rate to soil particles compared to non-motile cells. This is thought to be due to the morphological features used to 'move, i.e. flagella or pilli, also aiding in attachment (McCaulou and Bales ,1995).

Another factor that can affect the adsorption of bacteria to soil particles is the bacterial cell surface charge. Bacterial cells have a net negative electrostatic charge (at pH >2) due to the presence of teichoic acids or lipopolysaccharides present within the cell wall, meaning they would repel negatively charged soil particles, decreasing adsorption, but be attracted to soil particles with a positive charge, which would increase adsorption (Han et al., 2013). For example, bacteria including *Campylobacter*, are often attracted to clay particles. Clays can accumulate nutrients, reversibly bound NH4+ cations, which can then be passed to bacteria via ion exchange- therefore providing bacteria with nutrition to aid survival in the soil (Burford et al., 2003; Cuadros, 2017). However, bacteria can actually alter cell surface properties in a soil environment due to the secretion of extracellular polymeric substances (EPS) (Du et al., 2021), with studies finding conflicting evidence as to whether the EPS increases or decreases repulsion, and thus ability for bacteria to adsorb, due to electrostatic charge (Kim et al., 2009; Tong et al., 2010). The EPS and the ability to alter cell surface properties, including charge, is also important in biofilm formation (Donlan, 2001). EPS can neutralise the repulsive electrostatic forces between like-charged bacterial cells. By becoming part of the EPS matrix, the charge of an individual bacterial cell is reduced, allowing adhesion, rather than repulsion, between cells (Mahto et al., 2022).

Permanent adsorption of bacteria to soil particles can lead to biofilm formation. Formation of a biofilm can aid survival of bacteria but also increase the bacterial resistance to antimicrobials and antibiotics via up/down regulation of portions of their genes (Prakash et al., 2003). This is relevant, as it has been found that FIB, along with opportunistic pathogens, can attach themselves to existing biofilms, integrate, and thus survive for several weeks longer than their free-living counterparts (Wingender and Flemming 2011).

Wingender and Flemming's study, although looking at drinking water, featured bacterial species similar to those found in this study- *Campylobacter, Pseudomonas, Legionella*, etc. This suggests that perhaps FIB, and pathogenic bacteria found in floodwater in this study could attach to biofilms present in the soil, and survive for long periods of time, especially since biofilms in soil have been found to be 'the most common mode of soil bacterium survival' (La-Sa et al., 2022).

As well as the factors discussed above, naturally there are a plethora of other factors that can affect movement of bacteria between the soil and water interface including additional physicochemical factors: temperature of both soil and water, salinity of water, flow rate of water passing over soil, and nitrogen content of soil among others. However, this discussion has focused on the observed , measured, factors believed to have influenced the species detected at the sample sites. Although it appears that some bacterial species found at urban flood sites do have the ability to move between soil and water, no pathogenic species detected in this study appear to do so. This is encouraging in

terms of the long term risk posed by such bacteria- however this study only focuses on a small sample area (2 sites) and so does not mean conclusively that pathogenic bacteria could not move from water to soil. In fact many pathogenic species are known to do so. For example *E.coli* has been found by several studies to migrate from surface water into top soil layers- despite unfavourable conditions (Xing *et al.,* 2017; van Elsas *et al*., 2010).

4.8.5 Seasonal changes in bacterial communities at urban flood sites

At both locations sampled, seasonal changes are seen in the bacterial community , seemingly driven by rainfall , and thus also the source of the flood- which is quite different at the two sites in this study.

In the floodwater sampled at Tongue Gutter, several faecal associated species were found to be most abundant following periods of heavy rainfall in March and August 2022. Such species that increased in abundance included: *Faecalibacterium prausnitzii, Bifidobacterium adolescentis*, and *Roseburia faecis.* All 3 species are known to be found in human faeces, and were detected in floodwater samples only (not in soil samples) (La Reau and Suen, 2018; Lopez et al., 2017; Nie et al., 2021).It can be theorised that higher abundance of such species following heavy rainfall is due to contaminated water and surface water discharging from the outlet pipe at a higher rate and volume during heavy rain (Xu et al., 2017). Higher numbers of faecal bacteria- *E. coli* and intestinal enterococci, after heavy rainfall have similarly been seen in a study tracking contamination of a river by a discharging CSO in Paris (Passerat et al., 2022). The outlet pipe at Tongue Gutter spilled on every visit, even during days with no rainfall (thought to be due to a misconnection), and so the faecal contaminants present in the spilled water (*F.prausnitzii, B. adolescentis, R. faecis*) were still present, even on visits when rainfall amount had been lower. Overall at Tongue Gutter, rainfall amount appeared to be a driver of changes in the bacterial community of soil, but for the bacterial community in water especially, the role the source of the flood, and thus contamination plays, is also of importance.

At Endcliffe Park, in the soil samples, several species had an increased abundance during the two months the highest rainfall amounts were seen - February and August 2022. Such species included, *Flavobacterium succinicans, Mycobacterium vaccae*, and *Cellulomonas xylanilytica*. However, these species also showed similar increases in background soil over the two months. This could indicate that the change in abundance is driven generally by rainfall amount, and not necessarily impacted only by the application of floodwater onto the soil. Especially since it has seen shown that increased rainfall can alter soil geochemical properties, increasing soil moisture, nitrogen and carbon content, creating a more favourable enviroment for bacteria to survive (Cruz et al., 2012,Van Gestel et al., 1993) However, *F. succinicans* was also detected in floodwater samples, with an increase in abundance in February and August too. *F. succinicans* has been found in both soil and freshwater environments (Good et al., 2004; Zhang et al., 2006) , and so it's source cannot be determined, however it could be theorised that due to the increase in abundance when rainfall amount is higher that this species originates from floodwater spilling from the gully. However this species is also found in background soil samples, and so may be present in the soil community naturally, and moves into the water during a flood event.

For water samples from Endcliffe Park, due to limited data, seasonal changes could not be distinguished (data only available for December 2021, and January, February and August 2022).

### 4.9 Conclusions

From this research it can be concluded that:

- The bacterial structure of urban floodwater is diverse, and often influenced by the source of the flood's pollution. Sites that have floodwater contaminated by sewage have higher (alpha) bacterial diversity.
- The variability of bacterial structure of soil samples is due to a large number of species at a very low relative abundance (rare species) but in water , the most abundant species are responsible for the majority of the composition of the community .
- Urban floodwater can contain pathogenic species that can compromise environmental and public health, along with those that can cause human disease, including species of *Arcobacter* and *Bacteroides*.
- Species of the *Pseudomonas* genera, known to exhibit antimicrobial resistance were detected in soils at urban flood sites.
- Soils at urban flood sites have a long term altered bacterial structure compared to soils not affected by flood water, quantified by measuring beta diversity and determining statistically with PCoA and PERMANOVA analysis.
- Species found in urban floodwater are also found in soils at the site, meaning the bacteria may be transferred between the soil and water interface, using the floodwater as a way to infiltrate into soil, or via active movement- using morphological features to aid motility.
- Bacterial communities in both soil and floodwater at an urban flood site show seasonal changes, including changes in diversity, and bacterial abundance, which may be driven by bacterial preference to different moisture and temperature conditions that fluctuate over the seasons.
- Bacterial communities show a response in the short term to individual 'heavy' rainfall events, evident as increases in bacterial abundance.

This study provides new insights on the presence, risk and behaviour of bacteria at urban flood sites. Although it has been shown that potentially pathogenic and disease-causing bacteria are in fact present in the soil and water at urban flood sites- further epidemiological and medical research would need to be carried out to fully evaluate the risk they may pose to public health. This could include research into the bacterial survival rates under different conditions, but also research into the exposure routes, and also dose required for illness to occur. Although focusing on the long term effects, this study did also find bacterial community responses to short term heavy rainfall events. To try and understand the short terms effects further another field study could be carried out, sampling more frequently (perhaps daily) after each heavy rainfall event, to collect data on the dynamics at play over a shorter timescale- complimenting the data collected from this long term study. With urban flood events likely to increase in frequency over the coming years, data from this study consolidated with information from aforementioned suggested studies, could be used by authorities to determine the safety of areas subject to frequent urban flood events, especially in terms of bacterial risk, and risk posed to public health.

# Chapter 5

# Exploring short-term microbiological hazards of urban floods in the UK through bacterial community analysis

## 5.1 Abstract

This study investigated the short term microbial risks associated with urban floodwaters, focusing on the presence and dynamics of faecal associated bacteria and potential pathogens in the 3 days following a flood event.

Microbial dynamics were investigated in urban floodwater and soil samples taken from two field sites over 3 days following an urban flood event. Methods used included coliform counts, measurement of physico-chemical factors, and Illumina sequencing of the 16s rRNA gene. Bacteria associated with faeces including *Faecalibacterium prausnitzii* , and potentially pathogenic bacteria- *Dialister invisus* were detected in floodwater at one of the urban flood sites sampled (*F. prausnitzii* at both visits, *D. invisus* at October visit only). Despite the bacterial risk found in flood water, monitored soil samples showed no evidence of faecal contamination or pathogenic species in the 3 days following a flood event, suggesting that flood water rather than soil serves as the primary source for bacterial risks. Changes in the bacterial community were seen over the 3 sampling days in both soil and water, with the presence of faecal contamination thought to have altered the soil bacterial community structure, including reducing diversity. Evaluation of traditional coliform counts and DNA-based sequencing outcomes revealed the limitations of relying solely on culture-based methods for assessing faecal contamination in urban flood sites. While coliform counts indicated faecal pollution, 16s rRNA gene sequencing identified a broader range of bacterial species of concern, including potential pathogens.

This study contributes valuable insights into the short-term bacterial dynamics of urban flood events, emphasising the importance of, flood origins, and surveillance methodologies in assessing microbial risks to public health.

**Keywords:** short term risk, urban floods, soil-water interface

## 5.2 Introduction

In the UK, sewer systems, once considered pioneering feats of engineering during their conception in the Victorian era, now face significant challenges due to ageing infrastructure, blockages, and pressures arising from increasing urbanisation- impermeable surfaces and increased population densities (Campos et al., 2025). These factors combined with increased occurrence of extreme weather events (mainly due to climate change effects) means spills from combined sewer systems are happening more frequently (Hirabayashi et al., 2013). Spills that occur in an urban area- forming

substantial areas of standing water, are referred to as an urban flood (O'Donell and Thorne, 2020). Urban flooding poses significant dangers to both people and infrastructure. Rapidly rising water levels can lead to property damage, displacement of members of the public, and often disruption to transport if streets become flooded. The inundation of public spaces and sewage systems with urban flood water can create health hazards due to contamination with faecal bacteria (from the sewage system) present in the flood water.

The bacterial communities present in urban floodwaters can pose a significant threat to public health; this is mainly due to the fact that urban floodwaters often contain sewage, and thus faecal associated bacteria and pathogens. Among the most prevalent pathogens in the public realm are *Escherichia coli (E. coli*), *Salmonella* spp., and *Campylobacter* spp., which thrive in the nutrient-rich environment of floodwaters (Ten Veldhuis et al., 2010). These bacteria can enter the human body through ingestion, inhalation, or contact with contaminated surfaces, leading to disease or infection (Arnone and Perdek Walling, 2007).

Not only are the bacteria a risk in the floodwater but it is known that bacteria from the water column can transfer onto surfaces the floodwater moves over- depositing bacterial contaminants (Taylor et al., 2013). For example(*E.coli* commonly present in wastewater, has been found by several studies to migrate from surface water into top soil layers- despite the unfavourable conditions (Xing et al., 2017; van Elsas et al., 2010). *E. coli* has been found to survive at temperatures of less than 30 degrees, with very little nutrition, for almost a year in soils, with it taking 2-3 months for *E.coli* concentrations in soil to significantly decrease following an input event (application of wastewater or manure) (Jiang et al., 2005). In contrast, looking at *E. coli* transferring onto other urban surfaces, one study , that of Scoullos et al., 2019 found that when *E. coli* was deposited via simulated floodwater on impermeable urban surfaces -concrete and asphalt- that no viable *E.coli* cells were recovered after 24 hours. This study found this was mainly due to the effects of sunlight on soil moisture content (Scoullos et al., 2019). Another study looked at *E. coli* and *Enterococcus faecalis* (*E. faecalis*) survival on wood and brick surfaces after a simulated flood. A significant increase in bacterial die off was observed over 8 days, with the biggest reduction in viable bacteria in the first 3 days- namely due to lack of nutrients and again, reduction in moisture (Taylor et al., 2013).

Once transferred to the soil environment, bacteria can persist for extended periods, remaining viable and potentially re-emerging during subsequent flood events or even leaching into groundwater reservoirs (Casonovas et al., 2018; Nieder et al., 2018). This increases the potential risk to public health- causing infections when in contact with pathogenic bacteria.

Bacterial risk in flood situations is usually quantified using traditional methods- culture-based methods, and detecting faecal indicator bacteria (FIB) (Gronewold et al., 2008) . Culture-based methods offer simplicity and versatility, allowing for the quantification and isolation of bacterial colonies, at a low financial cost. However, these methods underestimate bacterial concentrations due to factors such as uneven distribution of bacteria on agar plates and the presence of viable but non-culturable (VBNC) bacteria (Ferguson et al., 2012). Such methods usually use selective media to identify a selected number of FIB species- such *E. coli* and *enterococci*, which are indicative of faecal contamination and serve as proxies for pathogen presence (Holcomb and Stewart, 2020). FIB may not always correlate with the presence of specific pathogens, and so these methods may be useful

to identify contamination but do not present the full picture of the risk the bacterial community may pose (Saingam et al., 2020). To overcome these limitations, this study uses 16s rRNA sequencing allowing for the direct analysis of all bacterial DNA present in a sample. Sequencing has many advantages, including high sensitivity - particularly valuable when studying complex environments such as soil and floodwater where many bacterial species coexist at varying abundance levels (Price and Wildeboer, 2017). The use of this molecular method allows a more comprehensive picture of the bacterial diversity present in the samples to be created, without having to account for growth requirements of each species that may potentially be present.

This risk posed by bacteria present in flood water, has been studied previously, but usually focusing on the long term bacterial effects- studying flooded areas, for months, or only studying 'one-off' extreme flood events (Casteel et al., 2006; Castro-Ibáñez et al., 2015; Strauch et al., 2005) . Long term studies are useful for examining trends over extended periods, helping any seasonal patterns or drivers to be identified. However, short term studies on microbial dynamics are needed to complement this existing data. Firstly, short term studies allow for the immediate effects to be captured, including the rapid changes in dynamic microbial populations, source of pollution, levels of risk, and exchange between soil and water, and environmental conditions in the direct aftermath of an urban flood event. A short-term study is also useful in immediately being able to identify risk factors such as faecal contamination. If the timescale of the risk is determined , authorities could use this data to advise and create guidelines to safeguard public health in the area. Finally, short term data could - be used to validate existing long term data, adding to databases and models, improving the ability to forecast future flood risks and inform better flood management strategies.

To build a better understanding of the short term bacterial risk at urban flood sites, this study will focus on the specific aspects of : identifying the bacterial communities present in soil and water in the 3 days following an urban flood event, and determining movement of bacteria across the soil-water interface. This study will also look at how potential contamination of floodwater impacts the bacterial community structure of receiving soils, and whether traditional culture-based methods (Heterotrophic plate counts and faecal indicator bacteria) are most appropriate for assessing contamination at an urban flood site.

## 5.3 Materials and Methods

### 5.3.1 Experimental rationale and overview

In order to investigate the short-term dynamics of bacterial communities structure present in urban floodwater sites ,a field study was carried out, visiting sites in the days immediately after an urban flood event, and taking samples every 12 hours. This allowed shifts in the community structure to be tracked over a period of 3 days.

Short term flood studies have been performed previously, however these tend to look at floodwater only, often sampling at a singular time point to determine the bacterial community captured at one moment, rather than looking at the temporal changes over the soil-water interface (Addison-Atkinson et al., 2022; Kapoor et al., 2018 ). As well as this, studies typically tend to focus on the analysis of a small number of pathogenic species-usually those used as faecal indicators, such as *E.coli*, and coliform species (Edberg et al., 2000).

Collecting data on the bacteria present in soil and water in the immediate onset and short term aftermath of an urban flood event will not only provide an insight into bacterial dynamics but will

also give information on how the bacterial community changes within and between the soil and water interface. This will help to identify the most 'critical' time following a flood- on which day, if any, do bacteria from urban floodwater pose a potential risk.

5.3.2 Sample site selection

The same two sample sites; Tongue Gutter and Endcliffe Park, were selected as outlined in Chapter 4.

5.3.3. Sampling regime

Sampling for this study was opportunistic, with visits planned when consistent rainfall had been observed in the area). Samples were then only able to be taken if a 'flood' had actually formed and surface water was visible – increased rainfall did not always mean surface water was present to be sampled. With this in mind, samples were taken from Site 1 - Tongue Gutter on two occasions, once in September 2022, and once in October 2022 (Table 9). At Site 2 -Endcliffe Park samples were only able to be taken on one occasion in June 2023 (Table 9).

**Table 9**. Site visits in relation to cumulative rainfall amount

| **Site Visited** | **Month Visited** | **Cumulative Rainfall Amount 5 days before day 1 of 3 day visit/mm** |
|---|---|---|
| Tongue Gutter | September 2022 | 41.2 |
| | October 2022 | 56.2 |
| Endcliffe Park | June 2023 | 30.8 |

On each occasion the site was sampled for three consecutive days. On each visit, triplicate soil samples were taken, from 0-10cm depth, at both sites using a basic handheld soil auger. This sample depth was selected, as this study was focusing on the short term dynamics of the bacteria in the 3 days following a flood, and thus it would be expected any transferred bacteria would be present towards the soil surface, rather than deeper into the soil profile (Eilers *et al.,* 2010; Petersen *et al.,* 2012). As well as this, the soil surface was of more interest in this study as members of the public are more likely to encounter the soil's surface and potential bacterial pathogens, rather than the deeper portions of the soil. This is important to consider especially given the overall objective of whether a hazard is posed to public health in the short term following an urban flood. Soil samples were taken in the same approximate 1m x1m area at each flood site in a section observed to be affected by flood water regularly. Control (background) soil samples were also taken located close by to the sample site but in an area not known to flood as outlined in Chapter 4. Samples were placed in Ziploc plastic bags and sealed to prevent drying-out. The samples were stored in the dark and transported within 4 hours of collection to the laboratory and stored at 4 ºC (again in the dark) until analysis was undertaken. This minimised any potential changes in the microbial communities in the soil samples. For the molecular genomic work, 50 g sub-samples of soil were sealed in small plastic bags and were preserved at –20 ºC for DNA analysis.

Flood water samples were also obtained when possible, at each visit. For Tongue Gutter, samples of flood water were taken at every visit. This was due to the fact the surface water outflow pipe was always spilling, regardless of rainfall amount (Figure 36).


**Figure 36**. Sample location at Tongue Gutter. A= Visit in September 2022. B=visit in October 2022.

Whereas, at Endcliffe Park, water samples were only able to be taken on one visit, after rainfall of >30mm (cumulative amount 5 days before sample visit) (Table 1). This was due to the origin of the flood at Endcliffe Park being a gully- this only overflowed, creating a surface flood when rainfall was higher (Figure 37) (see Chapter 4 for image of sampling location) .


**Figure 37**. Sample location at Endcliffe Park . A= A visit in August 2022 when surface floodwater was not sufficient for samples to be taken rainfall was <20mm 5 days prior to visit, . B= a visit in June

2023 when surface floodwater samples could be taken, and rainfall amount was >30mm 5 days prior to visit.

5.3.4. Rainfall Monitoring

To understand the influence of precipitation on bacterial dynamics, rainfall was monitored using Detectronics' remote GPRS rain gauge as described in Chapter 4. Daily rainfall for each sample period can be seen plotted on Figures 38, 39, and 40. In these figures, bars indicate the level of precipitation several days before the sampling took place, and black bars are showing sampling days. Both sites were flooded when the daily rainfall depth 2-3 days before sampling was above 10-15 mm.

Only the declining phase of the flood cycle was monitored (sampling 3 days after the flood event), this was due to the main aim of the study focusing more on the potential bacterial hazards posed by floodwater in the time the flood has already formed, and also after a flood. Therefore the arrival of the flood was not as important. As well as this, from a practical point of view, it was more difficult to plan sampling and monitoring during the rising limb of the flood cycle, this would have required more intense monitoring of both sites to witness the flood formation- something which would have been difficult to predict based on weather forecasting and rainfall monitoring alone. It would also have been difficult to take a large enough volume of floodwater to perform adequate analyses, during the rising limb of the flood cycle.

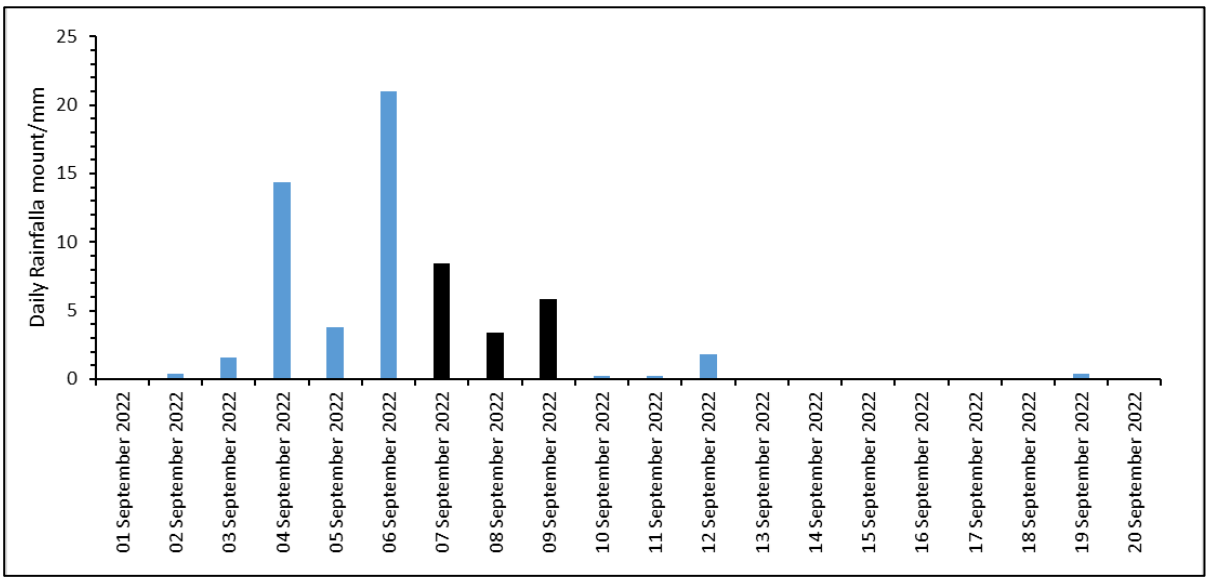


**Figure 38.** Daily rainfall depth over September 2022. Black bars indicate days of site visit to Tongue Gutter- 7th, 8th, 9th September.

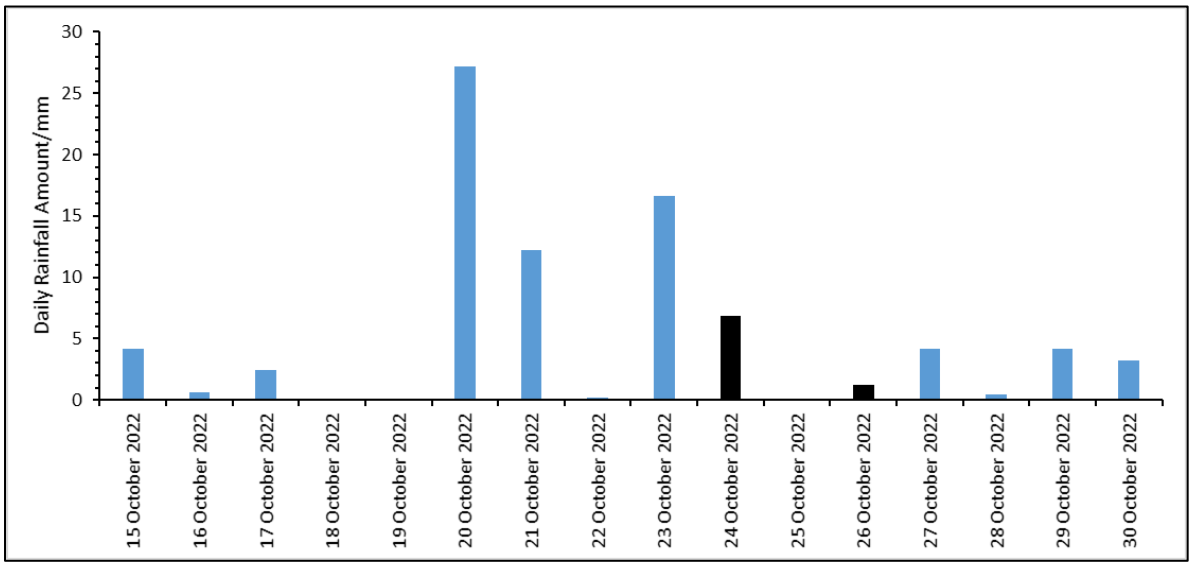


**Figure 39.** Daily rainfall depth over October 2022. Black bars indicate days of site visit to Tongue Gutter- 24th, 25th, 26th October.

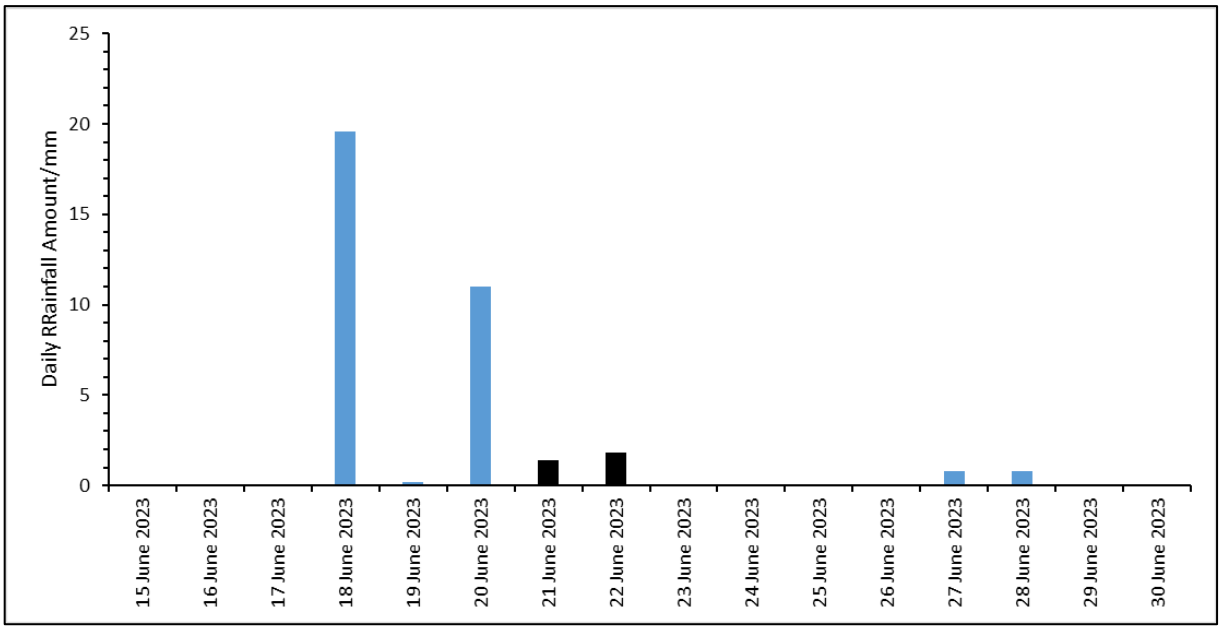


**Figure 40.** Daily rainfall depth over June 2023. Black bars indicate days of site visit to Endcliffe Park- 21st, 22nd, 23rd June.

5.3.5. Sample analysis

5.3.6 Physico-chemical factors, CFU counts, and soil classification

Soil and water pH , water temperature, soil moisture content, soil organic content and soil particle analysis were all carried out as outlined in Chapter 4. Coliform CFU counts, using a selective media to grow only *E. coli* and *Citrobacter freundii* were also carried out as described in Chapter 4.

5.3.7 Molecular genomic methods

The use of these molecular techniques allows the characterisation of those components of the soil - microbial community that currently- that cannot be cultured (Hosokawa et al., 2022). DNA extraction, quantification and Illumina sequencing of the 16s rRNA gene: including a test PCR and gel electrophoresis, was carried out as described in Chapter 4.

*5.3.7.1 Bioinformatics: DNA sequences analysis*

Bioinformatic analysis was carried out as outlined in Chapter 4, following the steps shown in figure 41, with the only exception of sampling depth. For this study the sampling depth was set at minimum - 2000 and maximum- 680000 reads. The sampling depth is different to that of the study in Chapter 4, due to these samples being sequenced in a different 'batch' and thus having different rarefaction curves, of which the sampling depth is based.

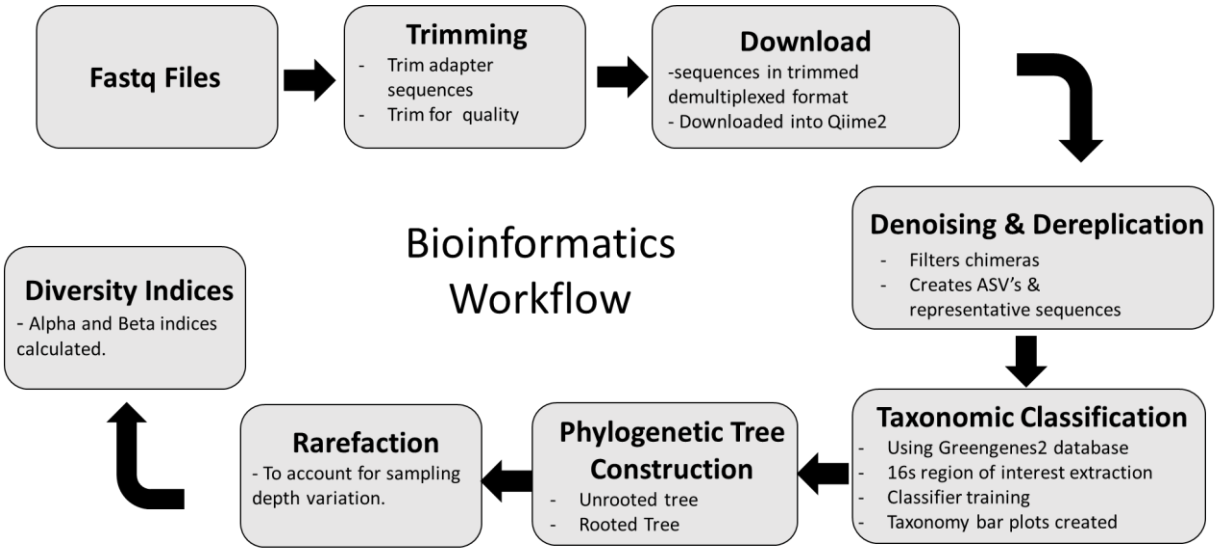


**Figure 41**. Pipeline of workflow used for bioinformatics analysis.

5.3.8 Diversity Indices

To fully investigate the bacterial diversity within the samples both alpha and beta diversity metrics were calculated in Qiime2, as described in Chapter 4.

5.3.9 Statistical Analysis

In order to determine differences and correlations between physico-chemical factors, coliform counts and abundance of species, statistical analysis, including Pearson's correlation coefficient, independent t-test, Principal Coordinate Analysis and PERMANOVA analysis , was carried out as described in Chapter 4.

**5.4 Results**

This section describes in detail the results of analysis of physico-chemical factors, soil analysis, taxonomic analysis, and diversity indices measured. Any differences or correlations between factors are outlined. The section is split by sample site (with the exception of soil characterisation data).

5.4.1 Soil characterisation: Soil particle size and class/type

Soil was sized and classified according to British Standard 1377 Part 2 (British Standards Institute, 1990). Results including a soil gradation curve are shown in Chapter 4. In short it was found that the soil at 'Tongue Gutter' was predominantly coarse sand, with fine gravel, and the soil at 'Endcliffe Park' was similar, also being classified as a coarse sand with fine gravel, but with more particles present of a slightly smaller size than that of Tongue Gutter.

### 5.4.2 Tongue Gutter

Soil and water samples were collected during two, three day visits to this site. The first visit was in September 2022. Although monitored soil, floodwater and background soil samples (soil samples taken in an area at the site not affected by surface flooding) were taken on each day, not all samples - once DNA was extracted - were of high enough quality or of high enough concentration for subsequent sequencing (according to the thresholds set by the NEOF Liverpool sequencing lab). As such, physicochemical data, and coliform counts data is available for all samples, however taxonomic data is only available for background soil samples from Day 3 of the September visit, monitored soil from Days 2 and 3 and water samples from Days 1 and 3 (Table 10).

For the second visit in October 2022, monitored soil and water samples for all 3 days were analysed and also sequenced (bacterial taxonomic data available for all samples on all 3 days). Taxonomic data is available for background soil samples only on days 2 and 3 (Table 10).

**Table 10**. Samples taken from Tongue Gutter over September and October 2022 in relation to whether taxonomic data is available for each sample.

| Month of Visit | Sample Day | Sample Type | Taxonomic data availability |
|---|---|---|---|
| Sep | Day 1 | Background Soil | No |
| | | Monitored Soil | No |
| | | Water | Yes |
| | Day 2 | Background Soil | No |
| | | Monitored Soil | Yes |
| | | Water | No |
| | Day 3 | Background Soil | Yes |
| | | Monitored Soil | Yes |
| | | Water | Yes |
| | Day 1 | Background Soil | No |

| | | | |
|---|---|---|---|
| Oct | | Monitored Soil | Yes |
| | | Water | Yes |
| | Day 2 | Background Soil | Yes |
| | | Monitored Soil | Yes |
| | | Water | Yes |
| | Day 3 | Background Soil | Yes |
| | | Monitored Soil | Yes |
| | | Water | Yes |

5.4.2.1 Physicochemical analysis and estimation of bacterial abundance of monitored soil, background soil, and floodwater

Table 11 includes physicochemical data obtained from soil and water samples, along with coliform counts for all 3 sample days in September and October 2022.

Firstly, the September visit will be discussed. For soil samples, the pH range of monitored soil was significantly lower than that of background soil (t-value = 4.52., p=0.005.). The highest pH value was observed on day 3 for both background and monitored soil. For background soil, the soil moisture content increased over the sampling days, however monitored soil had the highest soil moisture content on day 2 at 39.7. Organic content of background soil was variable but was highest on day 3 at 13.35%. The organic content of monitored soil was similarly variable, with the highest value of 10.77% on day 2. Organic content was higher in background soil than in monitored soil on days 1 and 3 only.

Coliform counts for background soil ranged from 10000 to 20000 CFUs per gram of soil. Coliform counts in monitored soil decreased over the sampling days, but a significantly higher count was seen for monitored soil compared to background soil ( t-value= 5.95. p= 0.0003).

Water pH was significantly lower than monitored soil pH (t-value = -9.67. p = 0.0003) and decreased over the sampling days. Water temperature did not vary over the sampling days. Coliform counts increased over the sampling days, increasing from 410000 on day 1 to 700000 CFUs/ml on day 3. This is opposite to the trend seen in monitored soil- a decrease in coliform counts over the sampling days.

**Table 11**.Results of physiochemical analysis, Coliform counts, and rainfall amount (n=18).

| Month of Visit | Cumulative rainfall amount 5 days before first sample visit/mm | Sample Day | Sample Type | pH | Soil Moisture/ % | Temperature/ ℃ | Organic Content of soil/% | Coliforms* /per gram of soil or ml of water |
|---|---|---|---|---|---|---|---|---|
| Sep | 41.2 | Day 1 | Background Soil | 8.24 | 20.65 | N/A | 11.58 | 20000 |
| | | | Monitored Soil | 7.88 | 28.12 | N/A | 6.59 | 780000 |
| | | | Water | 7.37 | N/A | 10 | N/A | 410000 |
| | | Day 2 | Background Soil | 8.20 | 18.58 | N/A | 8.72 | 10000 |
| | | | Monitored Soil | 7.88 | 39.70 | N/A | 10.77 | 610000 |
| | | | Water | 7.35 | N/A | 10 | N/A | 580000 |
| | | Day 3 | Background Soil | 8.50 | 38.87 | N/A | 13.35 | 10000 |
| | | | Monitored Soil | 7.90 | 32.33 | N/A | 8.72 | 280000 |
| | | | Water | 7.18 | N/A | 10 | N/A | 700000 |
| Oct | 56.2 | Day 1 | Background Soil | 7.38 | 30.36 | N/A | 14.20 | 0 |
| | | | Monitored Soil | 7.09 | 26.97 | N/A | 23.55 | 50000 |
| | | | Water | 7.64 | N/A | 10 | N/A | 250000 |
| | | Day 2 | Background Soil | 8.01 | 31.64 | N/A | 14.58 | 0 |
| | | | Monitored Soil | 7.30 | 40.92 | N/A | 8.16 | 780000 |
| | | | Water | 7.31 | N/A | 10 | N/A | 500000 |
| | | Day 3 | Background Soil | 7.77 | 31.19 | N/A | 20.65 | 0 |
| | | | Monitored Soil | 7.18 | 30.12 | N/A | 6.85 | N/A |
| | | | Water | 7.47 | N/A | 10 | N/A | 600000 |

Coliforms*= *E. coli* and *C. freundii* colonies

Next, the October visit is discussed. The pH of monitored soil was variable- highest pH was observed on day 2, with the lowest on day 1 (Table 11). Monitored soil pH was significantly lower than background soil pH (t-value = -2.74. p-value =0.026). Moisture content in background soil samples ranged from 30.12-31.64%, whereas monitored soil had a larger range of 26.97% - 40.92%. Organic content of soil was variable for both background and monitored soil with no clear trend, and no significant difference in the values between the two data sets.
No coliforms were detected in background soil samples, but coliforms were detected in monitored soil on every day of sampling. The highest number of coliform counts was found on day 2 (780000), with the lowest on day 1 (50000). pH of water was highest on day 1 (7.64) and lowest on day 2 (7.31). Water temperature did not vary over the sampling days. Coliform count in water was highest on day 3 (600000) and lowest on day 1 (250000). Coliform count in water did not correlate with any physico-chemical factors measured.

5.4.2.2 Bacterial Community Structure

This section describes the analysis of the bacterial community present in each sample at phylum, genus, and species level.

*Phylum*

Figure 42, shows the percentage of relative abundance at phylum level of background soil, monitored soil and water samples from both site visits- September and October -based on assigned ASV's.

For the September visit, taxonomic data is available for background soil on Day 3 only.. All 26 phyla detected in background soil were also found in monitored soil. In addition to the 26 phyla monitored soil shares with background soil, an additional 10 unique phyla were also present . On both sample days, the most abundant phylum in monitored soil does not change and remains as Proteobacteria at >20% each day.

Much like the monitored soil samples, Proteobacteria was the most abundant phyla in water over both sample days, >35%. Although at a higher abundance in water than that of monitored soil (~10-20% higher). Proteobacteria relative abundance in both monitored soil and water samples decrease over the sample days. All phyla found in water were also present in monitored soil, apart from Riflebacteria, found only in water on day 1 at <0.1%.

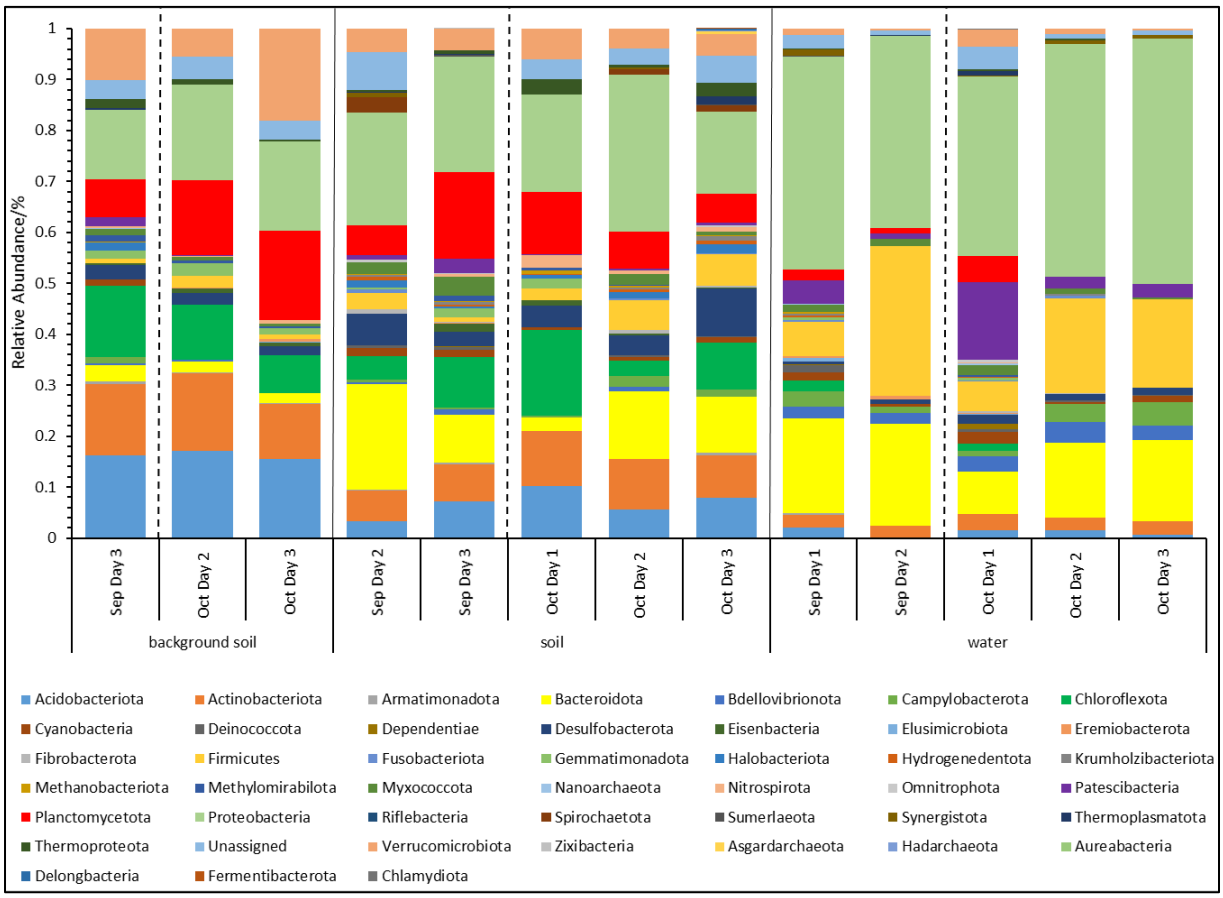


**Figure 42**. Comparison of relative abundance of bacterial communities present in monitored soil ,background soil , and water samples at phylum level from both site visits (n=13). Unassigned taxa were grouped into 'Unassigned' category. Y- axis scale 0-1 represents 1-100% relative abundance.

For the October visit, the number of phyla increased over the days in monitored soil from 20 phyla on day 1 of sampling, to 36 phyla on day 3. On all 3 sample days, Proteobacteria is the most abundant phylum at 15%-30%. 18 phyla are present on all 3 sample days in monitored soil, including Acidobacteria, Bacteroidota, and Campylobacterota among others. All phyla detected in background soil were found in monitored soil, however 19 phyla in monitored soil are unique to the monitored soil samples, and thus are not found in background soil samples. These phyla unique to monitored soil include phyla such as: Methanobacteriota, Halobacteriota, and Aureabacteria.

For water samples, the number of phyla decreased over the days from 35 on day 1 to 14 phyla on day 3. Whereas in soil the number of phyla increased. Like with monitored soil, the most abundant phylum in water over time was Proteobacteria. On all 3 days, the relative abundance of Proteobacteria was much higher (~15-32% higher) in water than in monitored soil. 14 phyla were present in water on all 3 days of sampling, including phyla such as: Firmicutes, Deinococcus, and Bdellovibrionota. Water and monitored soil shared all but 4 phyla. The 4 phyla unique to water were: Chlamydiota, Dependentiae, Elusimicrobiota, Zixibacteria. All of which appear on day 1 only, at very low relative abundance (0.01-1%).

*Genus*

At genus level, sequences that were not assigned to a genus were removed, along with any sequences that had a relative abundance of <0.5% on all days sampled to simplify the presentation of results.

Firstly, looking at the data from the September visit (Figure 43) on day 2, 45 genera were present in monitored soil representing 26.5% of the actual bacterial community. The number of genera decreased by over half on day 3 and included *Flavobacterium, Clostridium*, and *Cellvibrio.* . In total only 14 genera were present in soil for both sampling days. Of these genera, only 3 increased in relative abundance over the 2 days. These were *Nakamurella, Ohtaekwangia*, and *Flavobacterium*. However, the increase was relatively small, between 0.01 and 0.06%.

For background soil from day 3, only 22 genera were present. Only 5 genera detected in background soil were also present in monitored soil, including *Mycobacterium, Nakamurella, Desulfobacca, Hyphomicrobium*, and *Terrimicrobium.*

On day 1 46 genera were present in water, with the number of genera increasing to 50 on day 2 representing almost three quarters of the total bacterial community- 71.5%. On day 1, the most abundant genus was *Flavobacterium* at almost 5%. On day 2, the most abundant genus was *Phocaeicola,* increasing from 1.2% on day 1 to >3% on day 2. On day 2, 23 new genera - not found on day 1- appeared. Monitored soil and water shared 10 genera including *Pseudomonas, Cellvibrio,* and *Azonexus*, amongst others. Of the genera found in water only, many faecal associated genera were present, including: *Phocaeicola, Bifidobacterium, Roseburi*a, and *Faecalibacterium*- all of these genera increased in water over time.

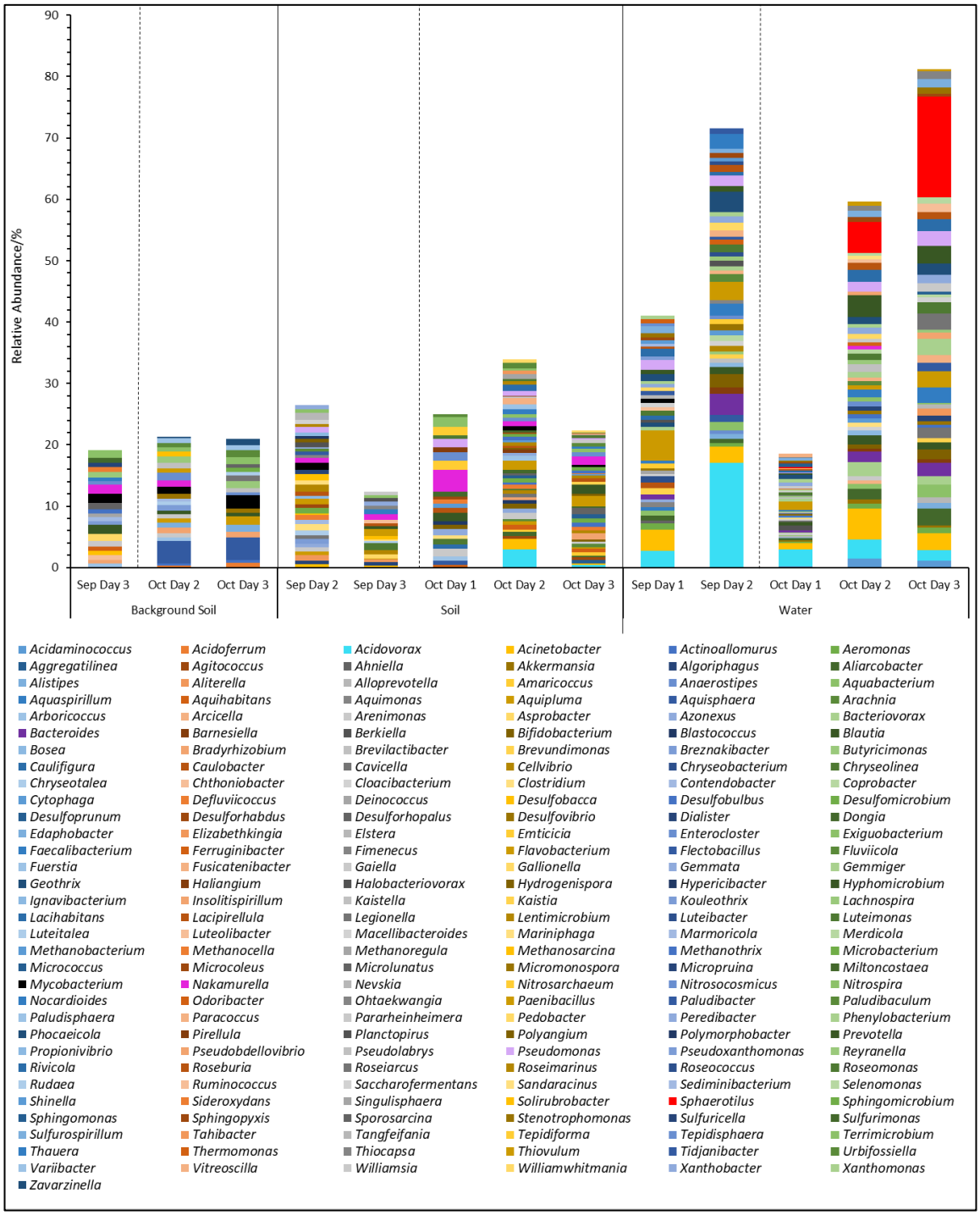


**Figure 43.** Comparison of bacterial communities present in monitored soil, background soil , and water samples at Genus level (n=13). Filtered to include only genera with a relative abundance of >0.5% on at least one of the sample days.

During the October visit, up to 33% of the total bacterial community was represented by monitored soil samples. On day 1, 25 genera were present in monitored soil, (Figure 43). On day 2, the number of genera in monitored soil increased by over double to 52, including the appearance of 35 new genera such as: *Acinetobacter*, *Paracoccus* and *Flavobacterium* at >1%. On day 3, the number of genera decreased by 10, to 42. Only 9 genera were present in monitored soil over all 3 days. Of these genera, 8 , including; *Methanobacterium, Arenimonas*, and *Pseudomonas* had the highest

relative abundance in soil on day 1, at 0.7%- 1.5% respectively. 12 genera present in monitored soil were also present in background soil.

In water samples, a much higher number of different genera was detected on all three days, compared to monitored soil samples. 51 genera were detected in water on day 1 , but only representing 18% of the total bacterial community. This indicates the community was made up of lots of low abundance genera.

On day 3, an even higher proportion of the bacterial community was represented (81.25%), but with the lowest number of genera present at 49. This indicates the community in the water on day 3 was made up of a lower number of unique genera that had a higher relative abundance than those on days 1 and 2. 5 genera appeared on day 3 only: *Micrococcus*, *Flectobacillus*, *Elizabethkingia*, *Brevundimonas*, and *Insolitispirillum* at 0.5-1.3%. 33 genera were present in water on all 3 sample days, this includes many genera associated with faeces and/or pathogenic species . These genera included: *Legionella, Faecalibacterium, Clostridium, Akkermanisa*, *Bifidobacterium,* and *Roseburia*. 12 genera found in water were also found in monitored soil, including: *Pseudomonas, Flavobacterium, Bifidobacterium, and Clostridium.* These genera were present on all 3 sample days in water, however abundance levels and presence in soil were variable, fluctuating for each genus over the sample days.

*Species*

At species level, all samples were filtered to include species with a relative abundance of >0.5% on at least one sample day. This was due to the high number of low abundance species present in the samples. To accurately describe and analyse the plethora of species present background soil and monitored soil data will be presented first in Figure 44, with water data separate in figure 45.

Looking firstly at the September visit, for monitored soil, on day 2, 20 species were present- representing >13% of the total community (Figure 44). The number of species decreased to just 9 on day 3. . Only 4 species were present on both sampling days. Of these 4 species, only one increased over time - *Peredibacter starrii,* the other three: *Nakamurella flavida, Chryseotalea sanaruensis*, and *Terrimicrobium sacchariphilum*- decreased. *N. flavida* and *T. sacchariphilum*, were the only species in monitored soil that were also found in background soil. Both species have higher relative abundance in background soil than in monitored soil.

At the October visit, for monitored soil, 35 different species were detected over the 3 sampling days- representing up to 16% of the total bacterial community present (Figure 44). This means that only a small proportion of the community was represented after unassigned species were removed and filtered . On day 1, 16 species were present. On day 2, the number of genera present in monitored soil increased to 28. On day 3, the number of genera in monitored soil decreased to 21 species, however, 2 new species appeared: *Paludibaculum fermentans* and *Aggregatilinea lenta*.

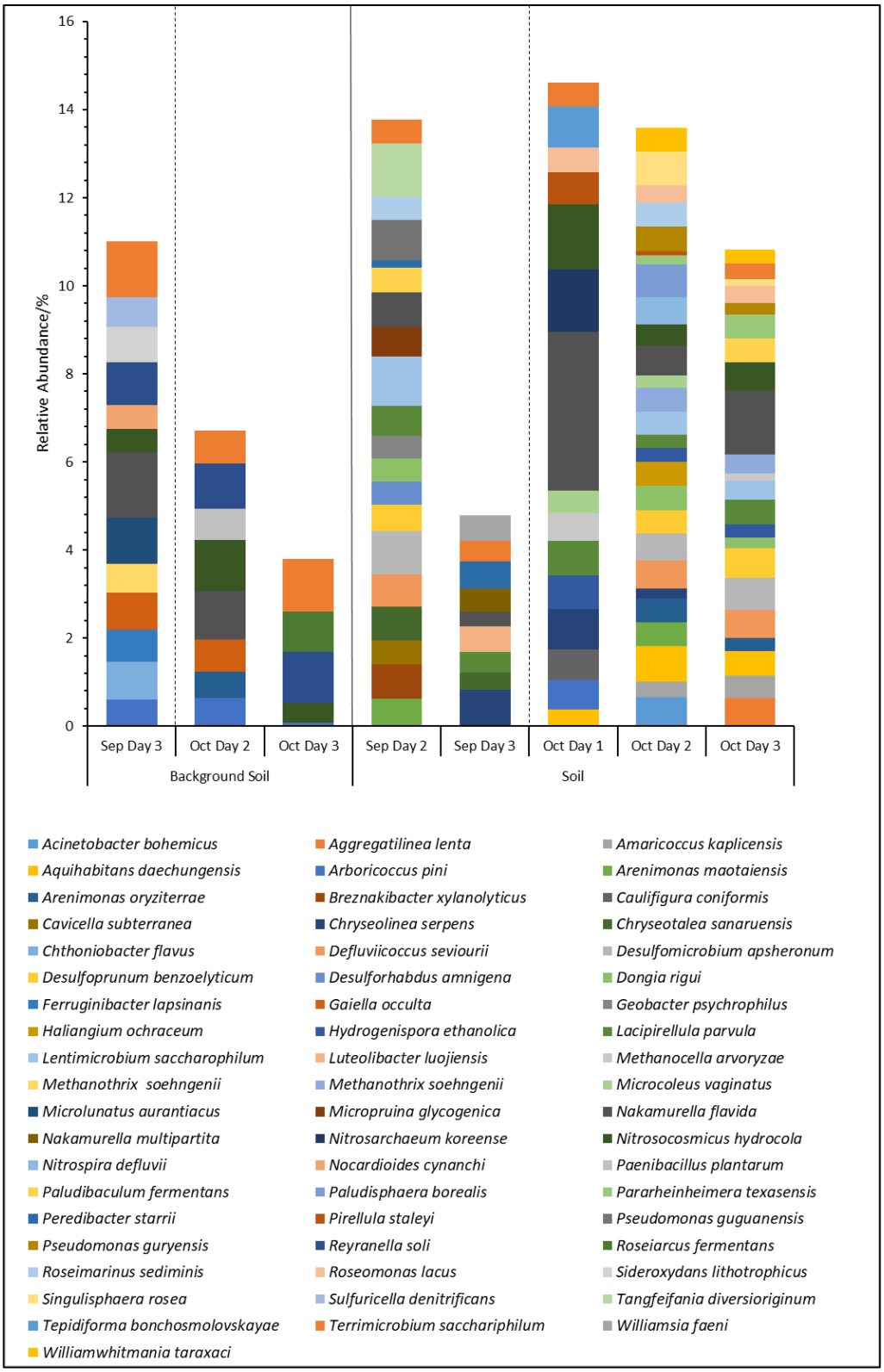


**Figure 44**. Relative abundance of species present in monitored soil and background soil samples at both visits (n= 8) . Filtered to include only species with a relative abundance of >0.5%, on at least one of the sample days.

Regarding the water samples for the September visit. On day 1, >10% of the total community was represented, with 17 species (Figure 45). This increased to 22 species on day 2. On both sampling days *Phocaeicola vulgatus* was the most abundant species, at 0.9% relative abundance on day 1, and increasing to >3% on day 2. This leaves only 6 species present in water over both sampling days. All of those species, which includes *Phocaeicola vulgatus*, *Faecalibacterium prausnitzii* and *Prevotella copri,* increased over time. Only two species found in water were also found in monitored soil: *Peredibacter starrii* and *Pseudomonas guguanensis. P. starii* increases in both water and monitored soil over the time.

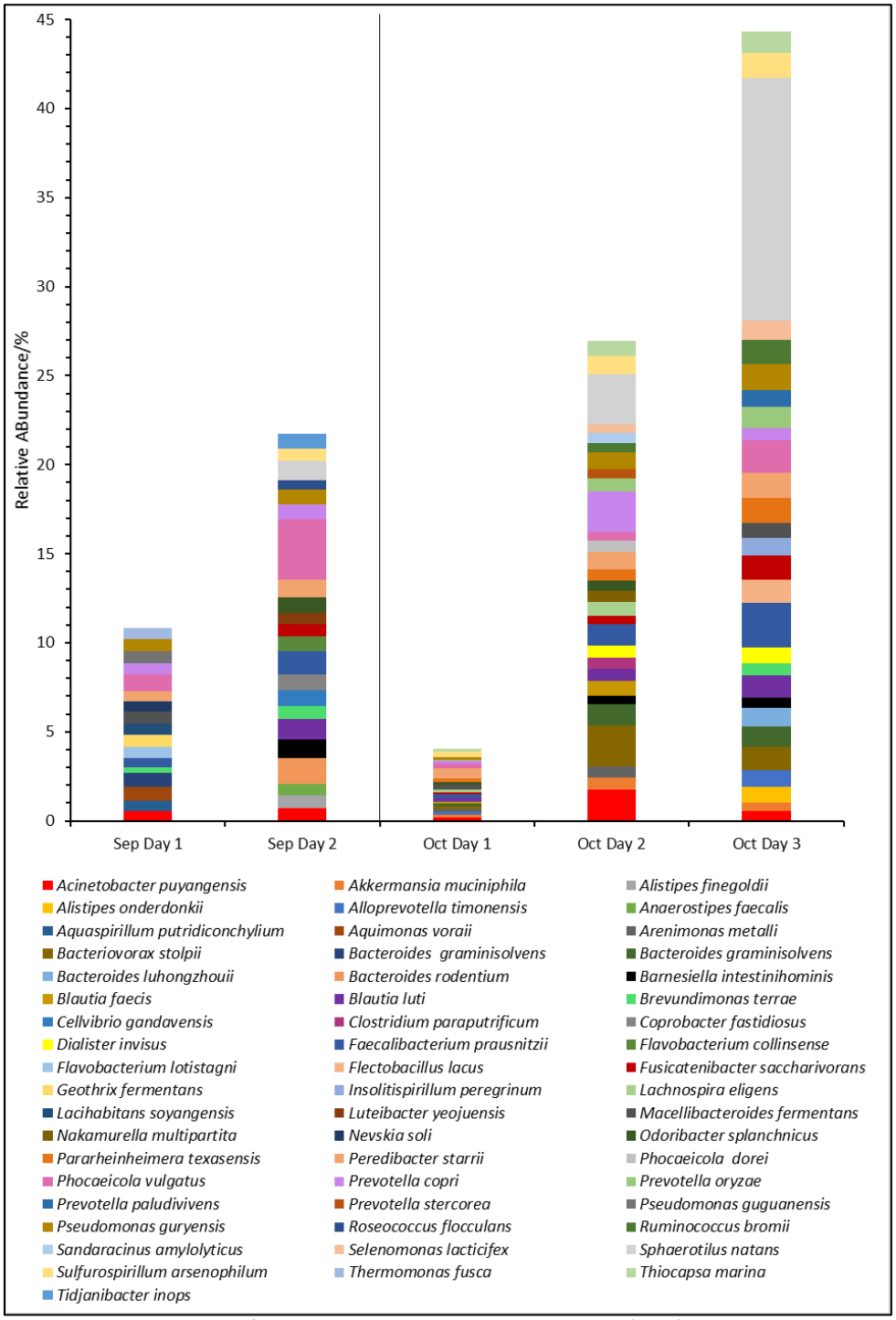


**Figure 45**. Relative abundance of species present in water samples (n=5). Filtered to include only species with a relative abundance of >0.5%, on at least one of the sample days.

For the water samples from the October visit, on day 1, 22 species were detected in water samples, representing just < 5% of the total bacterial community (Figure 45). Once again, this percentage is very low, and indicates that the rest of the community was composed of rare species. The most abundant species in water on day 1 was *Phocaeicola vulgatus* at just 0.27%. On day 2, the number of species in water increased to 28. On day 3, 28 species were present in water, now representing >44% of the total bacterial community- a 40% increase from day 1. On day 3 6 new species not found in water on the two previous days appeared. These species were: *Bacteroides luhongzhouii, Prevotella paludivivens, Alistipes onderdonkii, Flectobacillus lacus, Brevundimonas terrae*, and *Insolitispirillum peregrinum*- present at a relative abundance between 0.5-1.3%.

15 species were present in water over all 3 sampling days, these included : *Phocaeicola vulgatus,* and *Prevotella copri*, which increased in abundance over time. Only two species found in water were also found in monitored soil: *Pseudomonas guryensis* and *Pararheinheimera texasensis.* Both species increase in water over the sampling days. Although present continuously in water, these two species only appeared in monitored soil on days 2 and 3 , at lower abundance than in water.

**5.4.2.3 Diversity Indices**

Alpha Diversity

To determine alpha diversity, 3 unique measures were calculated at species level using all ASV's (Table 12). For the September visit in monitored soil, ACE Species richness increased over time from 237 to 372, and Faith's Diversity index increased from 31.5 to 45.24. However, Simpson's Evenness index decreased from 0.78 to 0.73. When comparing background and monitored soil (day 3 only), all alpha diversity indices were higher in monitored soil than background soil. ACE species richness was higher in monitored soil by 121, Evenness was higher by 0.08, and Faith's Diversity was higher by 12.39. For water samples, all alpha diversity indices were higher on day 1, than day 2. When comparing water to monitored soil on day 2, all diversity indices were higher in monitored soil than in water (ACE, higher by 111, Simpson's higher by 0.19, and Faith's higher by 11.04).

**Table 12.** Alpha diversity indices for background soil, monitored soil, and water samples (n=13) taken at Tongue Gutter during September and October 2022.

| Month of visit | Visit Day | Sample Type | ACE Species Richness | Simpson's Evenness Index | Faith's Diversity Index |
|---|---|---|---|---|---|
| Sep | Day 1 | Water | 268 | 0.68 | 34.64 |
| | Day 2 | Water | 126 | 0.59 | 20.46 |
| | | Monitored Soil | 237 | 0.78 | 31.50 |
| | Day 3 | Monitored Soil | 372 | 0.73 | 45.24 |
| | | Background Soil | 251 | 0.65 | 32.85 |
| | Day 1 | Monitored Soil | 115 | 0.70 | 19.94 |

| | | | | | |
|---|---|---|---|---|---|
| Oct | | Water | 766 | 0.68 | 75.90 |
| | Day 2 | Monitored Soil | 276 | 0.77 | 32.61 |
| | | Background Soil | 329 | 0.67 | 40.68 |
| | | Water | 143 | 0.72 | 20.80 |
| | Day 3 | Monitored Soil | 181 | 0.74 | 23.45 |
| | | Background Soil | 148 | 0.79 | 20.34 |
| | | Water | 102 | 0.70 | 15.86 |

For the October visit, monitored and background soil samples, both had highest Simpson's Evenness on day 2, at 0.77 and 0.67 respectively. On day 2, monitored soil also had the highest Faith's diversity index (32.61) and ACE species richness (276). On average, monitored soil had higher evenness and faith's diversity values compared to background soil, however this is not statistically significant. Water had its highest evenness (75.90), and ACE species richness (766) values on day 1, with Faith's diversity being highest on day 2 (0.72). On average, water has higher Faith's diversity and ACE richness values when compared to monitored soil; this is not a statistically significant difference.

Beta diversity

PCoA analysis could not be performed on beta diversity to compare samples from each site visit (September and October individually). This method requires at least 3 data points in each 'group', and unfortunately samples for all 3 days from each group (monitored soil, background soil, water) were not sequenced and thus could not be analysed with the PCoA method. Instead, background and monitored soil data were collated from both visits and presented in figure 46 . This allowed the comparison of beta diversity between monitored soil, background soil, and water from both visits to look at the site as a whole. When looking at the site as a whole, a difference was seen in the beta diversity of monitored, and background soil. However, as can be seen on Figure 50, the data points do not clearly separate according to type of sample . When comparing monitored soil and water at the site, more discrete groups can be seen, with a clear separation between the monitored soil and water data points.

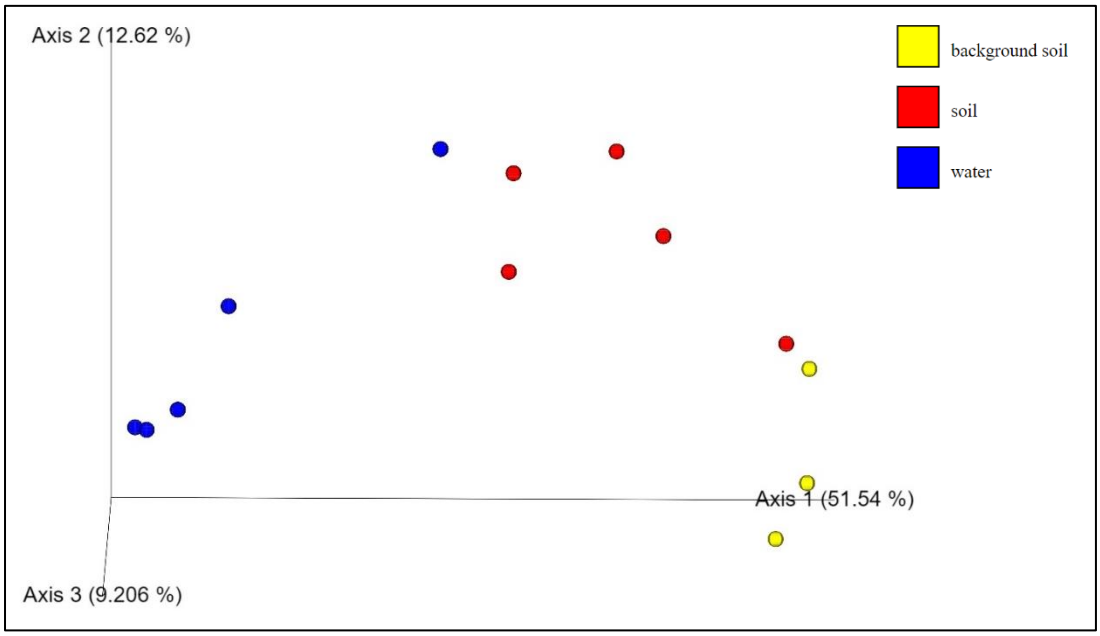


**Figure 46**. PCoA (Principle Coordinate Analysis) Emperor plot based on weighted unifrac distance analysis ( n= 13).

### 5.4.3 Endcliffe Park

Soil and water samples were collected during one, three-day visit to this site in June 2023. Samples were not able to be collected on other visits due to the lack of surface water present, despite rainfall. On unsuccessful visits (no samples taken) it was noted that the gully in which the floodwater usually originated appeared 'blocked' with vegetation (Figure 47). It is thought this may have been the reason why no flood water was present, and so further samples were not able to be taken and thus analysed.

**Figure 47**. Gully blocked with vegetation at Endcliffe Park in which flood water spills from. Picture taken on 17th August 2022.

As mentioned, on the visit to this site over 21st to 23rd June 2023, samples were successfully collected. Monitored soil, floodwater and background soil samples were taken on each day. Physicochemical data, and coliform count data is available for all samples, as well as full, taxonomic data (down to species level).

5.4.3.1 Physicochemical analysis and estimation of bacterial abundance of monitored soil, background soil, and floodwater

Table 13 includes physicochemical data obtained from soil and water samples, along with coliform colony forming unit counts for all 3 sample days in June 2023.

**Table 13**. Results of physiochemical analysis, coliform counts, and rainfall amount (n=9).

| Cumulative Rainfall amount (5 days before day 1 of sampling)/mm | Day | Sample Type | pH | Soil Moisture/% | Temperature /°C | Organic Content of soil/% | Coliforms*/CFUs per gram of soil or ml of water |
|---|---|---|---|---|---|---|---|
| 30.8 | Day 1 | Background Soil | 7.47 | 24.56 | N/A | 11.00 | 0 |
| | | Monitored Soil | 7.34 | 25.48 | N/A | 9.09 | 720000 |
| | | Water | 6.64 | N/A | 16 | N/A | 430000 |
| | Day 2 | Background Soil | 7.80 | 21.91 | N/A | 11.06 | 250000 |
| | | Monitored Soil | 7.92 | 26.01 | N/A | 10.5 | 20000 |
| | | Water | 7.10 | N/A | 16 | N/A | 170000 |
| | Day 3 | Background Soil | 8.03 | 23.75 | N/A | 9.23 | 0 |
| | | Monitored Soil | 7.50 | 31.27 | N/A | 9.40 | 310000 |
| | | Water | 7.30 | N/A | 16 | N/A | 570000 |
| Coliform CFUs= *E. coli* and *C. freundii* colonies | | | | | | | |

Monitored soil pH was lowest on day 1 and highest on day 2, with background soil sharing a similar range (~7.4-8.0). Overall, monitored soil had a higher moisture content over the sampling period (not enough data points per group to test differences statistically) than background soil samples. Monitored soil moisture content increased over time, whereas the moisture content in background soil was variable- highest on day 1 (24.56%), and lowest on day 2 (21.91%). Organic content of monitored soil did not vary much over the sampling period, with a range of 9.09% - 10.5%. Background soil had a higher organic content than monitored soil on the first two sampling days (higher by 1.91% on day 1, and higher by 0.56% on day 2), but this decreased on day 3 to below that of monitored soil (lower by 0.17%). Coliforms were only found in background soil on day 2, but they were present in monitored soil samples on every day of sampling ( > 20000 coliform CFUs per gram of soil). Coliform counts in monitored soil were highest on day 1, and lowest on day 2, with counts decreasing to levels below those found in background soil on this day (lower than background soil by 260000 CFUs).

Water pH was lowest on day 1 at 6.64, and highest on day 2 – 7.10. A difference between monitored soil and water pH was observed, with monitored soil having a higher pH over time. Water temperature over the sample period did not vary – remaining at 16˚C over the 3 days. This is unusual as it was expected temperature would change slightly. Coliform counts in water were highest on day 1 (430000 CFUs per ml,), and lowest on day 2 (170000 CFUs per ml). Coliform counts in water were lower than in monitored soil on day 1 (lower by 290000 CFUs/ml) however the opposite is true on days 2 and 3 (coliform count was higher in water than monitored soil by 150000 CFUs on day 2, and by 260000 CFUs on day 3).

5.4.3.2 Bacterial Community Structure

In this section, results from bacterial community structure obtained from sequencing are shown at different taxonomic levels.

*Phylum*

Figure 48, shows the percentage of relative abundance at phylum level of background soil, monitored soil and water samples, based on assigned ASV's.

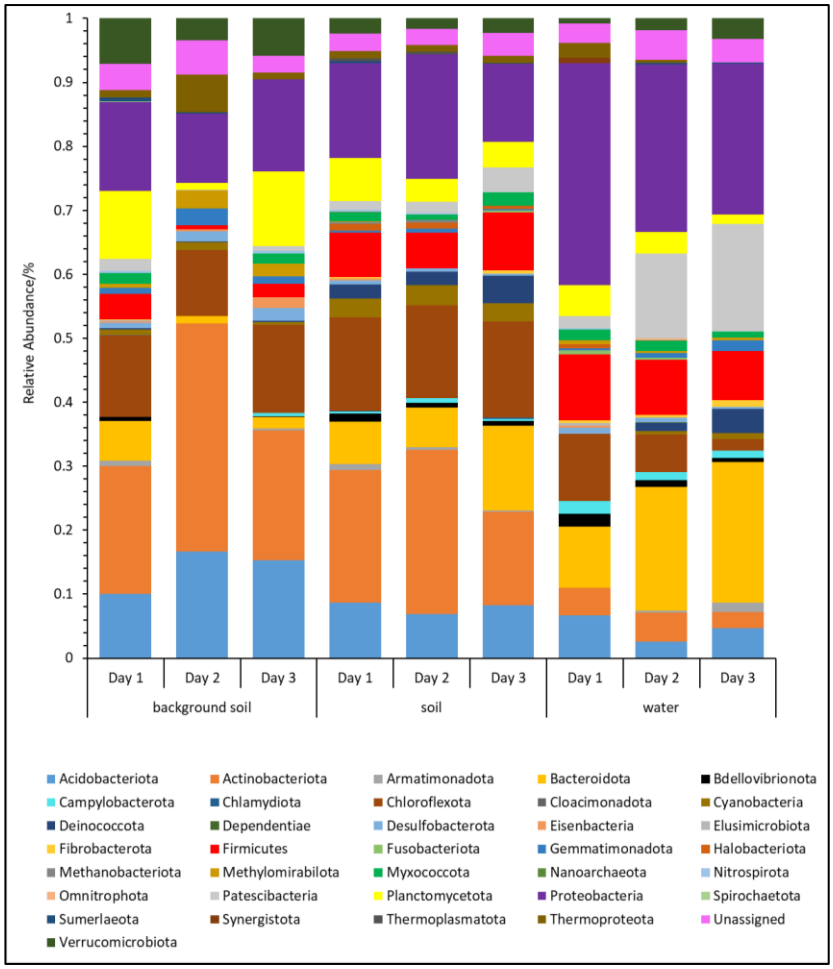


**Figure 48.** Comparison of bacterial communities present in monitored soil , background soil , and water samples at phylum level (n=9). Unassigned taxa grouped into 'Unassigned' category. Y- axis scale 0-1 represents 1-100% relative abundance.

On day 1 of sampling, 28 phyla were present in monitored soil, this decreased over time to just 13 phyla on day 3. On all 3 sampling days Actinobacteria was the most abundant phylum at >20% on days 1 and 2, decreasing to 14.5% on day 3. 26 phyla were found in both background and monitored soil over time. On day 1 of sampling, 24 phyla were present in water. On day 2, the number of phyla increased to 27 phyla in water, with 8 new phyla appearing, including: Deinococcus at >1%, Cyanobacteria and Armatimonadota at <1%. On day 3 the number of phyla decreased to 20, with no new phyla appearing.

Over all 3 sample days, Proteobacteria was the most abundant phyla , decreasing from > 35% on day 1, to >23% on day 3. Proteobacteria also decreased in monitored soil over time, but at much lower levels - 12-19% relative abundance. Soil and water shared 28 phyla, with the exception of : Elusimicrobiota, Omnitrophota, Synergistota,and Dependentiae, which were present in water but not in monitored soil.

*Genera*

At genus level, sequences that were not assigned to a genus were removed, along with any sequences that had a relative abundance of <0.5% on all days sampled.

On day 1, 48 genera were present in monitored soil (Figure 49). On day 2, 41 genera were present, a decrease of 7 when compared to day 1. 4 new genera appeared in monitored soil - *Clostridium* , *Flavobacterium, Desulfobulbus* and *Rubellimicrobium* at 0.5%-1%. On day 3, the number of genera increased by 9 to 50- with the most abundant genus being *Deinoccous* , at its highest relative abundance over the 3 sample days, of >3% (1-1.5% on days 1 and 2).

For water samples, 43 genera were present on day 1. 16 genera were present in water on day 1 only: including *Bifidobacterium ,Streptomyces* and *Bacteroides* at <1%. The most abundant genus in water on day 1 was *Acidovorax* at >9.5%- with *Acidovorax* not being present on any other days. *Acidovorax* represented around 20% of the sequenced bacterial community on day 1.

On day 2, the number of genera present in water increased by 4 to 47, with the appearance of 25 new genera that were not present on day 1. These genera included: *Armatimonas, Paludibacter*, and *Skermanella* amongst others. On day 3, the number of genera decreased by 16 , to a final total of 31. 4 genera were present on day 3 only: *Cellulomonas* , *Aliterella* , *Acetivibrio* ,and *Dendrosporobacter* at 0.5-0.8%. 13 genera were present in water over all 3 sample days. Some of those genera included: *Acinetobacter, Sphingomicrobium*, and *Paracoccus*- all of which had lowest abundance on day 2 and highest on day 3.

Over the sample period, 16 genera were found in both water and monitored soil. Those genera included: *Flavobacterium, Clostridium*, and *Rubellimicrobium*. *Flavobacterium* increased in abundance in soil over time, whereas in water, it decreased in abundance. *Clostridium* decreased over the sample days in both soil and water. *Rubellimicrobium* was highest on day 2 in both soil (0.5%) and water (1.4%).

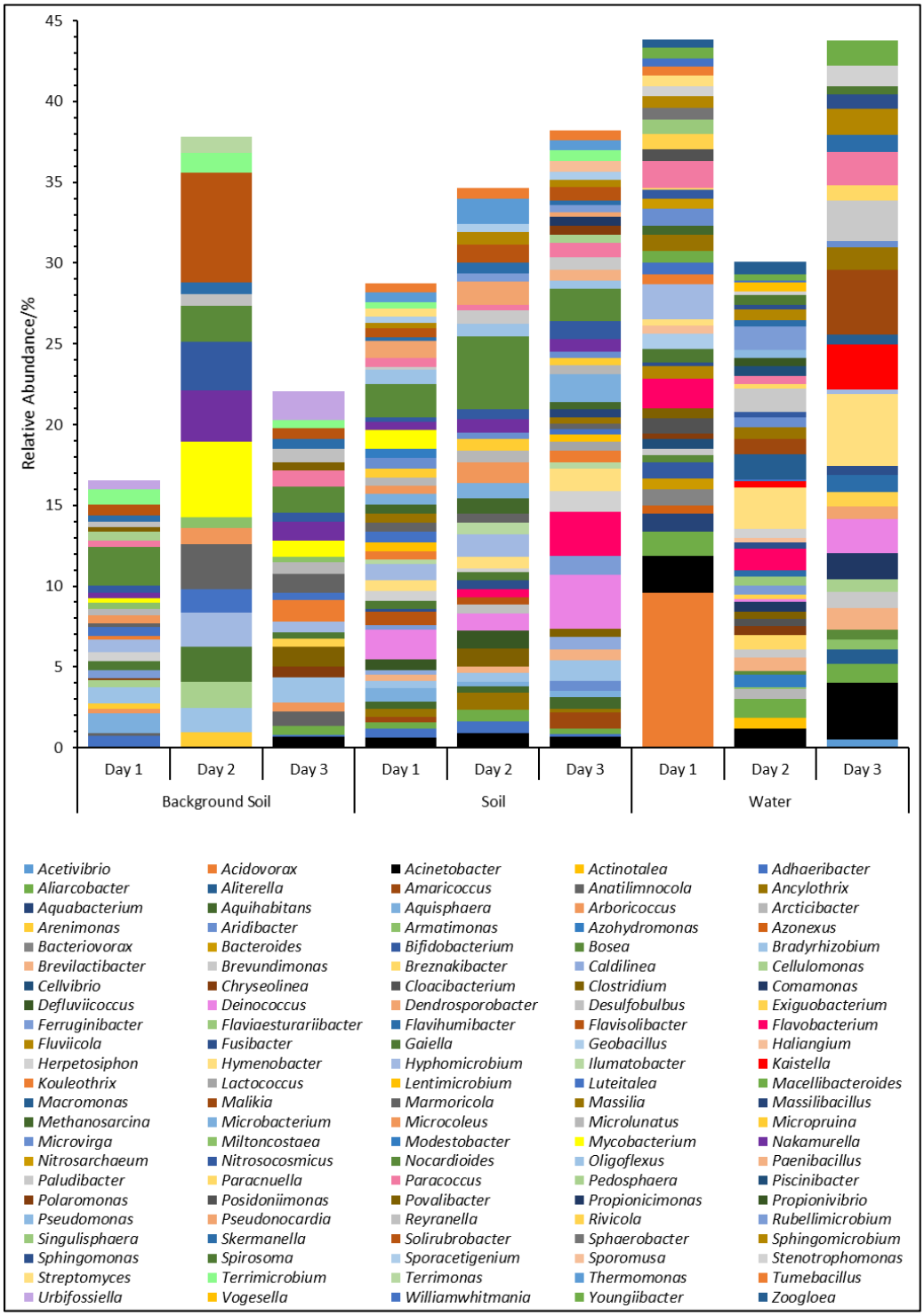


**Figure 49.** Comparison of bacterial communities present in monitored soil ,background soil , and water samples at Genus level (n=9). Filtered to include only genera with a relative abundance of >0.5%, on at least one of the sample days.

*Species*

At species level, all samples were filtered to include species with a relative abundance of >0.5% on at least one sampled day. To accurately describe and analyse the plethora of species present, background soil and monitored soil data will be presented first in Figure 50, with water data separate in figure 51.

After filtering, 29 species were detected in monitored soil, over the 3 sample days (Figure 50). 23 species were detected on day 1. On day 2, the number of species decreased by 6 to 17. On day 3, the number of species increased slightly from 17, to 20. The most abundant species was *Nitrosocosmicus hydrocola,* at >1%. This species was present over all 3 sampling days, with the lowest abundance on day 1- 0.27%, and highest abundance on day 3. 10 species were present in monitored soil for all sampling days, including: *Ancylothrix terrestris, Skermanella aerolata, and Aliarcobacter aquimarinus*- all of which had highest relative abundance on day 2 ( 0.6% -1%).

46 species were found in background soil samples over the sampling days. Of all of those species, 13 were found in both background and monitored soil- leaving 33 species unique to monitored soil.

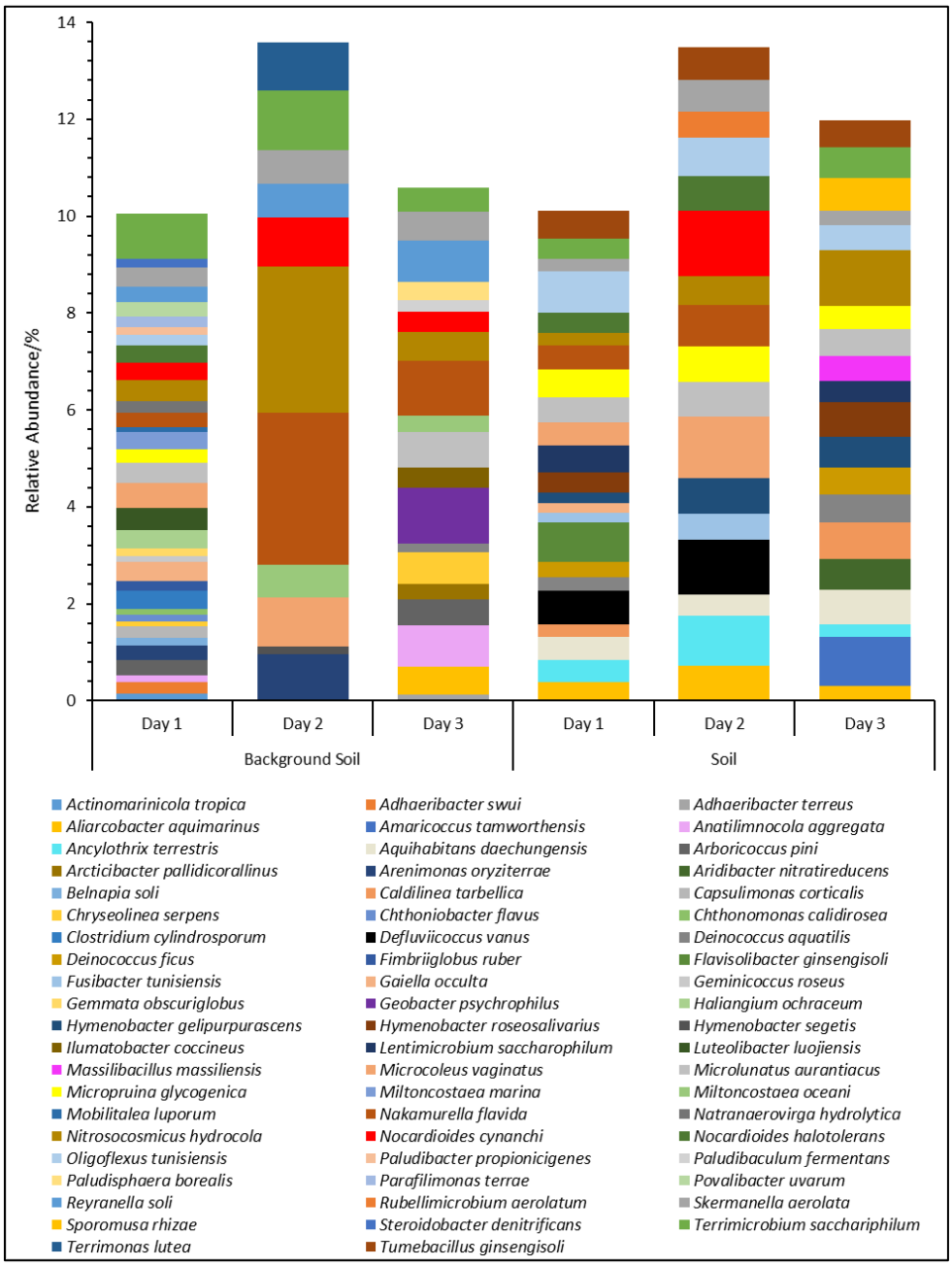


**Figure 50**. Relative abundance of species present in monitored soil and background soil samples only (n=6). Filtered to include only species with a relative abundance of >0.5%, on at least one of the sample days.

31 species were found in water samples over time, representing up to 15% of the total bacterial community (figure 51). On day 1, 18 species were present in water, with 11 of those species present for day 1 only. On day 2, the number of species present decreased by 2 to 16 with 9 'new' species appearing- not present on day 1. The most abundant of which was *Macromonas bipunctata* at >1.5%. On day 3, the number of species decreases again to 12 species. Only 3 species were present in water over all three sampling days*: Paracnuella aquatica, Hymenobacter gelipurpurascens, and Fusibacter tunisiensis.* All 3 species increased in relative abundance over the sampling period- being most abudnant on day 3 at 0.5% -1.5% respectively.

7 species detected in water were also detected in monitored soil. Of those species 5 were not found in background soil. Those species were: *Hymenobacter roseosalivarius, Aliarcobacter aquimarinus, Deinococcus ficus, Fusibacter tunisiensis,* and *Rubellimicrobium aerolatum.*

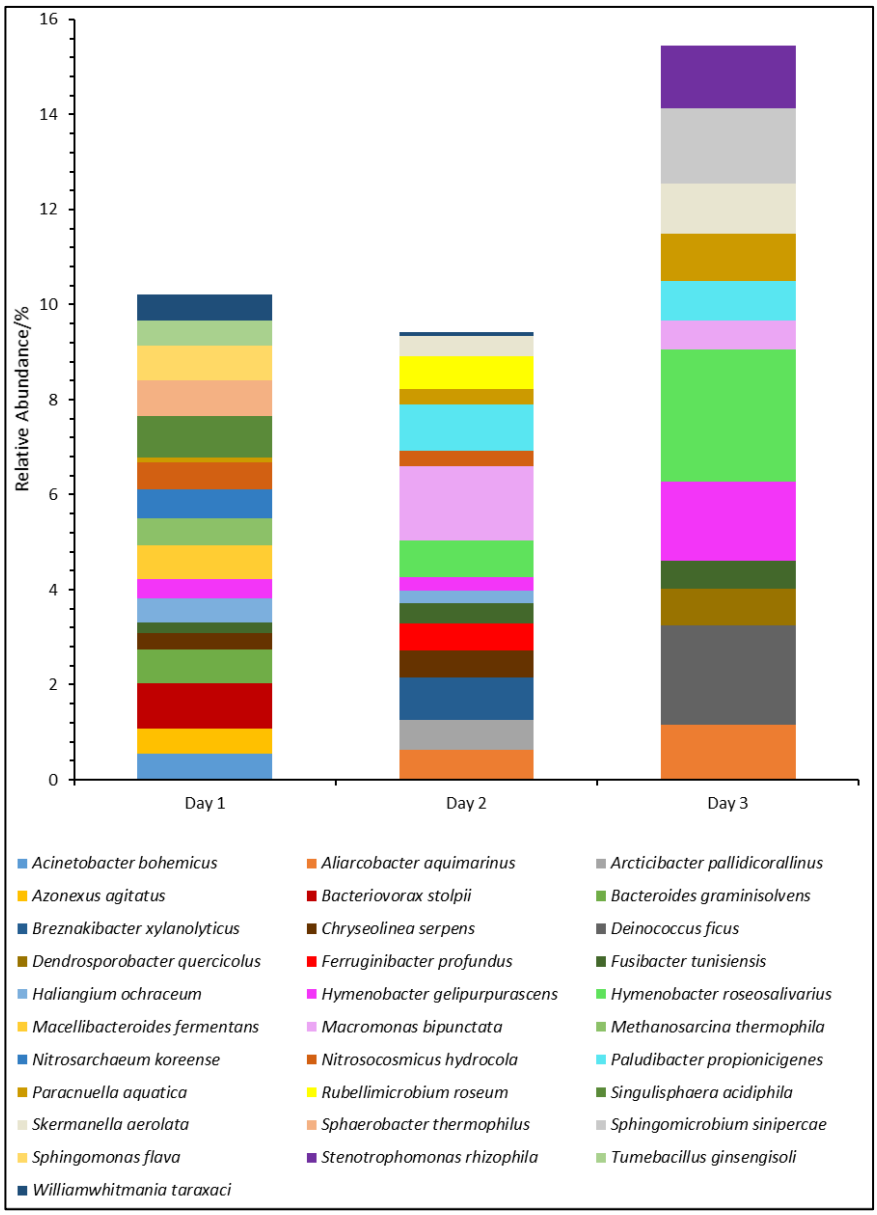


**Figure 51**. Relative abundance of species present in water samples only (n=3). Filtered to include only species with a relative abundance of >0.5%, on at least one of the sample days

#### 5.4.3.3 Diversity Indices

Alpha Diversity

For monitored soil, both ACE Species Richness and Faith's Diversity index values were highest on day 1, and lowest on day 2 (Table 14). Simpson's Evenness showed the opposite trend, with the lowest value on day 1, and highest on day 2. For background soil samples, ACE species richness was variable, ranging from 83-448. Despite this, monitored soil on average had a higher species richness over the sampling period (172-380). In background soil, Faith's diversity index is higher than monitored soil on day 1 only- with its highest index value of 67.84, compared to 46.85 for monitored soil. Simpson's Evenness in background soil ranges from 0.7-0.77, highest on day 3, where it is higher than that of monitored soil (background soil= 0.77, monitored soil=0.75), and lowest on day 1, where the evenness value is equal to that of monitored soil (0.70).

For water samples, ACE species richness varies over time (129-368), but was highest on day 2 at 368. On average, over the sampling period, monitored soil had a higher species richness than water- higher by 43.3. Simpsons evenness index for water remained unchanged for the first two sampled days at 0.71, and then decreased to its lowest value on day 3 of 0.62. Evenness for water samples was only higher than that of monitored soil on day 1, when water evenness was 0.71, and monitored soil was 0.70. Faith's Diversity index for water was highest on day 2 at 43.48, and lowest on day 3 at 22.95. Faith's diversity in water was lower than that of monitored soil on all sample days, apart from day 2, when it was almost double that of monitored soil (monitored soil was 24.28, with water at 43.48).

**Table 14**. Alpha diversity indices for background soil, monitored soil, and water samples taken at Endcliffe Park during June 2023 (n=9).

| **Day** | **Sample Type** | **ACE Species Richness** | **Simpson's Evenness Index** | **Faith's Diversity Index** |
|---|---|---|---|---|
| Day 1 | Monitored Soil | 380 | 0.70 | 46.85 |
| | Background Soil | 448 | 0.70 | 67.84 |
| | Water | 187 | 0.71 | 26.87 |
| Day 2 | Monitored Soil | 172 | 0.81 | 24.28 |
| | Background Soil | 83 | 0.74 | 22.61 |
| | Water | 368 | 0.71 | 43.48 |
| Day 3 | Monitored Soil | 262 | 0.75 | 33.42 |
| | Background Soil | 202 | 0.77 | 31.03 |

| | | | | |
|---|---|---|---|---|
| | Water | 129 | 0.62 | 22.95 |

Beta Diversity

PCoA analysis was performed on beta diversity values for all samples from June 2023. This allowed the comparison of background and monitored soil samples, as well as comparing monitored soil against water. Figure 52, shows the separation of all samples. Background soil sample data points did not clearly cluster together, whereas monitored soil data points formed a more discrete group. When comparing beta diversity of monitored soil and water samples, very clear separation was observed between the two groups.

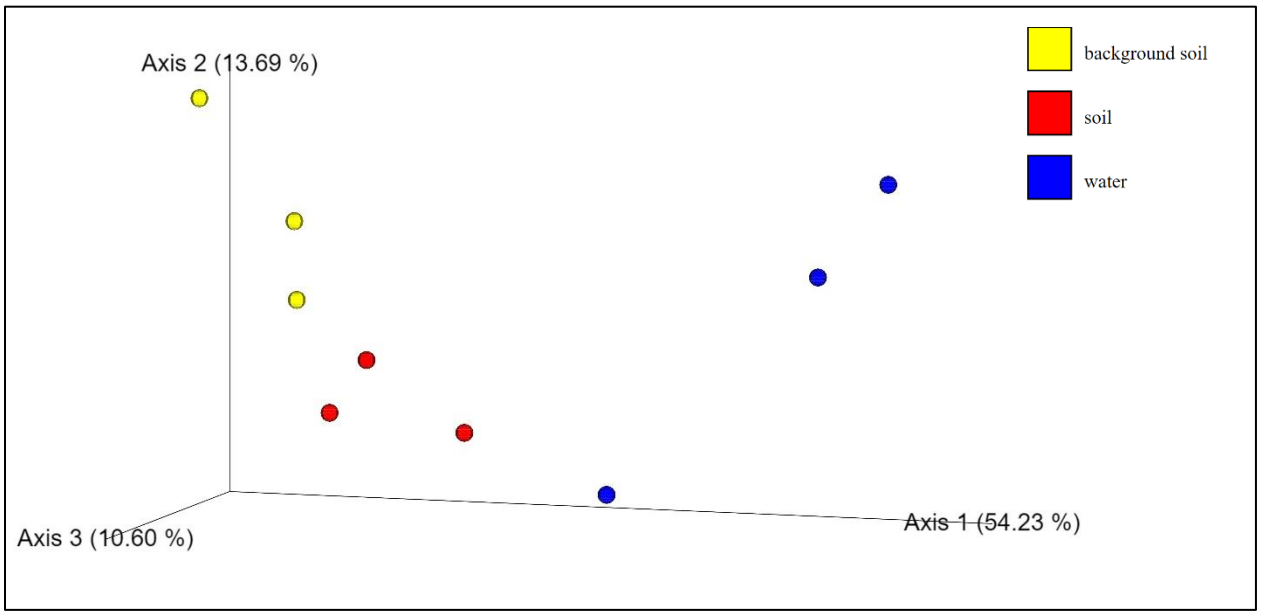


**Figure 52**. PCoA (Principle Coordinate Analysis) Emperor plot based on weighted unifrac distance analysis ( n= 9).

**5.5 Discussion**

5.5.1 Changes in bacterial community at urban flood sites over a 3 day sampling period

It is evident that over the 3 sampling days in both soil and water sampled at both sites that the bacterial community is diverse and dynamic, with bacterial genera often appearing for one day only, before disappearing. Understanding the changes over this short time period will help in determining whether an immediate risk to public health is present and at what timescale the risk could be most prominent.

Changes in the bacterial community structure were seen in floodwater sampled at Endcliffe Park, with the most abundant genus in floodwater on the first sampled day- *Acidovorax*, no longer detected on the subsequent days, despite having high abundance on day one. A similar trend was observed for other genera, for example *Streptomyces* was present on day one only and was absent over subsequent days in soil or water. The genus *Streptomyces* is known to have species that are

pathogenic to humans (including *S. somaliensis* and *S. sudanensis*) (Quintana et al., 2008; Sgorbati et al., 1995) consequently, with *Streptomyces* present, a potential risk to public health could be inferred. Despite this, at species level, no pathogenic *Streptomyce*s bacteria were detected, and thus no risk to public health can be assumed

Looking at the floodwater at Tongue Gutter, the number of different genera present increased over time, as did the abundance of bacteria associated with faecal pollution (during the September visit) (Cao et al., 2014; Wang et al., 2021). This included *Phocaeicola, Roseburi*a, and *Faecalibacterium*-present at species level as *Phoecicola vulgatus, Roseburia faecis, and Faecalibacterium prausnitzii.* The increase in abundance of these bacteria over time could be due to the different source of flood and rainfall pattern. The September sample visit was carried out after several days of rainfall, and the rainfall continued over each sample day. The continuous rainfall could be linked to the increase in the aforementioned bacteria over the days, with the outflow pipe in Tongue Gutter continuously spilling due to rainfall duration. This could be suggested, since many studies have found correlations between rainfall amount and faecal indicator (*E. coli*, enterococci) abundance in floodwaters , as well as incidence of gastrointestinal illness among populations following flooding caused by heavy rainfall (Andrade et al., 2018; Dzodzomenyo et al., 2022; Fang et al., 2018).

At both sites, soil bacteria communities were comprised of a 'core' group of phyla-present in all samples on each sample day. These phyla included: Proteobacteria, Actinobacteria, and Planctomycetes, often found in high abundance in natural soils, irrespective of the land use (Gschwend et al., 2021;Karimi et al., 2018). Slight fluctuations in relative abundance of these phyla were evident over time, but the main differences were seen through the appearance and disappearance of lower abundance phyla over the sample days, such as Synergistota, Dependentiae, and Fusobacterota, thought to have transferred from the floodwater. However, no faecal associated bacteria or pathogenic bacteria were detected in soil on any of the sample days at Endcliffe Park.

Changes in bacterial community of both soil and floodwater over 3 days have been discussed, with no clear overarching trend over both sites and both sample types (soil and water). Instead, it is evident that the bacterial community is dynamic and fluctuates over 3 days. Fluctuations are thought to be influenced by the source of flood (whether faecal contamination is present or not), and also the duration of rainfall. Transfer of bacteria between the floodwater and soil would also explain the changes seen in bacterial community composition over the 3 days and this topic merits further discussion.

### 5.5.2 Dynamics of bacteria across soil-water interface and implications for short term risk to public health

The interface between soil and water is a unique habitat where frequent flood events can shape microbial communities and their dynamics. Among these microorganisms, bacteria play a significant role due to their diverse metabolic activities and potential impacts on public health. Understanding the dynamics of bacteria across the soil-water interface is essential for assessing short-term risks to public health, as these environments can serve as reservoirs for potentially pathogenic bacteria.

In this study, faecal associated bacteria, along with pathogenic species were present at one of the sampling sites – Tongue Gutter - in the flood water during both visits- including *Faecalibacterium prausnitzii, and Dialister invisus*. *F. prausnitzii* is one of the most abundant bacterial species in the human gut and is present in faeces representing around 5% of the total microbial community (Cao et al., 2013; Parsaei et al., 2021). *D. invisus* has only been isolated from humans- in which it is known

to play a role in causing infections of the mouth/gums, and is also present in the gut- but has never been previously isolated from non-anthropogenic sources (Downes et al., 2003). Therefore, this species has previously been identified as a potential indicator of faecal contamination in rivers (Jeong et al., 2011). This study confirms that this bacteria has potential to being used as a faecal indicator of critical contamination conditions (or thresholds) that might compromise public health.

Interestingly, no faecal associated or pathogenic species were detected in monitored soil at any visit, at either site. It is known that water can facilitate the transport, or deposition of bacteria from water to soil, and vice versa (Huysman et al.. 1993; Jamieson et al., 2004; Jefferey and van der Putten, 2011). Although no faecal bacteria appeared to have transferred from water to soil, this study did find evidence that other type of bacteria may have been transported out of floodwater and into the soil, including *Peredibacter* and *Acinetobacter.* These bacteria are both involved in the degradation of organic compounds and plant materials, and were present in soil and water samples, but not in soil background samples in this study (Jung and Park, 2015; Paix and Jacquet, 2017). Thus, indicating movement out of water and into soil. This is confirmed by similar behaviour observed in other bacteria- such as *Flavobacterium* - being present in soil and water samples, but not in background samples at the second sample site- Endcliffe Park. The fact that *Flavobacterium* abundance also increased in soil over time, whilst also decreasing in floodwater, further confirms the theory of movement out of floodwater and into the soil at this site.
It is evident that some transfer of bacteria between soil and water may be happening over the 3 day sampling period. However, this was only observed for certain types of non-faecal bacteria. This raises several questions, including: whether 3 days is a sufficient time period for transfer to occur, or if some bacteria simply do not transfer across the soil-water interface at all. Therefore, future studies should investigate whether certain types of pathogenic bacteria can be transferred from the contaminated flooded water to the soil over longer periods, (> 3 days ), and also if some bacteria are not selectively transferred between the soil-water interface due to certain biochemical/physiological characteristics (cell size, sorption behaviour, chemotaxis etc.) (Bolster et al., 2006; Liu et al., 2017; Mesibov and Adler, 1972).

Previous studies, including results from the long-term study- Chapter 4 of this thesis, have found FIB present in soil in areas that have been flooded, thought to have transferred from floodwater. However, there are no robust guidelines or recommendations for how long after a flood should soil samples be taken to accurately assess the bacterial contamination, or at what time period is bacterial contamination highest, and thus a more concerning risk. Answering these questions would help in understanding at what time after a flood do certain bacteria seem to transfer from floodwater to soil and what type of characteristics make them more prone to be transferred or retained in the soil. To answer this, a future study could be carried out sampling at selection of different time points. For instance, a long term study, taking soil samples 8 months after a major flood event along the river banks of the Elbe and Mulde Rivers (Germany), found that multiple FIB species ( *E. coli*, *Pseudomonas aeruginosa*, *Clostridium perfingens* and *enterococci)*, were extremely variable in soil . Not only this, but the FIB species were not present in levels significantly higher than those found in pre-flood (background samples) (Strauch et al., 2005). It can be suggested that testing 8 months later is not an appropriate time point for sampling and capturing the bacterial activity, and more frequent monitoring in the short term is required.

Looking at a reduced time scale- one week after a flood event- one study found *Clostridium perfingens* levels were significantly higher than in background soil, following a flooding event in North Carolina, USA (Casteel et al., 2006). Similarly, this study showed that at genus level, *Clostridum* was also detected in the monitored soil samples at Tongue Gutter in the 3 days following the flood event, during the September visit, but was never detected in background soil. This indicates movement of *Clostridium* out of floodwater and into the soil- indicating faecal pollution and presenting an infection risk, from not just floodwater , but also the contaminated soil, with species such as *C. perfringens* causing gastrointestinal illness in humans (Labbe, 1991; Matches et al., 1974).

Another study found elevated *E. coli* levels in the first week, following flooding of growing fields in the east of Spain. However, soil tested in later weeks showed dramatically declining *E. coli* numbers (Castro-Ibáñez et al., 2015). This suggests that FIB from floodwater can transfer from water to soil, and survive, for at least one week after a flood, but could there be changes and transfer of bacteria even before one week ? A simulated flood study, by Shiraz et al., 2020 spiked artificial flood water with *E. coli*, and sampled affected soil every 48 hours after the flood water had receded. This study found that an elevated *E. coli* level was not found in the soil until 96 hours post flood (Shiraz et al., 2020). These previous studies, sampling at various time points, could explain why no FIB were detected in soil samples from this study- despite their presence in sampled floodwater. This is because the aforementioned studies showed elevated levels of bacteria between 4-7 days following a flood, not in the time period before (1-3 days). Overall, the other studies discussed above appear to show that 4-7 days are the optimum time to sample affected soil in order to see if any FIB have transferred from flood water (Casteel et al., 2006;Castro-Ibáñez et al., 2015; Shiraz et al., 2020) . This hypothesis on shorter sampling times is interesting, however given the plethora of other factors (sorption behaviour, soil type, cell morphology etc.) that can affect transfer of bacteria from flood water to soil (as discussed in chapter 4), further intensive testing, taking more frequent soil samples would need to be performed to properly assess risk. It can be suggested, that future studies could be carried out in a controlled lab environment- similar to the study undertaken by Shiraz et al, 2020 , using synthetic spiked floodwater to monitor bacterial community changes, in order to determine the transfer time, whilst accounting for and controlling other interfering factors.

It has been discussed that various bacterial genera, including *Flavobacterium,* and *Clostridium*, appear to have moved between the soil-water interface during this 3 day study, but that no pathogenic or faecal associated bacteria were involved. Based on the available literature, this may be due to the timescale over which the study was performed- with movement of faecal bacteria potentially happening after 3 days. It is also plausible that there could be other factors, not investigated in this study, that could determine the moment between the soil-water interface, such as bacterial function or properties.

### 5.5.3 Effects of faecal contamination and flood events on bacterial community structure

Faecal associated bacteria were present in floodwater samples during this study, and in previous studies in aquatic environments suggest that presence of faecal bacteria can influence the microbial community structure as a whole. For example, one study on the Arga River in Spain, found the

presence of *Acinetobacter spp*. only in samples downstream from a wastewater discharge point, with this species not found in any upstream samples (Goñi-Urriza et al., 1999). *Acinetobacter* was also detected in the water samples analysed at Tongue Gutter in this study- the site in which faecal bacteria were also detected, but not at Endcliffe Park, where no faecal bacteria were detected. This indicates that this difference in community structure (presence or absence of *Acinetobacter*) could be due to the presence of faecal contamination at one site, but not at the other. Not only this, but alpha diversity of the soil bacterial community was higher at the site not found to have faecal bacteria, in comparison to the site with faecal bacteria- indicating, that faecal contamination of floodwater may affect the microbial diversity of soils. Bacterial diversity has also been found to be reduced in contaminated waters too- with one study , looking at the effect of faecal contamination on several water bodies in Norway (creek, urban stream, reservoir), finding that severely polluted waters had a greatly reduced microbial diversity (Paruch et al., 2019). This could be due to bacteria originating from wastewater 'outcompeting' the usual phyla present. Especially since the same study found that water from an urban stream known to receive human waste, had *Faecalibacterium* constituting over 50% of the total relative abundance of the bacterial community. To our knowledge, this genus has not been detected in other samples taken from less polluted sources, or sources thought to be polluted with animal waste (Paruch et al., 2019). The introduction of faecal pollutants unbalances the water biochemistry and ecological balance, causing stress to some bacterial populations (changes in pH, temperature, organic content etc) leading to faecal bacteria thriving, and other bacteria becoming suppressed- reshaping the structure of the bacterial community (Yadav and Sharma, 2020). This could also be the case for our study, with faecal contamination at Tongue Gutter introducing 'new' bacteria to the environment that become dominant, suppressing the existing bacterial community, and thus reducing diversity.

As well as contamination effecting taxonomic bacterial structure, it can also affect microbial activity, with changes in the both often happening in parallel (Liu et al., 2021). After irrigating soil with treated wastewater effluent, one study found that all measured microbial parameters- including enzyme activity, nitrogen mineralisation and total organic carbon – were negatively impacted by pollution (Kayikcioglu, 2012). Another study, investigating flooding effects on flood-plain soil, also found changes in enzyme activity- with enzymes involved in carbohydrate metabolism significantly increasing in flooded soils, and taking up to 24 days to return to pre-flood levels (Wilson et al., 2011). Microbial activity was not measured in our study, however going forward this would be a useful parameter to measure to compliment taxonomic data.

It can be also hypothesised that bacterial community structure is not only affected by contamination, but also by the actual flood event. Frequent flooding, and drying out cycles- like that observed at our study's sample sites, can have a significant impact on the bacterial soil community (Fierer et al., 2003; Yu et al., 2014). Such cycles create physiological stress for bacteria, by changing factors such as pH, nitrogen content, and aerobic/anaerobic conditions (Gordon et al., 2008; Li et al., 2021). For example, it has been suggested that under wetting and drying conditions, bacterial community structure in a wetland sediment changed to favour specialist bacteria that can cope with such conditions (Foulquier et al., 2013).

It can be concluded that faecal contamination affect bacterial communities- altering structure and composition, with contaminated water often introducing dominant bacteria that then suppress the

existing populations, reducing diversity, as seen at Tongue Gutter. Contamination can also disrupt the physico-chemical balance of receiving waters and soils, at detriment to bacteria present, inhibiting microbial activity such as the metabolism of carbohydrates and nitrogen. Bacterial community structure and activity can also be impacted by the changing environment created by frequent floods- namely wetting and drying cycles, causing stress for bacteria.

### 5.5.4 Evaluation of the suitability of coliform counts and next generation sequencing for detecting faecal contamination at urban flood sites

Despite the well-known limitations of culture-dependent methodologies, they are the current regulatory requirement used to routinely monitor microbial quality of water, including the detection of faecal contamination in bathing and drinking waters (Drinking Water Inspectorate, 2018; European Union, 2006) . In this study, despite coliforms being detected using a culture-based method, when the same samples were sequenced, neither *E. coli*, or *C. freundii* were found at species level. Not only this, but faecal bacteria were actually only detected via sequencing in water samples at one site- Tongue Gutter, despite coliforms being present in water and soil at both sites. This raises questions of the suitability of using traditional FIB to inform on risks in soil and water at an urban flood site. This question has been raised by many previous studies and so will be discussed below, together with the lack of reliable standards for flooding.

In this study coliform counts ranged from $1.7 \times 10^5$ to $7 \times 10^5$CFU/ml in flood water samples. Although no legal 'limits' for urban flood water coliform counts exist, the counts here can be compared to the limits set out for maintaining water quality of bathing waters, to understand the level of contamination. The Bathing Water Directive states a mandatory standard of 10000 total coliforms per ml is the limit for bathing sites to be considered safe for humans (European Union, 2006). This means the floodwater in this study contained coliforms at levels 17-70 times over that considered safe for humans to bathe in. Naturally, humans would not be expected to be exposed to floodwater as one might at a bathing site, but this is interesting for a perspective on coliform levels and faecal contamination. Standards exist for coliform loads for bathing/recreational water (albeit not for flood water), however the same standards do not exist for affected soils and sediments. This is despite many studies showing that post-flood, faecal coliforms are present (Casteel et al., 2006; Shiraz et al., 2020; Tandlich et al., 2016). This study found coliforms present in flood affected soil between 50000-780000 CFU's/g, with a similar urban flood study finding even higher levels,- $1.08 \times 10^7$ *E. coli* CFUs per gram of sediment (ten Veldhuis et al., 2010). This is of concern, as one of the sites sampled during this study, was a recreational area with a children's playground, meaning there is an increased chance of a child (which are more vulnerable to infections than adults) coming into contact with bacteria in the soil through play- and thus potentially becoming ill if said bacteria has the potential to cause disease.

Given that floodwater and affected soil have such a high number of coliforms, the question arises if is it efficient to use traditional methodologies designed for drinking and bathing water to assess contamination? Especially given the lower coliform counts usually found in bathing and drinking waters (Barrell et al., 2000; Karunakaran et al., 2024; Rees et al., 1998). Culture-dependent methods are convenient surveillance tools given that they are simple to perform and relatively low-cost ways

of detecting general microbial failures. However, they are only representative of a limited and specific fraction of microbial communities in environmental samples. As well as this, using such culture-based methods may underestimate the true levels of bacteria present, as only viable, culturable bacteria would be detected. Other bacteria could be there that are viable but due to stress encountered in the environment may not be cultured (Price and Wildeboer, 2017).
Based on results from this study, it can be confirmed that alternative molecular methods are more suitable for indicating not only the presence of faecal contamination but also to provide more reliable information on the presence of microorganisms including pathogens. Despite its clear advantages, NGS also has limitations. It allows identification of bacteria down to species level, but this technology relies on assignments made against existing databases, and so this limits the level of identification to known species and strains only (Acharya et al., 2019). In our study many sequences were 'unassigned', unable to be matched to the database, and so information on the entire bacterial community is potentially missed. However, it should be noted that only a fragment of the 16s RNA gene was sequenced (approximately 400 bp) to yield the results in this study, but today it is possible to sequence the whole 16s rRNA gene to improve identification based on taxonomical assignments (Johnson et al., 2019; Petti et al., 2005).

The limitations of using culture-based methods alone, were also seen in the data gained from samples from the Endcliffe Park site. Despite the samples from this site having no faecal bacteria present in sequencing data, coliforms (*E. coli and C. freundii*) were detected when using the culture-based spread plate method. This raises the question of how well coliform counts accurately detect faecal pollution, as in this study, they seem to have 'worked' at one site, but not at the other.
It can be proposed that the solution, for detecting and preventing faecal contamination, at floodwater sites, is to use a combination of the two methods, as done in this study. Using coliform counts as a general indicator as a relatively fast, cost -effective screening method of whether faecal contamination is likely to be present, and then utilising NGS to assess the real bacterial community present, as well as even exploring other microbial risks such as viruses or fungi. NGS has proven to be a useful monitoring tool for environmental surveillance, with one study finding NGS detected over double the amount of enterovirus serotypes (strains differentiated by surface molecules or antigens) in sewage compared to cell culture alone (Tao et al., 2020). Another study found that used NGS to monitor levels of *Bacteroides* efficiently detected faecal contamination in a river originating from a Combined Sewer Overflow (Ekhlas et al., 2021). NGS has proved a useful tool for water quality monitoring in many areas: detecting pathogens, highlighting potential indicator species, and even being used to monitor the dissemination of antibiotic genes in contaminated waters (Tan et al., 2015). With the advancements in sequencing methods, such as the portable Oxford Nanopore MinION (MinION, Oxford Nanopore Technologies, UK) , routine , *in-situ* field sequencing would be possible in the near future, making sequencing a more robust and effective method for environmental surveillance. The MinION has already proved as effective as the laboratory-based Illumina MiSeq sequencing method in detecting faecal bacteria in contaminated river water in Nepal, at a family and genera level (Acharya et al., 2019). As well as this, the MinION has even been used to successfully detect and quantify resistance genes for multiple antibiotic resistant (MAR) coliform bacteria at a wastewater treatment plant in Hong Kong (Xia et al., 2017). With further optimisation of MinION platforms, with additional advances in portable DNA extraction and cleaning, and especially reduction in sequencing errors (Hu et al., 2018; Werner et al., 2022), the MinION could become an essential tool for rapid detection of faecal contamination in urban floodwaters.

In summary**,** this study did find faecal associated bacteria present in floodwater at the Tongue Gutter site, with abundance of such species increasing over time. This indicates faecal contamination at this site, with the continuous rainfall potentially responsible for the increasing abundance of faecal associated species. However, this is theoretic and a wider range of rainfall data and water samples would need to be collected to determine if a relationship is present. The current data however, is useful in communicating that the floodwater at this site could pose a potential health risk in the 3 days following a flood. Quantitative analysis such as QPCR, along with epidemiological studies would need to be carried out to compliment our data in order to understand whether the bacteria exist in the floodwater in a high enough amount to cause illness. On the other hand, no faecal associated or pathogenic species were detected in soil at either site in the three days following a flood- so it seems that in the short term, the affected soil at the urban flood sites is 'safe' , with the floodwater being the main vector of potential public health risk. Finally, data from this study contributed to the discussion as to whether traditional methods for detecting faecal contamination are suitable in an urban flood scenario. It was found that coliform counts are useful I indicating whether faecal contamination is present, but a combination of methodologies, might be required to accurately and robustly assess, not only faecal contamination, but also the presence of potentially disease causing pathogens.

**5.6 Conclusions**

From this research it can be concluded that:

- Faecal bacteria and potentially pathogenic species (*F. prausnitzii, D. invisus*) were detected in floodwater at one sampling site- Tongue Gutter, increasing in relative abundance over the 3 sampling days.
- Transfer of bacteria, such as *Acinetobacter* and *Peredibacter*, between the soil-water interface is thought to have occurred at both sites.
- The presence of faecal contamination has an impact on the soil bacterial community, with alpha diversity lower at Tongue Gutter (site though to have faecal contamination)
- No faecal bacteria or pathogenic species were detected in monitored soil on any visit at either site.
- To accurately determine bacterial risks at an urban flood site a combination of traditional coliform counts and a more specific molecular method such as next generation sequencing is required.

This study provides new insights on the presence of bacteria , including faecal bacteria and pathogens, in the immediate time period of 3 days, following an urban flood event. Species likely to pose a risk to public health have been identified, and the levels of which are linked to rainfall amount, as well as flood origin. Much like the findings of Chapter 4- in order to determine how likely a member of the public is to not only encounter a disease causing bacteria, but become ill from the contact, further social and epidemiological studies on dose amounts and human behaviour around floods would need to be carried out. Another suggested study to further this research would be to sample soil at an urban flood site but for a longer time period of 5-7 days, as data from this study, and the literature on this topic seem to suggest transfer of bacteria from flood water tends to happen between 4 and 7 days after a flood event. Data from this study compliments the findings of Chapter 4, adding an immediate, short term perspective on bacterial risks following an urban flood.

# Chapter 6
# Investigating the distribution of bacteria through a soil column under urban flood conditions

## 6.1 Abstract

This study investigated the distribution of bacteria through a soil column, using urban floodwater collected from a field site and a series of laboratory 'column' experiments. First, the hydraulic properties of the sandy loam soil column were investigated using salt tracer experiments, with data fitted to an ADE model (advection diffusion equation). The spatial and temporal dynamics of bacteria found in both the floodwater and in the soil were then determined using traditional methods- heterotrophic plate counts, and molecular methods- sequencing of the 16s rRNA gene. Such methods allowed bacterial abundance to be determined, as well as identifying bacteria, taxonomically, down to species level. This allowed for the distribution and transport time of a solute (tracer) to be compared to that of bacteria suspended in water. Results of this study showed that after application of floodwater, bacterial abundance was generally highest between the soil surface and 0.4m depth. This study also found that bacterial transport through the soil was slower than that of a solute- largely thought to be due to the sandy loam soil type, aiding in slowing bacterial transport, acting as a filter, and potentially trapping bacteria. The slower transport of certain types of bacteria could be due to different cell properties that favour adsorption to soil particles or attachment to other microbes in soil biofilms. Although this study did find that bacteria may have transferred from soil to water and vice versa, it was also found that one pathogenic species- *Dialister invisus*, originating from the faecal contaminated floodwater, did not transfer into the soil column, and remained suspended in the water. This study contributes to the discussion around bacterial behaviour during an urban flood, investigating not only the dynamics governing bacterial transport-specifically in an urban flood scenario, but also the consequences bacterial transport may have to public health following a flood.

**Keywords:** soil characteristics, soil column, bacterial distribution, urban flood.

## 6.2 Introduction

Infiltration of urban floodwater into soils is an important process to understand when considering the potential health risks posed by an urban flood event. Urban floodwaters can become contaminated with sewage, often containing disease causing or pathogenic bacteria (Basri et al., 2015; De Man et al.,2014) . These bacteria can pose a risk if contact is made with the contaminated flood water, but may pose another longer term risk, if bacteria are transported into the soil, and persist in the soil, after the floodwater has withdrawn (Gelsomino et al., 2006). This could be from direct contact between a person and the contaminated soil, but also, a risk could be posed if bacteria from the contaminated soil, 're-enter' surface water following a subsequent rainfall event. This risk is related to whether any disease causing, or pathogenic species are present in the floodwater, and whether they become adsorbed to the soil matrix, or remain suspended in the water.

Column studies through porous material such as soil, have proven to be useful tools for investigating not only bacterial transport, but also the transport of chemicals, solutes, and viruses (Gilbert *et al.,* 2014; Smajstrla, 1985). When focusing on the specific aspect of bacterial transport (often referred to as colloid transport) in the context of an urban flood, a column study allows the behaviour and dynamics of the whole floodwater bacterial community to be investigated, helping to determine the risk of contamination posed to receiving urban soils.

There are many factors that can influence the transport of bacteria through a saturated porous media. This includes: biological factors, such as cell size and morphology, hydrophobicity or charge of the cell surface, and motility (Bai et al., 2016; Gannon et al., 1990. Chemical factors, such as solution composition (pH, temperature, concentration gradients etc) (Hong and Brown, 2006; Santore, 2022), and physical properties such as flow rate of the solution, advection and diffusion, and adsorption/desorption (Flemming and Wingender, 2010; Jefferson et al. 2005; Lewis, 2005). All of these factors combine to influence bacterial transport, and no one factor alone is responsible for the rate of transport. This makes the topic of determining the transport and thus distribution of bacteria in floodwater through a soil column even more complex.
Adding to the complexity is the composition of floodwater- with variable physico-chemical factors, a plethora of different species present, along with organic debris. Most transport studies tend to focus on determining the transport dynamics of one or very few species at a time (Cunningham et al., 1999; Petersen et al., 2012) - however this does not represent the real-life bacterial communities present in urban floods. It also does not allow for the investigation of bacterial interactions, such as agglomeration and biofilm formation, which can also affect transport (Fenchel, 2002; Young and Crawford, 2004).
With this in mind, this study aims to: determine the spatial and temporal distribution of bacterial communities present in floodwater when mobilised through a soil column, and also understand how in turn, this effects the soil bacterial community.

## 6.3 Materials and methods

In order to analyse the movement of bacteria originating from floodwater across the soil-water interface and into the soil profile, a column experiment was designed. This was designed to replicate a small section of soil from the urban environment and subject it to 'urban flood' conditions, by infiltrating contaminated water through the column, taking water samples temporally and soil samples at different depths once the floodwater had run through the column. This allowed changes in behaviour and distribution of present potential pathogens within the soil to be tracked.
Prior to the experiments using floodwater, salt tracer tests were carried out on the soil column to determine appropriate flow rates to use, find the infiltration rate of the soil and solute diffusivity characteristics, and also estimate when would be best to take samples from the column to track potential movement of pathogens (via water) through the column.

### 6.3.1 Column Design

It was not only important that the column depth represented field conditions and spatial/ temporal scales relevant to urban flood events, but it was also important that the depth of the column was not too large, as practicalities, such as taking samples, and filling and emptying the column may have

become too difficult. Although many soil column studies have been carried out, there is no standardised method for designing or running a soil column experiment (Lewis and Sjöstrom, 2010). Studies using this experimental design use columns constructed from inert materials, such as steel, glass, acrylic, and PVC (Gilbert et al., 2014). For this investigation, acrylic or PVC would appear to be the best choice, in terms of rigidity, durability, as well as being inert and transparent - allowing the infiltration through the column to be viewed visually (Hebig et al., 2014). Column diameter is another important factor, and in this case a larger diameter would be ideal, in replicating conditions in the field and reducing boundary effects. Although some studies do use columns with diameters of 10cm, It is recommended that the diameter is a minimum of 20cm to avoid the effects of the wall on flow (Sentenac et al., 2001). To reduce the effects of wall flow further, sand could be painted on the inside walls of the column, or the inner walls could be 'roughened' (Smajstrla, 1985). There is also the option of adding annular rings, around 10-15mm to the inner circumference of the column (Gilbert et al., 2014). In terms of column depth the most common depth to use is between 1.5 and 2.5m, although some studies have used columns with depths of only 1.4cm (Voegelin et al., 2003; Lewis and Sjöstrom, 2010).
Therefore based on a review of the literature mentioned above, a column made of inert acrylic was designed with a depth of 1.28 metres and an inner diameter of 24cm (Figures 53 and 54).

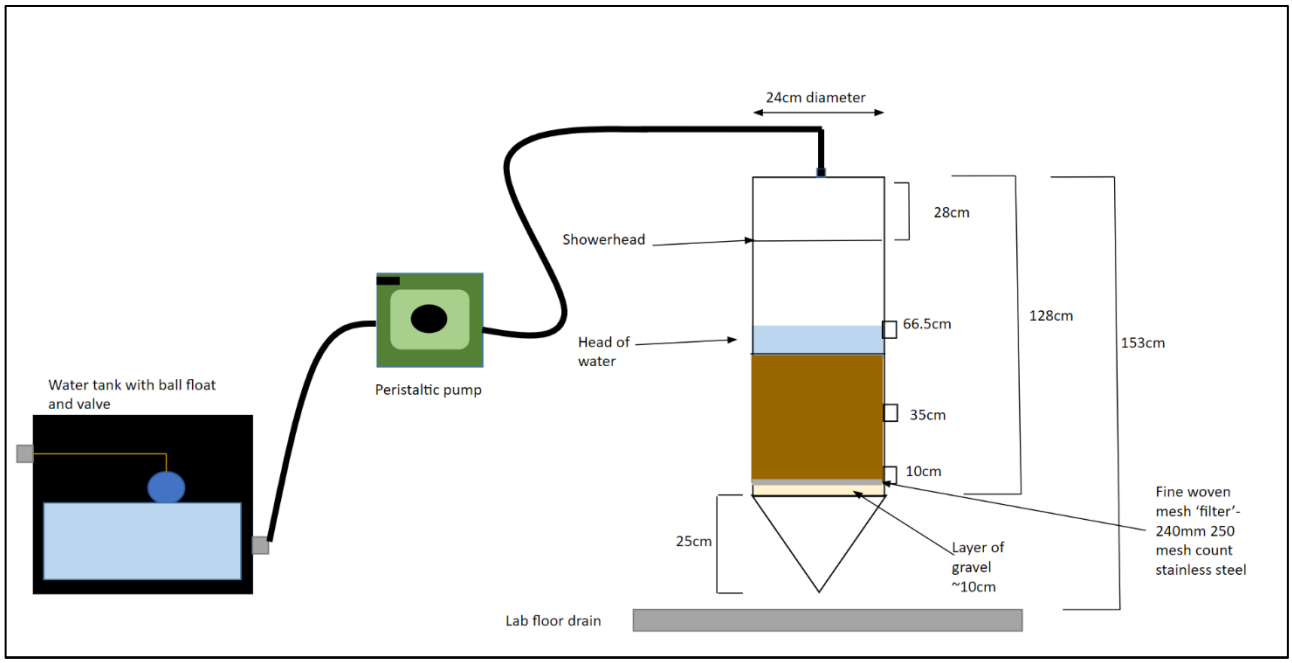


**Figure 53**. Illustrative diagram of column set up.

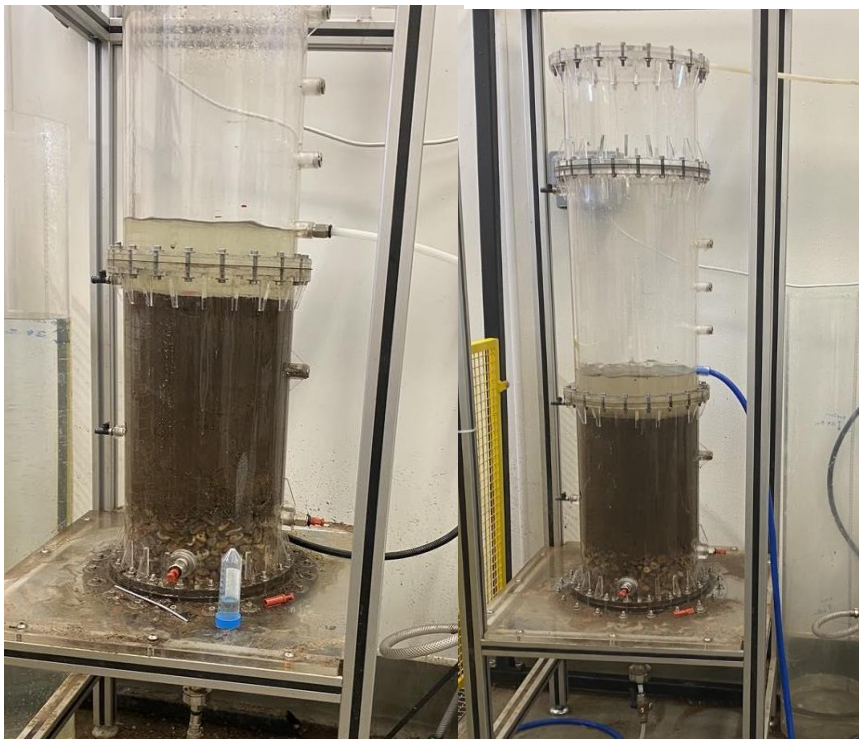

**Figure 54.** General column setup. Photo taken during a wastewater trial in September 2023.

The aim, was to fill the column with soil/gravel to a total depth of approximately 50cm (+/-5cm). This was because past work has shown that changes in pathogen distribution seem to slow as the soil depth reaches around 60cm, with more rapid changes in the distribution of pathogens occurring towards the soil's surface (Boyer *et al.,* 2009; Lance and Gerba, 1984; Petersen *et al.,* 2012).
The bottom of the column was fitted with a 240mm 250 mesh count stainless steel sheet-, across the diameter. This allowed water to drain out of the soil, without losing large amounts of the soil material (Gilbert *et al.,* 2014). Initial experiments showed that finer components of the soil were translocating to the bottom of the column-blocking the filter, and decreasing the hydraulic conductivity. Therefore, a 12 cm layer of pea gravel was also added to the bottom of the column with the mesh filter sitting above. This aided in evenly distributing the flow. A funnel with a tap attached was fitted at the bottom of the column- this allowed the column to be 'opened and closed'. Any effluent from when the column was 'open' would flow into a floor drain (Figure 55).

**Figure 55.** Funnel at bottom of column with tap attached. Pipe leads into lab floor drain.

At the top of the column was a 'showerhead' plate, which worked to disperse the incoming water (Figure 56). Water flow into the column was via a peristaltic pump (Verderflex EV 8000, Economy peristaltic cased tube pump, Scientific Laboratory Supplies, United Kingdom), with adjustable speed. The flow rate through the column was controlled, not by the pump settings, but by a combination of soil type, soil density, as well as the pressure exerted on the soil column by a head of water that sat on the soil surface. A constant head/depth was maintained (using an overflow pipe, see below), in order to create a consistent flow through the column for each test.

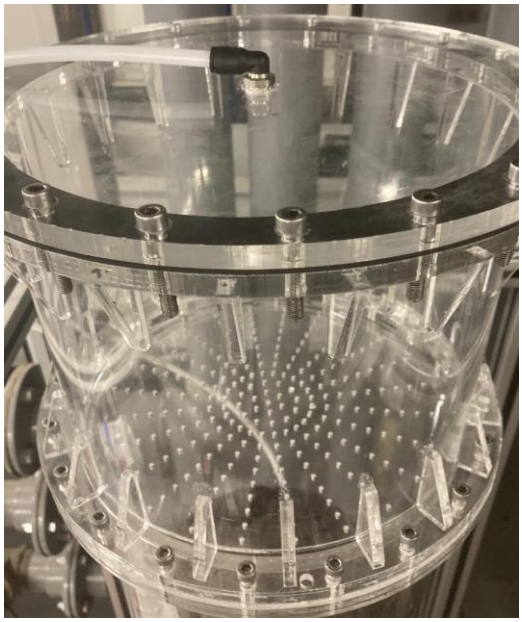

**Figure 56.** Shower head plate fitted at top of column.

Sampling ports were required at intervals along the column in order to take samples of soil . Soil ports were located (from the bottom of the column, upwards) at 10cm, 35cm, and 66.5cm. The port at 66.5cm was used to fit a pipe to allow excess water to flow out- maintaining a constant head of water sitting on the column's surface (Figure 57). Soil samples were taken horizontally through the

ports using a coring device (Figure 58), as well as vertically from the soils surface. Water samples were taken from the outflow tap at the bottom of the column (Figure 55).

**Figure 57**. Port used to allow flow of excess water to maintain a constant head/depth.

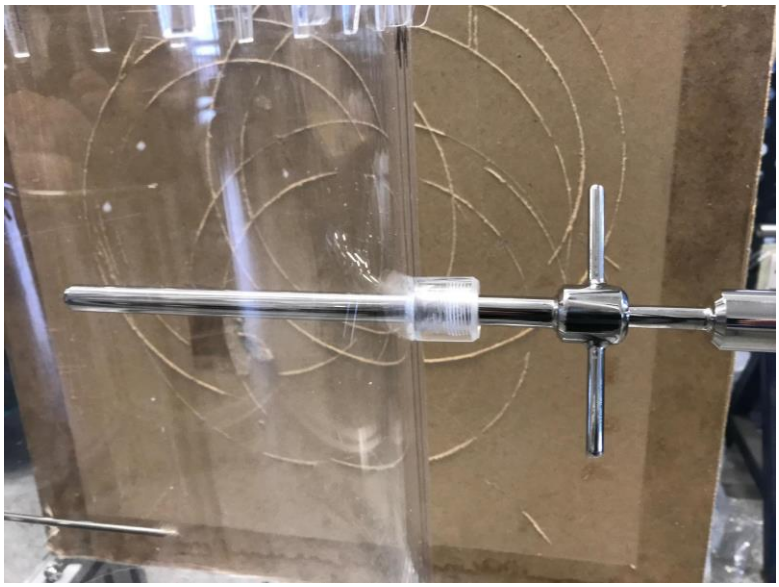

**Figure 58.** Soil corer in empty column. Used to withdraw soil samples from column.

6.3.2 Column Soil Type and Soil Classification

In terms of this study, a sandy loam soil was a good choice to use as a starting point, as it is composed of clay, silt, sand and organic matter mimicking a 'typical' soil found in park/garden areas of urban regions (Gill *et al.,* 2000). To ensure experiments could be performed practically, and samples taken within a reasonable time frame (≤8 hours), initial tests showed that a loam soil with a sand content of approximately 70% works best. The soil was mixed by hand in a large batch combining 30% loam soil (Hallstone Blended Loam, Travis Perkins, UK), with 70% coarse sand (0-4mm grain size) (Grit/Sharp sand, Travis Perkins, UK) .Once mixed, soil was sized and classified using

the sieve method as described in Chapter 4 to ensure the ideal soil type had been created (see soil gradation in Results section).

6.3.3 Flood water used for column experiments

Instead of creating synthetic floodwater, floodwater was taken from the field to use in the column experiments. Flood water was taken from the site 'Tongue Gutter' (described in Chapter 4), this was due to the fact the outflow pipe at this site spilled almost constantly and so it was guaranteed there would be enough water to collect. This site was also known to be contaminated with faecal bacteria from a typical UK urban drainage system, thus providing a diverse bacterial community to study and which humans may be exposed to from an urban flood event.

On the morning of each soil trial, approximately 100 litres of 'floodwater'- water originating from the pipe and forming a pool (Figure 59) was collected in sterile cans and transported to the laboratory within 30 minutes.

**Figure 59**. Outflow pipe and flood water at Tongue Gutter. Water was collected from here to use in soil column trials. Taken on a visit in August 2023.

Water was filtered through a 240mm 250 mesh count stainless steel sheet to remove any large pieces of debris to prevent any blockages, before being used in the column experiments.

6.3.4 Preparing column for a floodwater trial

A layer of round pea gravel (20mm sized gravel) was added to the bottom of the column. Any pieces of gravel with jagged edges were removed, and only smooth/round pieces were used. The layer aimed to reach around 12cm in depth. Once added the surface was lightly compacted using a custom made 'tamper tool'. This consisted of a metal pole with a weighted metal plate on the end with a diameter of one centimetre less than the column's diameter. This ensured a level surface and reduced any gaps in the layer. The mesh filter (240mm 250 mesh count stainless steel sheet) was placed on top of the gravel layer, with soil packed on top of this.
For all trials, the soil mixture was sieved into the column using a geotechnical sieve with an aperture size of 3.35mm. This was to remove any plant material or larger stones, as well as preventing 'clumps' of soil material, it aided in uniform packing. Sieved soil was added in 10cm increments. Each 10cm section was levelled and compacted using light pressure from the 'tamper' tool. The tamper helped to create consistent soil layers. After each layer was compacted gently, the surface was scarified with a scalpel- this helped the next soil layer adhere, creating a cohesive column of soil- avoiding the creation of discrete layers of soil. The column was packed in this way for each test.

*Saturation with clean water*

Once packed, clean tap water was added to saturate the column from the bottom up. This aimed to remove any air bubbles, and allow the soil to 'settle' into place. Water was pumped slowly into the column using the peristaltic pump through a port fitted at ~10cm from the bottom of the column (Figure 60). Water was added until the soil was saturated up to ~2cm below the soil surface. The saturated column was then left overnight (~16 hours), with the tap at the outflow point closed.

**Figure 60.** Soil column being saturated with water, using the bottom-most port. Note visible 'water-line' indicating level of saturation.

6.3.5 Salt Tracer Experiments

4 salt trials were carried out (referred to as Salt trials 1, 2, 3, 4), with the column setup as shown in figure 61.
As well as gaining practical experience and understanding typical transport timescales within the column, the aim of the experiment was to identify and quantify the scale of the physical/diffusion properties of the soil type setup using an inert tracer. This would enable the comparison of the inert tracer transport and bacterial transport (when comparing to floodwater trial data). Thus allowing for the differentiation of transport processes which are a function of the physical characters of the soil (via mechanical diffusion) and the biological properties of bacteria.

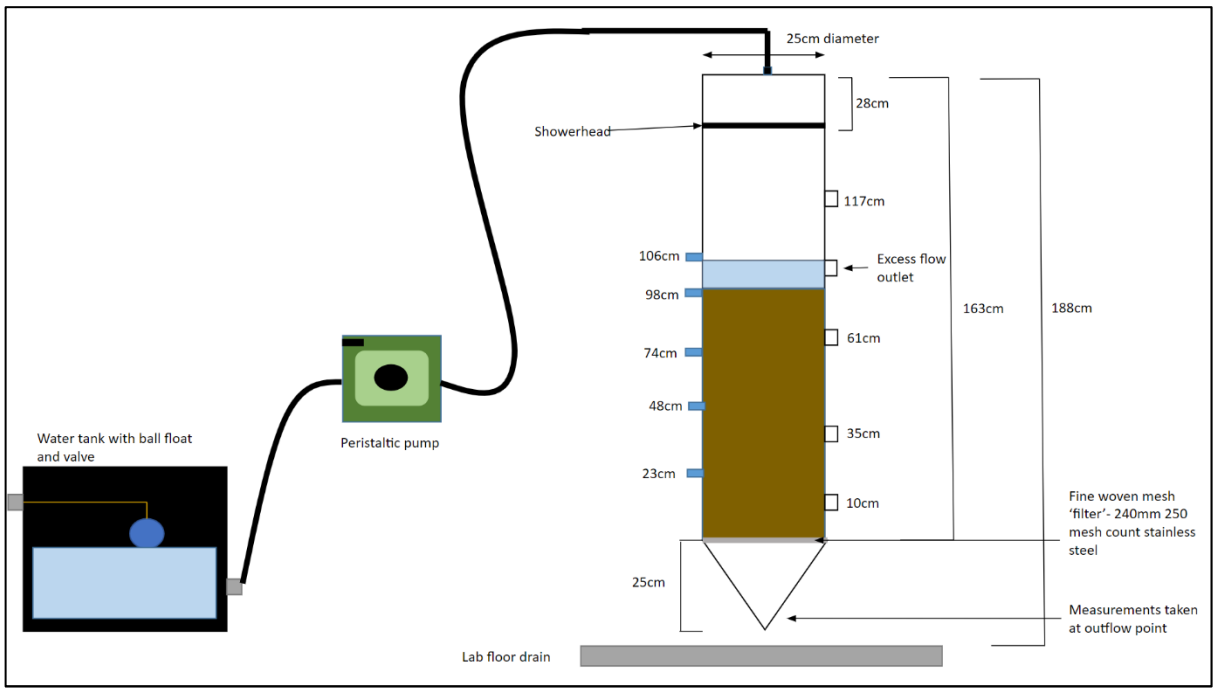


**Figure 61.** Illustrative diagram of column set up for preliminary salt tracer experiments. Sample port labels indicate the position of the port in centimetres from the bottom of column.

*Column set up and soil type*

For all trials, before any salt solution was added, clean water was pumped through the column until there was a steady head of water on top of the soil column and equilibrium was achieved (i.e. the column inflow was equal to the outflow less the excess). Flow rate was measured by weighing the amount of water collected in a bucket over 5 minutes, and dividing the weight by 5 to determine flow rate in litres per minute. The starting conductivity of the clean water outflow was measured using a conductivity probe (470 Conductivity Meter, Jenway, UK).

The salt solution itself was a mixture of clean tap water and basic 1 mole sodium chloride (salt) (Fisher Scientific,UK). The conductivity of the solution varied for each trial, however the conductivity was always higher than the clean water outflow conductivity, and lower than the probes maximum. All parameters for each trial are outlined in Table 15.

*Trial runs and conductivity measurements*

Following the acquisition of the equilibrium flow rate with clean water, the inflow was switched to the salt solution and measurements were taken regularly at the bottom of the column to characterise the breakthrough curve of the salt solution. There were no set guidelines in place for how frequently to sample during the trials- each trial was investigative, and as such, the results from one, informed the next. For example, for Salt trial 1, conductivity measurements were taken every 1-2 minutes, however for Salt trials 1 and 2 measurements were taken every 5-10 minutes, and for Salt trial 3, this was every 10 minutes.

*Data*

To plot and compare all results together, proportional conductivity was used- due to each trial using salt solutions of differing initial conductivities. To do this the following calculation was performed:

$$\frac{(c - c^0)}{(c^{max} - c^0)} \tag{6.1}$$

Where: $c$ = measured conductivity, $c^0$ = conductivity at time '0', and $c^{max}$ = maximum conductivity recorded.

In order to characterise the diffusivity of the soil column for each test, experimental results were fitted to the 1D-Advection Diffusion Equation (ADE) utilising a standard model routing procedure. The 1D ADE is given as (Rutherford, 1994):

$$\frac{\partial C}{\partial t} + V_x \frac{\partial C}{\partial x} = D_x \frac{\partial^2 C}{\partial x^2} \tag{6.2}$$

Where C = Concentration (l/l), Vx = Mean Velocity (m/s), x = Distance, Dx = 1D Diffusion Coefficient (m2/s). The conventional solution of the ADE is provided by Rutherford (1994) in terms of the temporal concentration distribution (C,t) at a known position downstream ($x_{ds}$) is given as:

$$C(x_{ds}, t) = \int_{\gamma=-\infty}^{\infty} \frac{C(x_{us}, \gamma) V_x}{\sqrt{4\pi D_x \underline{T}}} exp\left[-\frac{{V_x}^2 \left(\underline{T} - t + \gamma\right)^2}{4 D_x \underline{T}}\right] d\gamma \tag{6.3}$$

In terms of a spatial distribution (C,x) at a given time ($t_{ds}$), the solution is given as:

$$C(t_{ds}, x) = \int_{\gamma=-\infty}^{\infty} \frac{C(t_{us}, \gamma)}{\sqrt{4\pi D_x \underline{T}}} exp\left[-\frac{\left(x - \gamma - V_x \underline{T}\right)^2}{4 D_x \underline{T}}\right] d\gamma \tag{6.4}$$

Where $\underline{T}$ = mean travel time between measurement positions, (which can be related to mean velocity and distance between measurement positions), and $\gamma$ is an integration variable (in time or space as appropriate).

Assuming the injection can be represented as an instantaneous increase in solute concentration, equation 6.2 was used to produce a prediction of concentration profiles at the sampling location

(outflow point of column as shown in figure 1). An excel goal seek function (built in GRG solver) was used to fit model parameters (Diffusion Coefficient and Travel time) to each experimental test case, utilising a goodness of fit parameter ($R^2$) as the objective function to be minimised . $R^2$ describes the proportion of variance in the dependant variable that can be explained by the independent variable, and so the smaller this value, the better 'fit'. The results and ADE best fit to each experiment are shown in in Figure 62.

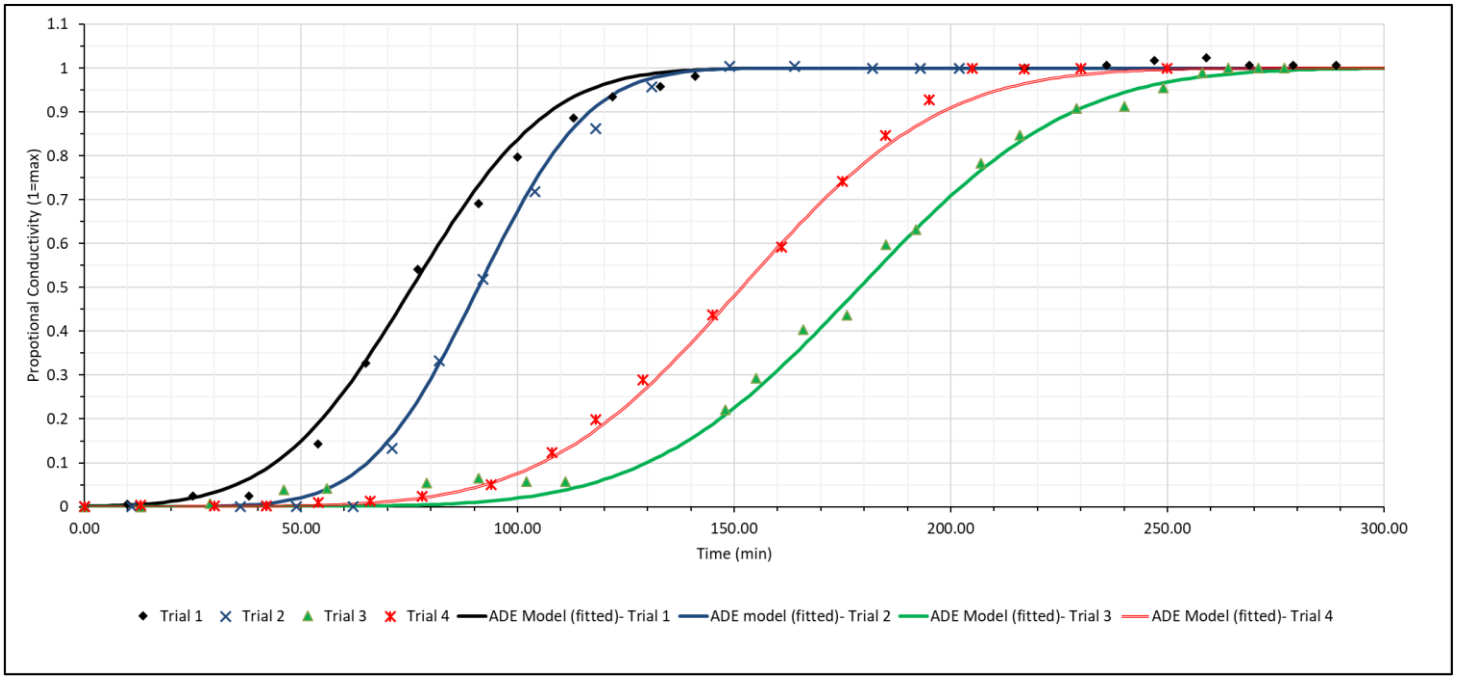


**Figure 62.** Proportional trial conductivity data plotted against a fitted ADE model for each trial.

Not only did this give an estimate of the diffusivity value, but to get the 'best fit' the flow rate was also calculated by the model via the fitted travel time and column characteristics fitting by the equation:

$$Q_f = V_x A_c = \frac{x_{us} - x_{ds}}{\underline{T}} A_c \qquad (6.5)$$

Where $A_c$ = Cross sectional area of the soil column ($m^2$), ($x_{us}$ - $x_{ds}$) is the distance between the water surface and measurement position (m).

Experimental results and other variables are presented in (Table 15).

**Table 15.** Observed parameters from each trial along with values estimated using ADE model.*= The total 'depth' included soil depth, but also the depth of the head of water present on top of soil- if present.

| **Trial** | **Total depth*/**m | **Pre-trial salt solution conductivity/** (μS) | **Flow rate (calculated by ADE model)/**l/min | **Calculated Diffusion Co-efficient/**$m^2$ /s | **$R^2$ value** |
|---|---|---|---|---|---|
| S1 | 1.2 | 1384 | 0.78 | $1.76 \times 10^{-5}$ | 0.997 |
| S2 | 1.05 | 1380 | 0.57 | $5.03 \times 10^{-6}$ | 0.997 |
| S3 | 1.06 | 967 | 0.29 | $2.44 \times 10^{-6}$ | 0.997 |
| S4 | 1.25 | 1914 | 0.41 | $4.93 \times 10^{-6}$ | 0.998 |

As expected, the shape of the experimental curves fit well to the ADE model lines (Figure 66). The diffusion coefficients for trials 2, 3, and 4 are also similar ranging between $2.44 \times 10^{-6}$ and $5.03 \times 10^{-6}$ $m^2$ per second. The diffusion coefficient value for trial 1 is much higher at $1.76 \times 10^{-5}$ $m^2$ per second. It is not surprising that data for Trial 1 specifically, seems different to that of the other trials, as this was the first trial performed- the soil packing technique had not been refined, and it is thought this would have affected soil density and thus travel time and diffusivity.

In terms of flow rate, some variation was seen, as expected between trials, but also compared to the flow rate predicted by the ADE model. Some variation would be expected, however it is thought the difference is mostly due to a leak in the column and blockages due to translocation of finer soil material building up over time. This leak was only discovered after the final salt trial, and so it is thought the recorded flow rates measured volumetrically for the salt trials are not reliable.
With the conductivity data validated by the ADE model, it could now be used to determine sample times (for water) and trial duration for the floodwater trials. Naturally, bacterial cells suspended in floodwater would move through the column differently to a solute (salt), however this was useful as a starting point. To do this, the model from salt trial 3 was used, as the difference between $r^2$ values was negligible, and the recorded flow rate for trial 3 was closest to the flow rate predicted by the model. Based on this it was estimated flood water trials would last for approximately 280 minutes, with it taking up to 50 minutes for the concentration to begin to increase, and then approximately 220 minutes from the concentration starting to increase, to the maximum concentration being reached (Figure 66). As such, it was determined samples should be taken every 30 minutes for the first 60 minutes, and then every hour thereafter. This would allow the increase to be tracked, and allow for the timely processing and analysis of samples.

## 6.4 Flood Water Trials

Three soil column trials were carried out using flood water collected from the field. Each trial was set up in the same way, and conditions for each trial replicated as closely as possible. Some variation, in terms of exact packing of column (soil depth), trial duration and of course contents of the flood water were expected. Trial durations, and column measurements can be seen in table 16.

**Table 16.** Column measurements, including soil, gravel, and water depths, as well as trial duration- sorted by trial.

| **Trial** | **Soil depth/** m | **Gravel depth/** m | **Surface water head depth/**m | **Total depth (soil + gravel + head)**/m | **Trial duration/**h and min |
|---|---|---|---|---|---|
| B1 | 0.31 | 0.13 | 0.13 | 0.57 | 6 h |
| B2 | 0.325 | 0.12 | 0.115 | 0.56 | 5 h 35 min |
| B3 | 0.365 | 0.12 | 0.70 | 0.555 | 5 h 30 min |

For each trial, after saturation of column (and leaving to settle overnight), clean water was pumped into the column, using the peristaltic pump until a head depth of between 7cm and 13cm was reached with the outflow tap open. The head was maintained by a constant inflow of water, and the open port. Once the inflow was equal to the outflow (less the excess from head) the peristaltic pump began pumping the flood water collected from the field site into the column. In terms of 'trial duration', the trial began once the pump switched to pumping flood water.

Water samples were then taken every 30 minutes or one hour ,from the outflow at the bottom of the column. Some trials were shorter in duration than others due to practicalities such as the lab closing for the day, or health and safety risks i.e. working alone out of hours. Water was collected into a sterile 15ml falcon tube, with at least 10ml collected per sample and immediately transferred to a refrigerator (4℃.). Sampling times were decided based on preliminary data taken from the salt tracer experiments (see 'Salt tracer Experiments' section 6.3.5).

Soil samples were taken from the column ~12 hours after the trial ended. This allowed the floodwater to naturally drain from the column, and so soil samples could be taken without disturbing the soil column, or creating 'holes' which would affect the flow path of the water while the trial was running. Soil samples were taken from each trial at the following depths: soil surface (top 10 cm), 40cm (from bottom of column), 33cm, 25cm, 17cm, mesh (10cm sitting on top of mesh). Each soil sample was extracted using the corer (figure 6) and transferred to a sterile 50ml falcon tube and placed into a refrigerator immediately (4℃). Each soil sample weighed at least 10g. For each trial the following pre-trial samples were also taken: clean water (used to saturate soil), floodwater, and a pre-trial soil sample (before being packed into the column).

### 6.4.1 Sample Analysis

#### *6.4.2 Physico-chemical factors and HPC*

pH was measured for Pre-trial soil, and pre-trial clean water samples, immediately after collection as described in Chapter 4. Room temperature and pre-trial floodwater temperature was recorded before contaminated floodwater was added to the column using a portable probe (Sension+ PH1 Portable pH Meter, Field Kit with Electrode for General Purposes, Hach, UK). Triplicate measurements were taken.

HPC CFU counts (Heterotrophic Plate Counts- HPC) for all soil and water samples were performed on LB Agar (Invitrogen™ LB Agar, powder (Lennox L agar) ThermoFisher Scientific, United Kingdom), as described in Chapter 4. Water samples were used to determine bacterial abundance using a spread plate method within less than 2 hours after they were taken from the running column, with soil samples plated as soon as they were taken on day 2 of the soil trial as explained in the 'Flood water trials' section of this chapter.

### *6.4.3 Molecular methods*

Molecular techniques were used to characterise the components of the bacterial communities present in samples that would not be represented by the CFU counts alone (Hosokawa et al., 2022). DNA extraction, quantification and Illumina sequencing of the 16s rRNA gene: including a test PCR and gel electrophoresis, was carried out as described in Chapter 4.

#### *6.4.3.1 Bioinformatics: DNA sequences analysis*

Bioinformatics analysis was carried out as outlined in Chapter 4, with the only exception of sampling depth. For this study the sampling depth was set at minimum - 2000 and maximum- 680000.

### *6.4.4 Diversity Indices*

To fully investigate the bacterial diversity within the samples both alpha and beta diversity metrics were calculated in Qiime2, as described in Chapter 4.

### *6.4.5 Statistical Analysis*

Statistical analysis, including Pearson's correlation coefficient, independent t-test, Principal Coordinate Analysis and PERMANOVA analysis, was carried out as described in Chapter 4.

## 6.5 Results

As explained in the previous section, soil and water samples were taken from the 3 column trials and physicochemical data recorded, along with HPCs). HPC data is available for all soil samples, and for the water samples from trial 3 only (due to the growth of 'lawns' on agar plates for trials 1 and 2 as explained in the subsequent sections of these Results).

In this section, HPCs from soil and water are also compared to the predicted movement of an inert solute based on the ADE model, as described in the previous section. This allows the relative significance of the microbiological factors of the transport processes to be quantified.
Not all samples, once DNA was extracted, were of high enough quality and/or concentration for sequencing (according to the thresholds set by the NEOF sequencing lab). This is due to the small volumes sampled and the success of the DNA extraction process. As such, taxonomic data is available for soil samples from all 3 trials, but not at every depth sampled (Table 17). Taxonomic data for water samples is available for trial 3 only.

**Table 17.** Table of soil samples from all floodwater trials in relation to taxonomic data availability.

| Trial | Depth soil sampled (distance from surface) | Taxonomic Data Availability |
|---|---|---|
| 1 | Pre-Trial | No |
| | 0.13m | No |
| | 0.17m | Yes |
| | 0.24m | Yes |
| | 0.32m | Yes |
| | 0.44m | Yes |
| | 0.57m | Yes |
| 2 | Pre-Trial | Yes |
| | 0.12m | Yes |
| | 0.14m | No |
| | 0.23m | No |
| | 0.31m | Yes |
| | 0.39m | Yes |
| | 0.56m | Yes |
| 3 | Pre-Trial | Yes |
| | 0.07m | No |
| | 0.15m | Yes |
| | 0.22m | Yes |
| | 0.30m | Yes |
| | 0.38m | Yes |
| | 0.55m | Yes |

6.5.1 pH and temperature

pH of the soil, clean water, and floodwater was taken before each trial, along with flood water and room temperatures. Flood water and room temperature was the same for all 3 trials, due to the lab being temperature controlled (set at 25℃. ) Flood water pH showed little variation- with pH for each trial being 7.59, 7.56, and 7.57 for trials 1-3 . Clean water pH did vary slightly at 8.02, 8.67, and 8.18 (for trials 1-3 respectively). Soil pH was lowest for trial 1 at 7.81, and highest for Trial 3 at 8.67. Trial 2 soil had a pH of 8.28.

6.5.2 Column Soil characterisation: Soil particle size and class/type

After soil sizing using the sieve method, and classification according to British Standard 1377 Part 2 it was found that the soil used during the column experiments was a sandy loam soil (Figure 63), with sand constituting ~70% of the total material, ~20% being silt, and very low levels of clay and gravel.

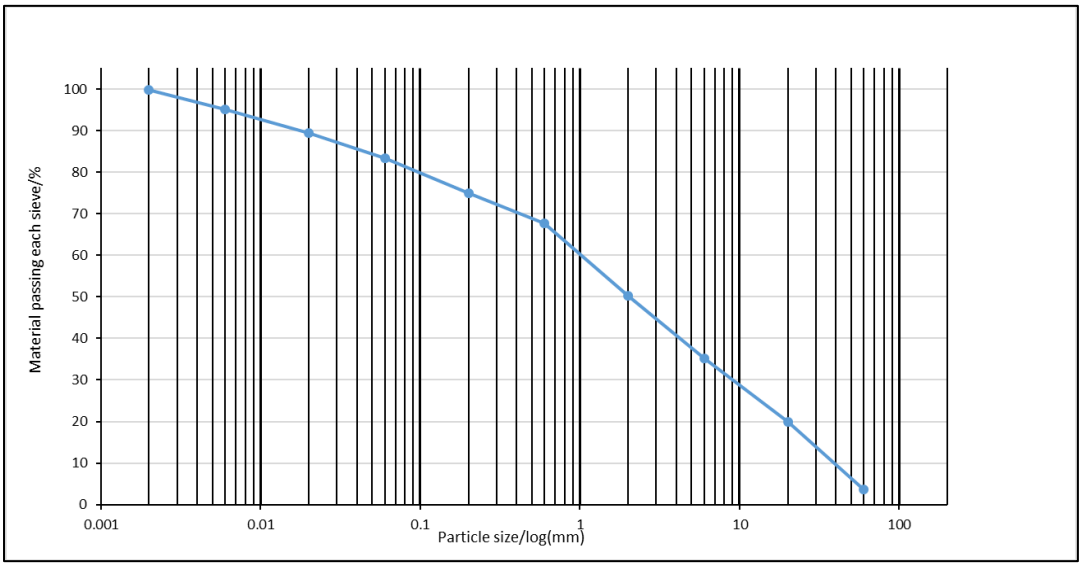


**Figure 63.**Soil gradation curve of column soil.

6.5.3 Bacterial characterisation of Pre-trial water

HPC data and taxonomic analysis is only available for water samples from Trial 3, and so the pre-trial water data is available and is discussed for this trial only. The pre-trial floodwater had a HPC of $3x10^{11}$ CFU/ml. Relative abundance of different phyla present in pre-trial floodwater was established based on assigned ASVs. 21 phyla were present in pre-trial floodwater, with the most abundant being Proteobacteria at 33% (Figure 64). Other abundant phyla included: Bacteroidota and Firmicutes at >20%. The 3 aforementioned phyla made up the majority – 77.64% - of the bacterial community present in the pre-trial floodwater sample. Other notable phyla were Bdellovibrionota at 4.03%, Campylobacterota and Fibrobacterota at >2%. The rest of the community was made up of 15 low abundance phyla- present at abundance levels of 0.01-1.3%.

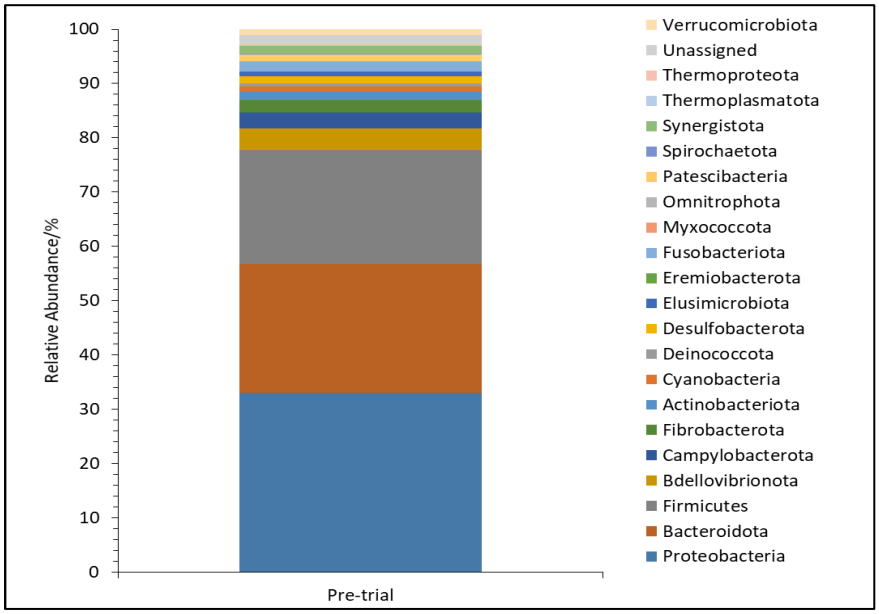


**Figure 64.** Bacterial communities present in pre-trial water sample at Phylum level (n=1). Unassigned taxa grouped into 'Unassigned' category.

At genera level, 25 genera were present, with *Acidovorax* the most abundant at >7%. Other abundant genera included: *Flavobacterium, Pseudomona*s, *Acinetobacter* and *Prevotella*, all at relative abundance >3.5%. Combined, these genera made up over half of the total bacterial community assigned to genera (46% of total bacterial community assigned to genera). Other genera were present at lower abundance, between 0.8 and 1.6% relative abundance. Several genera were identified that are known to be associated with human faeces, or to be pathogenic, including : *Bifidobacterium* , *Ruminococcus*, and *Phocaeicola* at >0.8%.

At species level, the bacterial community was composed of 30 relatively low abundance species, all present at <1.8% respectively. The most abundant species was, *Prevotella copri* at 1.78%. At this level, faecal-associated species were present: *Ruminococcus bromii* and *Akkermansia muciniphila* at >0.5%. As well as this, 2 potentially pathogenic species were detected: *Dialister invisus* and *Phocaeicola vulgatus* at >0.5%.

6.5.4 Bacterial characterisation of Pre-trial Soil

HPC data is available for pre-trial soil samples from all 3 floodwater trials. Taxonomic data is available for trials 2 and 3 only. The number of HPC CFUs in pre-trial soil samples was variable with $3x10^8$ CFU/g in trial 1, none detected in trial 2, and $7x10^9$ CFU/g in trial 3. Pre-trial soil samples from trials 2 and 3 shared 20 of the same phyla (Figure 65). Trial 2 had 23 phyla present in the pre-trial sample in total, whereas trial 3 had 21. Both trial samples had Firmicutes, Chloroflexota, Actinobacteriota, Proteobacteria, and Planctomycetota as the 5 most abundant phyla. However, there were variations in the most abundant phylum among trials, for example the most abundant phylum for trial 2 pre-soil was Chloroflexota at 17.64%, whereas for trial 3 it was Firmicutes at 28.41%. The samples from both trials had similar levels of Bacteroidota (> 3%).

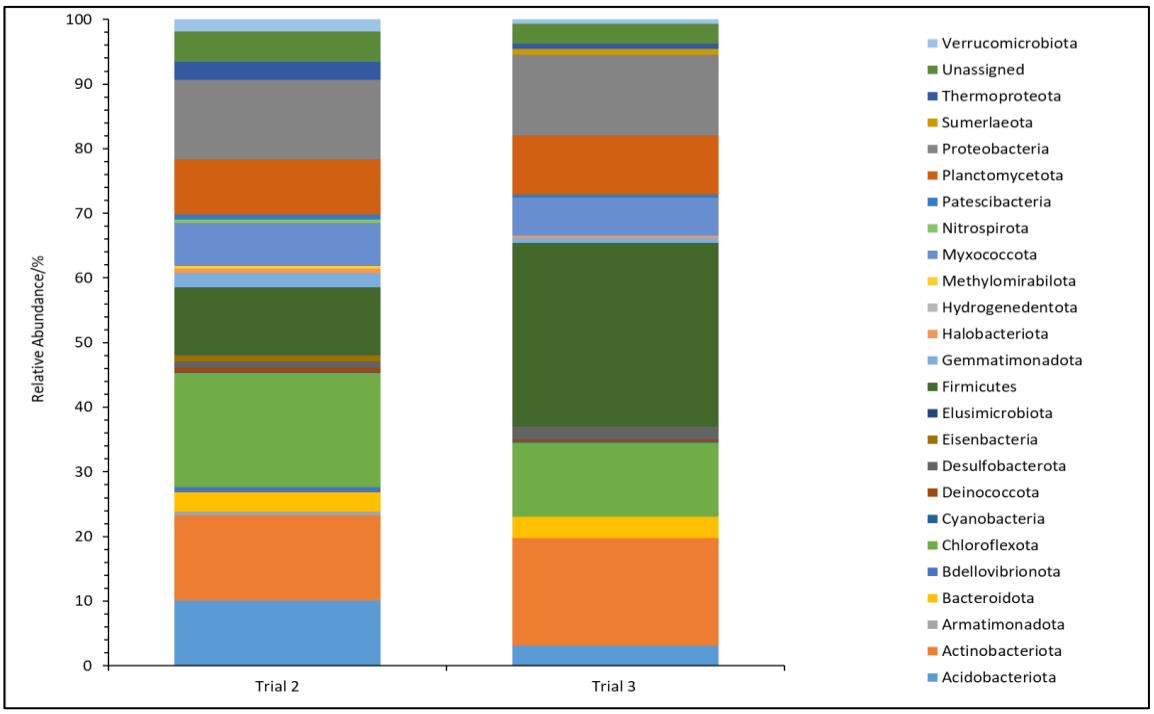


**Figure 65**. Bacterial communities present in pre-trial soil samples from trials 2 and 3 at Phylum level (n=2). Unassigned taxa grouped into 'Unassigned' category.

At genera level, samples from both of the trials shared 9 genera, including *Clostridium, Haliangium*, and *Caldicoprobacter*- all present at similar abundance levels in both samples. This left just 3 genera unique to the trial 2 sample, and 21 that feature in the trial 3 sample only. The most abundant genera in the trial 2 sample were *Povalibacter* , *Yinghuangia* , and *Streptomyces* at >1%. *Povalibacter* and *Yinghuangia* were at much lower abundance in the trial 3 sample, at 0.54% and 0.88% respectively. The most abundant genus in the trial 3 sample was *Planifilum* at >4%- much higher than the 0.89% abundance in the trial 2 sample. Other abundant genera in the trial 3 sample included: *Thermaerobacter* and *Ureibacillus* at ~2%.

At species level, samples from both trials were composed of many low abundance species- with abundance levels being <1.4% per species. The samples from both trials shared 10 species including: *Haliangium ochraceum* and *Clostridium thermopalmarium*. No faecal associated or pathogenic species were detected in either of the pre-trial soil samples.
When alpha diversity metrics were calculated for the pre-trial soil samples, ACE species richness and Simpson's Evenness values were higher for trial 3 (189 and 0.82), compared to that for trial 2 (174 and 0.71). However, Faith's Diversity index was actually higher for trial 2 (25.55), compared to that of the trial 3 sample (24.30).

6.5.5 Bacterial community analysis of column water samples, including transport times

After establishing pre-trial bacterial levels, it can now be determined from measured HPCs and taxonomic analysis, if any changes occurred in the bacterial community as it passed through the soil column. Water sample data (HPC and taxonomic data) is only available for trial 3. HPCs were measured for the water samples collected at various time points for trials 1 and 2 but all plates grew a bacterial 'lawn' and thus no changes in the HPCs could be quantified. DNA analysis was not possible for trials 1 and 2 as explained at the start of this 'Results' section.

6.5.5.1 Water Heterotrophic Plate counts and temporal distribution

To compare bacterial transport within the water through the column, to the anticipated transport behaviour of an inert solute , HPCs for water from Floodwater trial 3 were plotted against the ADE model curve (Figure 66). Parameters for the curve were: diffusion co-efficient from salt trial 3 (diffusion coefficient=$2.44 \times 10^{-6,}$ - used as this trial was thought to be least affected by the leak found in the column during salt water trials), travel distance from floodwater trial 3 (0.55m), and flow rate from floodwater trial 3 (0.18 litres per minute). The ADE curve was calculated using equation 6.4 (see methods section for equation). HPCs measured ranged from $1 \times 10^{9}$ CFU/ml at 30 minutes, and reached $2.3 \times 10^{10}$ CFU/ml at 270 minutes. None of the measured water samples had a HPC higher than that found in the pre-trial water sample ($3 \times 10^{11}$ CFU/ml), thus it cannot be said that the HPC had reached it's 'maximum' as the trial did not run for long enough for the number of HPC CFUs to plateau- indicating maximum had been reached. However, the sudden rise in CFU count after approximately 210 min provides comparative data regarding the 'arrival time' of contaminated floodwater at the bottom of the soil column.

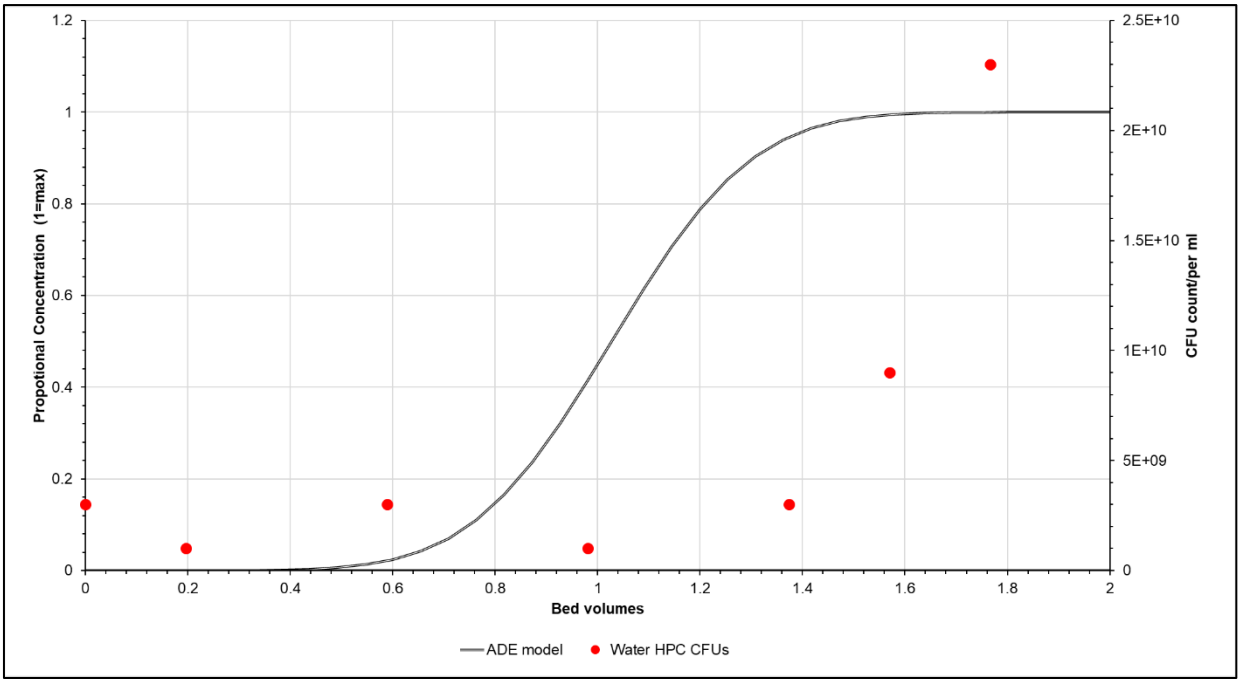


**Figure 66**. Proportional conductivity for ADE model based on salt conductivity trials, plotted against HPC CFU/ml for water samples from Floodwater Trial 3 (n=7 CFU count plotted on secondary axis.

For Figure 66, dimensionless bed volume was used to compare bacterial transport time to that calculated by the model for a solute. This was done by dividing measured time in seconds by hydraulic residence time (time for one bed volume to flow through the column) in seconds. From Figure 66, it is clear that there is a difference between the bacterial transport time, and that calculated by the model for a solute. The results indicate that the inert solute would move through the column more quickly than bacteria as the ADE predicted solute concentration begins to increase after around 25 minutes. However, the bacteria take longer to begin to increase (clear increase only seen after 210 minutes, lag of 185 minutes compared to the solute) thus moving more slowly through the column. This is supported by an estimation of the retardation factor (estimate only as full breakthrough curve is not available), which in this case is 1.55 due to the adsorption of the bacteria, thus meaning bacteria move through the soil more slowly than the solute.

6.5.5.2 Taxonomic classification

Relative abundance of different phyla present in floodwater samples taken from the outflow point of the column at different times was established based on assigned ASVs. The composition of each sample at phylum level, compared with the pre-trial floodwater sample is seen in Figure 67.

Proteobacteria was the most abundant phylum present in pre-trial floodwater and in all of the water samples taken , present at 35- 38% abundance at all time points. 3 phyla were found in the pre-trial flood water sample but not in any other water samples. These phyla were: Thermoplasmatota, Eremiobacterota, Spirochaetota , present between 0.1-0.2%. These phyla were not present in any soil samples from the same trial.

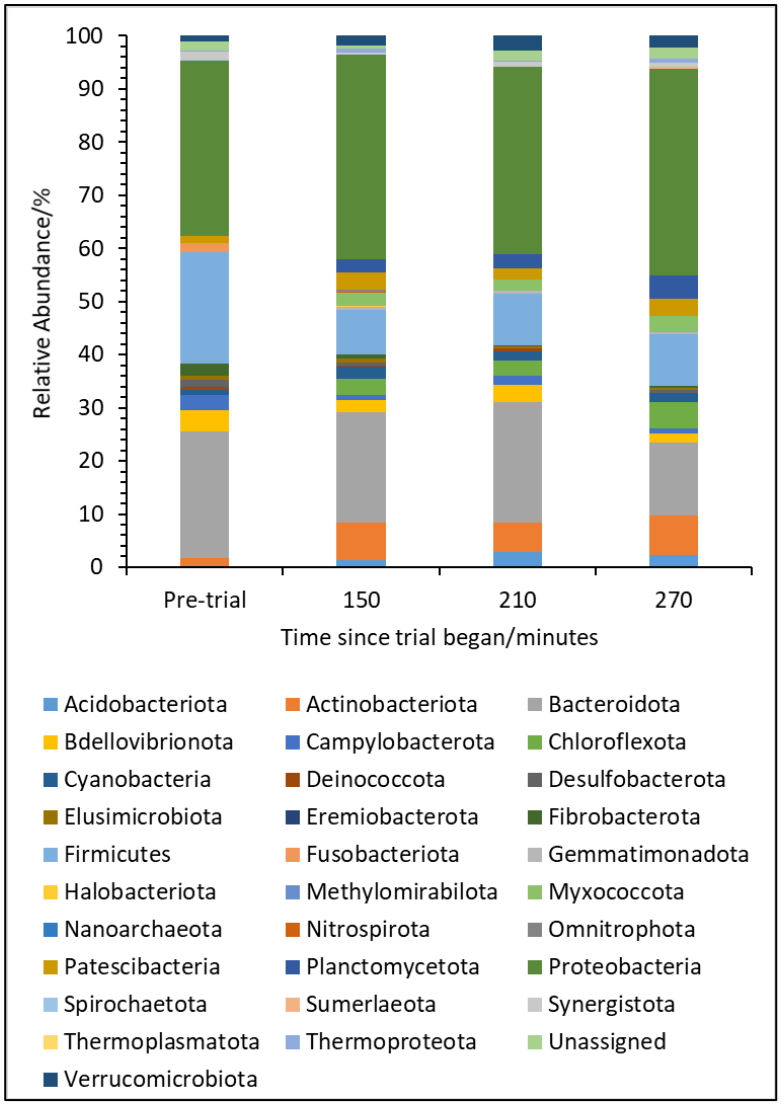


**Figure 67**. Comparison of bacterial communities present in water samples from floodwater trial 3 at phylum level (n=4). Unassigned taxa grouped into 'Unassigned' category.

9 phyla were detected in water samples taken from the column that were not found in the pre-trial floodwater sample. Of these 9 phyla, 3 increased in abundance over time. These phyla include Chloroflexi, Acidobacteriota, and Gemmatimonadota . Compared to the pre-trial sample, Firmicutes abundance was notably reduced, but increased over time in water samples from 8.35% at 150 minutes to 9.61% at 270 minutes. Both Bacteroidota (13.62%-22.63%) and Fibrobacterota (0.24%-0.84%) decreased in abundance in water samples. Myxococcota abundance was higher in the water samples at all time points (1.99%-3.09%) compared to the pre-trial sample (0.05%).

At genus level, 8 genera found in the pre-trial sample were not detected in any of the subsequent water samples (Figure 68). Some of these genera included: *Paracoccus* , *Ruminococcus* ,and *Clostridium* all present at ~1% abundance. The most abundant genus found in pre-trial water-*Acidovorax* (7.48%)- was not present in any of the water samples taken.

For water samples, the number of genera present increased over time, from 31 at 150 minutes to 36 at 270 minutes. *Flavobacterium* was abundant in water samples- present at an abundance higher than that of the pre-trial sample (5.05%). *Flavobacterium* in the water samples decreased in abundance over time from >7% at 150 minutes to <4% at 270 minutes. 12 genera found in water samples were also detected in soil from this trial, including genera such as: *Mycobacterium, Clostridium*, and *Sphingomicrobium*.

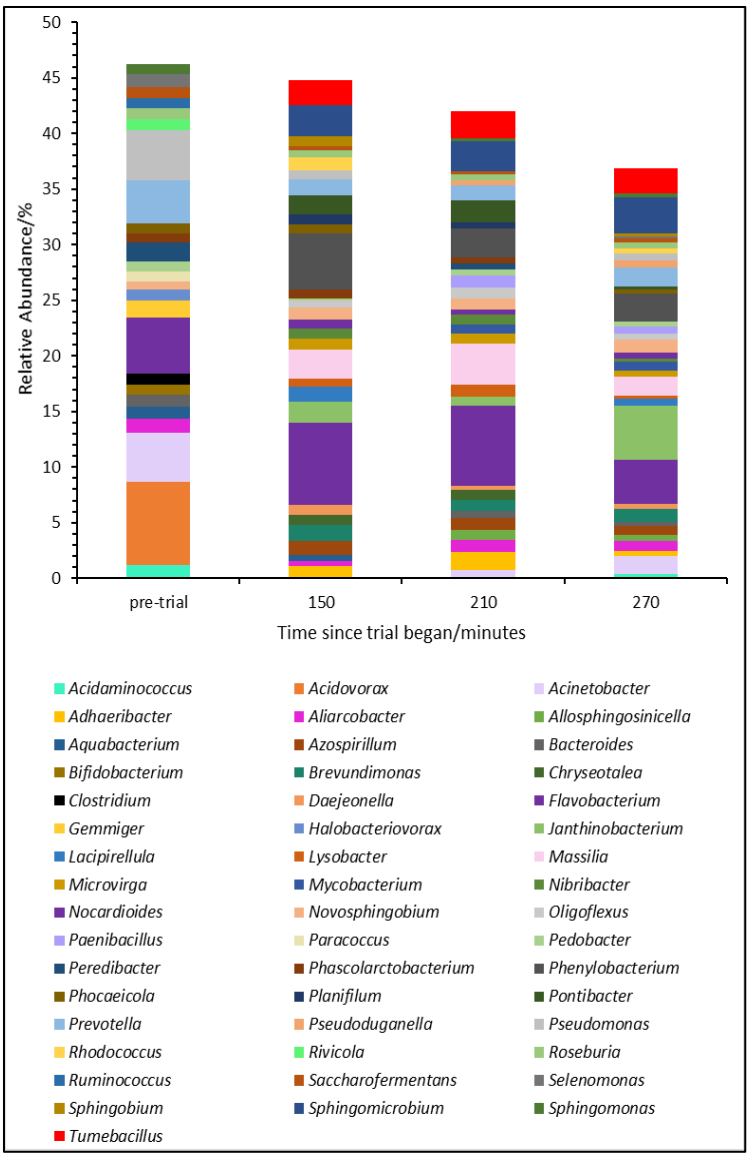


**Figure 68.** Comparison of bacterial communities present in water samples at Genus level (n=4). Filtered to include only genera with a relative abundance of >0.5%, in at least one of the samples taken .

At species level, samples were filtered by removing sequences that were not assigned to any genus and with a relative abundance of <0.5% for at least one time point sampled. Overall 12 species found in the pre-trial water sample, were also detected in other subsequent water samples. The number of species present in the water samples increased over time: starting at 27 at 150 minutes and increasing to 33 species at 270 minutes. The abundance of different species was variable over time.

Several species in water samples can be associated with faeces or potential waterborne diseases. These species included: *Akkermansia muciniphila, Phocaeicola vulgatus*, and *Dialister invisus*, present at <1%. *A. muciniphila* was found in the pre-trial water sample only, and, it was not found in any other water or soil sample from this trial. Other species were found in pre-trial and subsequent samples for example: *P. vulgatus* was present in the pre-trial sample (0.94%) and in water at 150 minutes. Similarly, *D. invisus* was found in the pre-trial water sample (0.61%), and at 270 minutes (0.18%), however, was not detected in any soil samples.

6.5.5.3 Comparison of solute transport to specific species

Bacterial transport (assessed as HPCs) has been compared to that of an inert solute (salt) earlier in this section. Here the inert solute was compared using sequencing results and taxonomic assignments. In order to accurately track the transport, two species were chosen to be plotted on Figure 69, that were present in water samples taken at all 3 of the sample times (not present in any soil samples). The relative abundance (%) of each species was plotted against time. These species were*, Flavobacterium album,* and *Prevotella copri.*

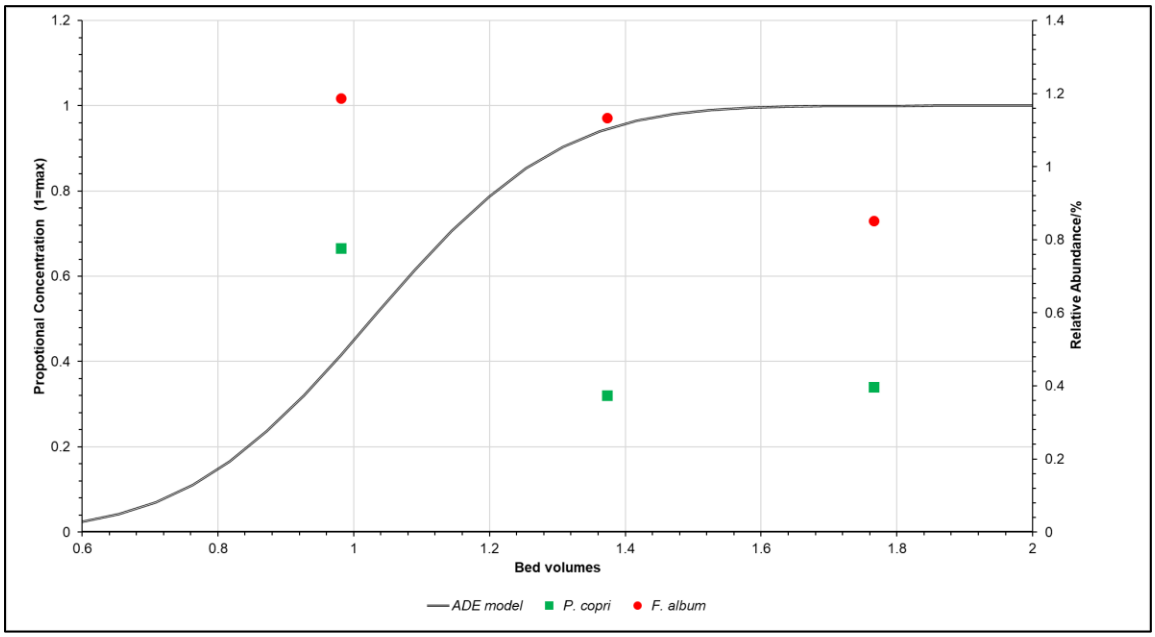


**Figure 69.** Proportional conductivity for ADE model based on salt conductivity trials, plotted against relative abundance of 2 species from water samples from Floodwater Trial 3 (n=3). Relative abundance plotted on secondary axis.

For Figure 69, dimensionless bed volume was used to compare bacterial transport time of two species to that calculated by the model for a solute. This was done by dividing measured time in seconds by hydraulic residence time (time for one bed volume to flow through the column) in seconds. It is clear that the transport of the two species is different to that of a solute (as predicted by the ADE model). Unlike with CFU count (Figure70), there is no clear arrival time or trend in abundance exhibited by *P. copri* or *F. album* over time.

6.5.5.4 Alpha Diversity Metrics

To determine alpha diversity, 3 unique measures were calculated at using all ASV's, for all water samples (Table 18).

**Table 18.** Alpha diversity indices for water samples taken from Trial 3 (n=4).

| **Time water sample was taken/**min | **ACE Species richness** | **Faith's Diversity Index** | **Simpson's Evenness** |
|---|---|---|---|
| Pre-trial | 200 | 24.23 | 0.75 |
| 150 | 202 | 27.83 | 0.73 |
| 210 | 235 | 28.67 | 0.74 |
| 270 | 259 | 30.73 | 0.77 |

Species richness and Faith's diversity metrics were higher in all water samples compared to the pre-trial sample (200 and 24.23). Not only this, but species richness and Faith's diversity increased in water samples over time. This infers that by infiltrating through the soil column, alpha diversity of the bacterial community in the floodwater increases. Despite this, Simpson's evenness showed a similar trend of increasing over time- starting at 0.73, and increasing to 0.77. However, Simpson's evenness was only higher than that of the pre-trial sample (0.75), at 270 minutes (0.77).

### 6.5.6 Bacterial community analysis of column soil samples, including spatial distribution.

After establishing pre-trial bacterial levels, HPC and taxonomic information were used to determine bacterial changes with depth. This section will analyse this on a 'trial by trial basis', beginning with Trial 3 (the trial in which water data was available for).

#### 6.5.6.1 Trial 3

Soil HPCs and spatial distribution

To compare the spatial distribution of HPCs in the soil vertically through the column to the predicted behaviour of an inert solute, HPC data from the floodwater trial, from varying depths down the column was plotted against solute concentration based on the ADE model (Figure 70). The ADE curve was calculated as outlined by equation 6.5 (i.e. spatial distribution). Inputs for the ADE model were: diffusion coefficient from salt trial 3 (diffusion coefficient=2.44 × 10-6), and then flow rate (~0.18 L/per min) and travel distance for the floodwater trial (Trial 3= 0.55m).

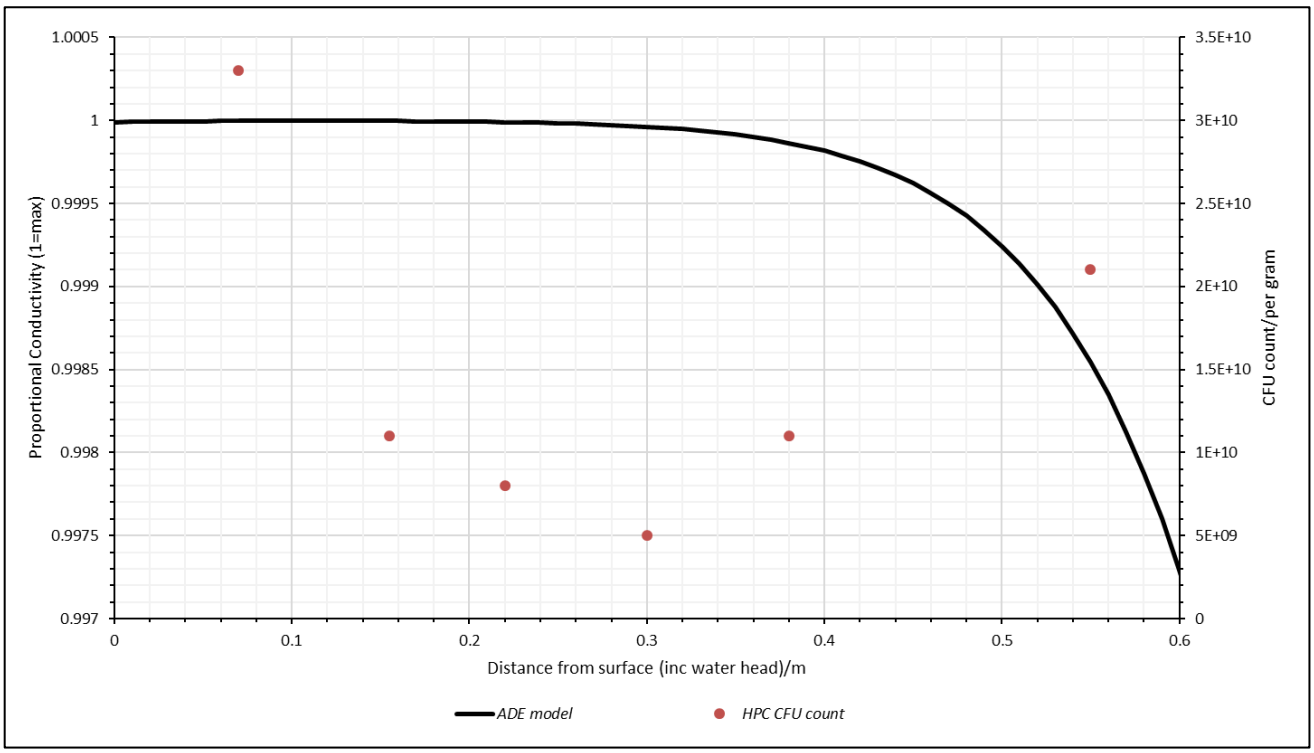


**Figure 70**. Proportional conductivity for spatial ADE model based on salt conductivity trials, plotted against number of HPCs per gram for soil samples from Floodwater Trial 3 (n=6). HPC data plotted on secondary axis.

For Trial 3, the highest plate count (3.3 x $10^{10}$) was seen at just 0.07m from the column surface. The lowest count (5 x$10^{9}$) was seen at 0.3m from the column surface. HPCs in the soil decreased from 0.07m to 0.3m from the surface, and then increased again at 0.38m and 0.55m from the surface. The HPCs follow a different trend to that of the inert solute. With the model estimating maximum concentration (1) of the solute up to approximately 0.2m from the surface, with the concentration then slightly decreasing to 0.997 at the bottom of the column (0.6m from column surface).

<u>Taxonomic classification</u>

Relative abundance of different phyla in soil samples taken from different depths down the column was established based on assigned ASVs. The composition of each sample at phylum level, compared with the pre-trial soil sample, and the pre-trial floodwater sample is seen in Figure 71.

Samples from all depths were used for taxonomic analysis, apart from the sample from 0.15m, for Trial 3. The most abundant phyla in most of the samples with soil depth (from 0.22 to 0.55) was Firmicutes representing >28% of the total bacterial community in the pre-trial soil sample (Figure 71). Firmicutes was not the most abundant phyla in the pre-trial water sample but it did have a high abundance at 20.8%. 7 phyla were present in soil samples that were not in pre-trial soil. These phyla included: Bdellovibrionota, Chlamydiota, and Methanobacteriota. As well as this, 2 phyla were present at 0.55m only: Methanobacteriota and Campylobacterota at <0.1%. Soil samples from this

trial (excluding pre-trial sample) and flood water samples shared 21 phlya, but only 9 were exclusively present in flood water, including: Elusimicrobiota, Spirochaetota, and Fibrobacterota .

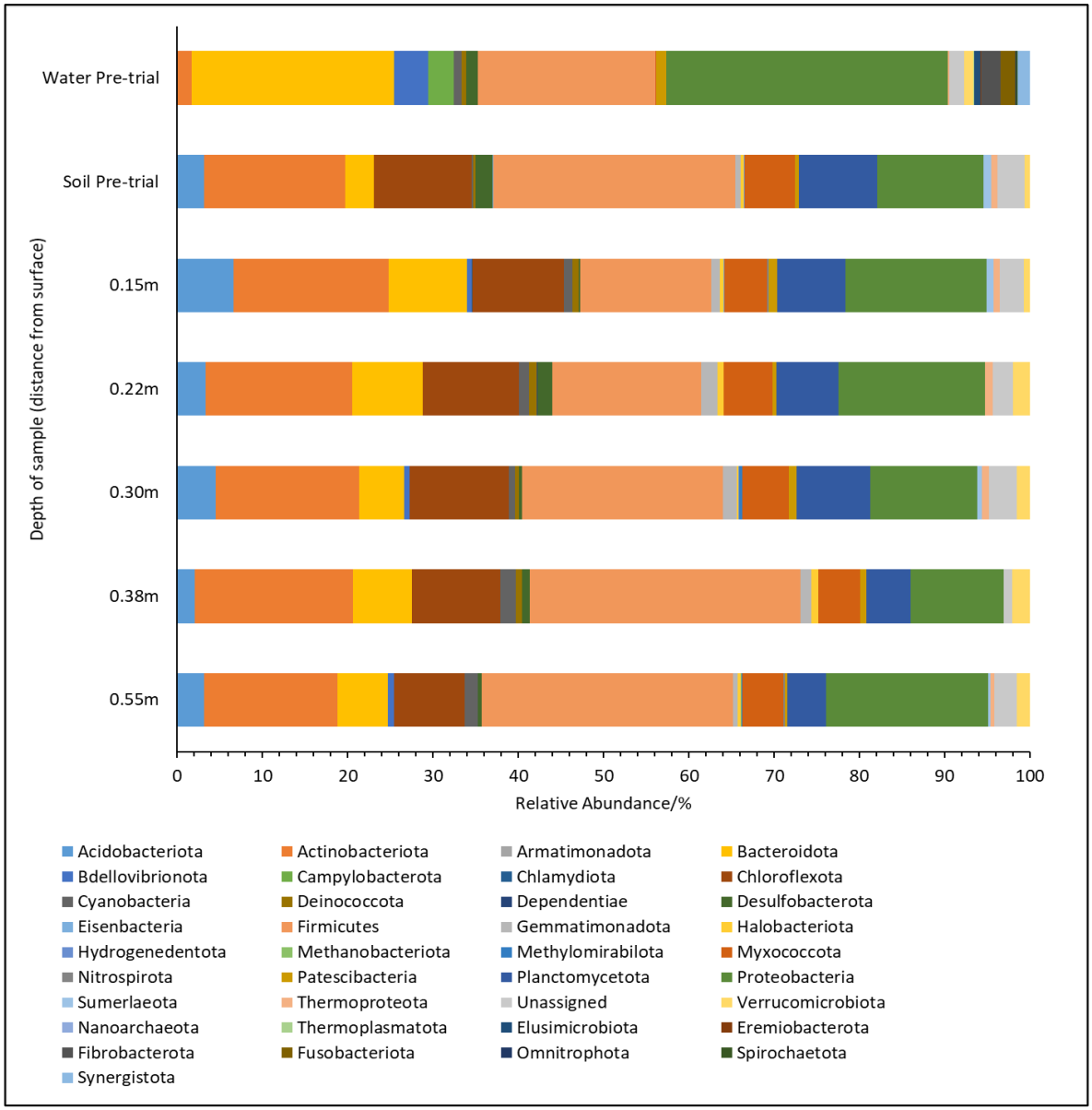


**Figure 71**. Comparison of bacterial communities present in Soil samples at different depths at phylum level (n=7). Unassigned taxa grouped into 'Unassigned' category.

In general, genera found in pre-trial soil were found in soil at different depths, however, 3 genera were found at all soil depths, but not in the pre-trial sample (Figure 72). These were: *Adhaeribacter, Pontibacter*, and *Flavobacterium.* The number of genera found in soil samples increased with soil depth, with the lowest number of genera being 27 at 0.15m, increasing to 30 at 0.55m. *Clostridium* was prominent in soil samples from this trial, present at all soil depths, and being most abundant (3%) at 0.55m. *Clostridium* was also present at lower levels in the pre-trial soil (0.8%) and pre-trial water sample (1%).

When comparing the pre-trial water sample to the soil samples, *Acidovorax* (most abundant genera in pre-trial water) did not appear in any of the soil samples, indictaing that this genus may not have transferred out of the floodwater and into the soil. In fact, only 3 genera from the pre-trial water sample, appeared in the soil samples- *Clostridium, Flavobacterium* and *Novosphingobium*. *Flavobacterium* had 5% abundance in the pre-trial water sample and appeared at all soil depths (not in soil pre-trial sample)- at lower levels, ranging from 1.9% at 0.15m, to 0.28% at 0.30m. Other genera such as *Novosphingobium* had 0.77% abundance in the pre-trial water sample (not present in pre-trial soil), and was present in soil samples at 0.22m (0.45%), and at 0.55m (1.15%).

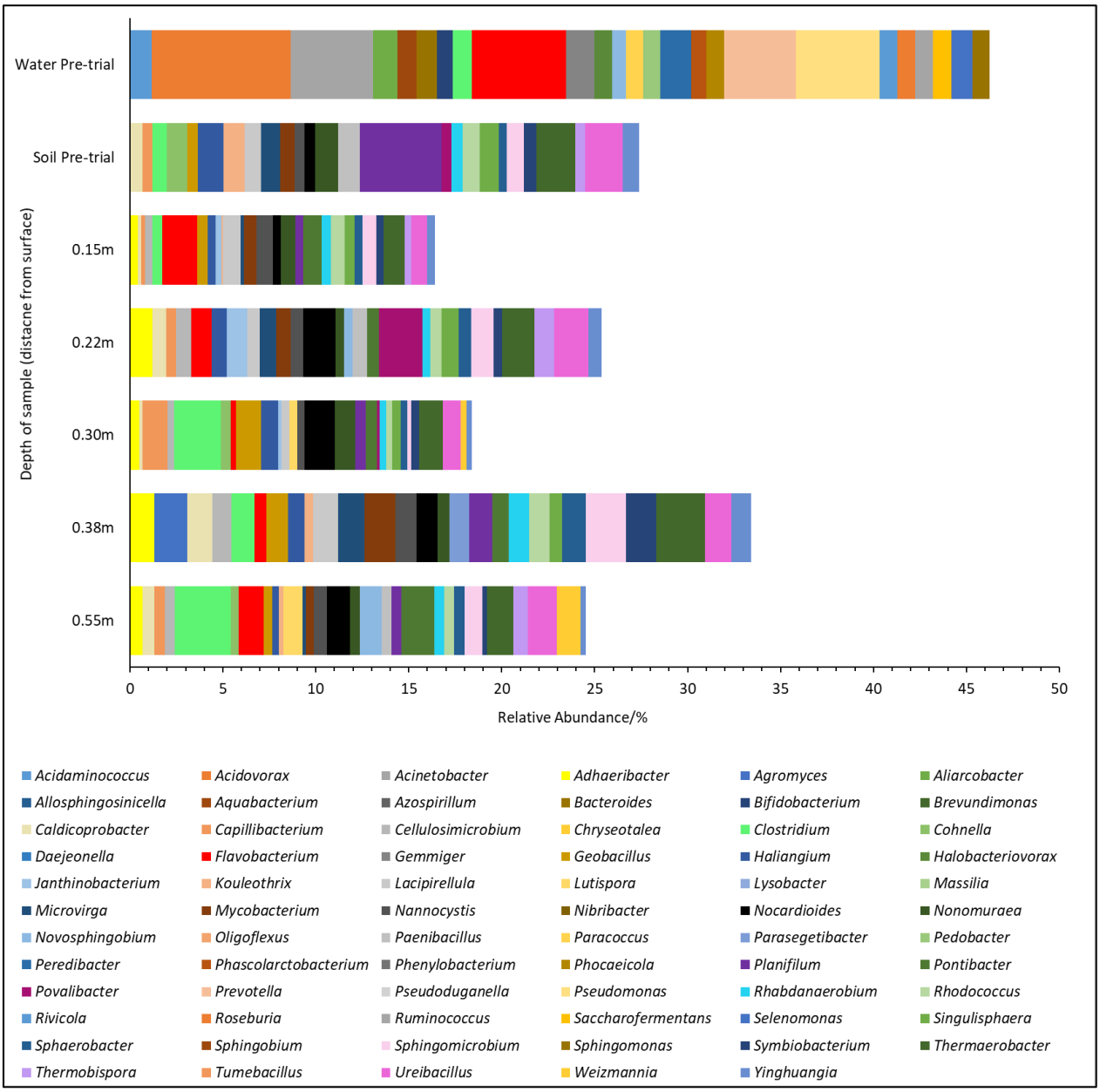


**Figure 72.** Comparison of bacterial communities present in soil samples at different depths at Genus level (n=7). Sequences that were not assigned to a genus were removed, along with any sequences that had a relative abundance of <0.5%.

At species level, the number of species present at each depth was variable with no clear trend of increase/decrease over the different depths. 4 species detected in water samples were also found in soil samples, however, no pathogenic or faecal related species were detected in any soil samples.

Comparison of solute spatial distribution to specific soil species

This section compares the taxonomic assignments made from DNA sequencing to the inert solute. In order to accurately assess the distribution of the biological component, two species were chosen to be plotted on Figure 73, that were present in soil at multiple depths in all three floodwater trials (not found in pre-trial soil). The relative abundance (%) of each species was plotted against time. These species were, *Caulifigura coniformis,* and *Methanosarcina thermophila*. Based on this, and the fact the species are not present in background soil samples, it is assumed that *C. coniformis* and *M. thermophila*, both originated from the floodwater in this study.

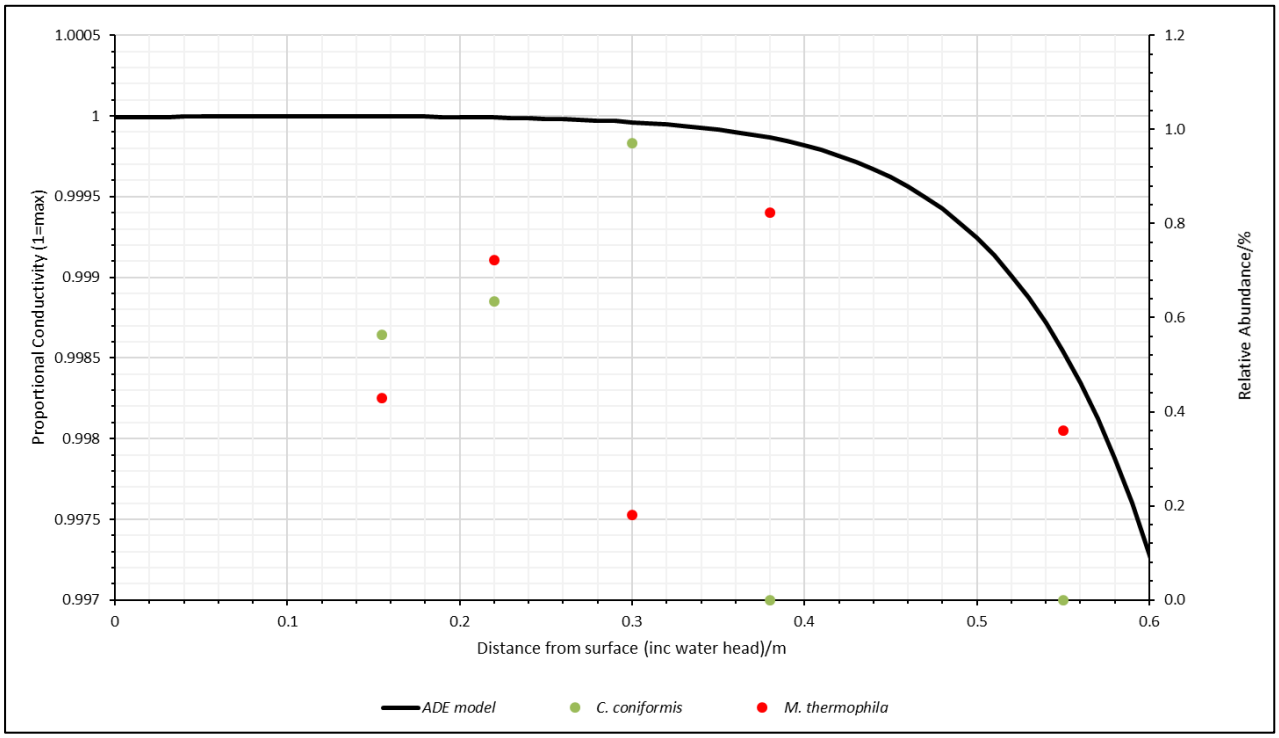


**Figure 73.** Proportional conductivity for spatial ADE model based on salt conductivity trials, plotted against relative abundance of 2 species from soil samples from Floodwater Trial 3. Relative Abundance plotted on secondary axis.

*C. coniformis* and *H. ochraeum* showed similar abundance trends with depth. Both species saw abundance increasing as distance from the surface increased, peaking at 0.3m from the soil surface (0.97% for *C. coniformis*, and 0.9% for *H. ochraeum*). After 0.3m *C. coniformis* was not present at any other depths, however *H. ochraeum* was present and decreased over the last 0.3m of the column. This indicates that *C. conformis* bacteria may have transferred from floodwater to soil towards the column surface- removed from the floodwater and thus is not present in soil samples after 0.3m. Whereas *M. thermophila*, although some transfer may have happened appears to have travelled further into the soil.

A difference in spatial distribution was seen between the species and the inert solute. The solute was at maximum concentration up until around 0.3m, where it began to steadily decrease. However, this depth is where 1 of the 2 species (*C. coniformis*) actually reached the highest (recorded) abundance. All species abundance and the solute concentration seemed to decrease after this depth.

Alpha Diversity Metrics

To determine alpha diversity, 3 unique measures were calculated at species level using all ASV's, for all soil samples (Table 19).

**Table 19.** Alpha diversity indices for soil samples taken from Floodwater trial 3 (n=6).

| Trial | Depth of sample (distance from surface) | ACE Species Richness | Faith's Diversity Index | Simpson's Evenness |
|---|---|---|---|---|
| Trial 3 | Soil Pre-trial | 189 | 24.30 | 0.82 |
| | 0.15m | 339 | 38.26 | 0.71 |
| | 0.22m | 185 | 25.56 | 0.74 |
| | 0.30m | 349 | 35.84 | 0.74 |
| | 0.38m | 111 | 17.53 | 0.76 |
| | 0.55m | 274 | 30.96 | 0.63 |

For trial 3, Species richness and Faith's diversity were highest at 0.15m (339 and 38.26), and lowest at 0.38m (111 and 17.53). However, Simpson's Evenness was highest at 0.38m (0.76) and lowest at 0.55m (0.63). For the pre-trial sample, the species richness (189) was higher than that of the samples taken at 0.38m (111) and 0.22m (185), but lower than other samples. Faith's diversity for the pre-trial sample (24.30) was lower than the value for all samples apart from at 0.38m (17.53). Simpson's evenness was higher in the pre-trial sample at 0.83 than in any of the other samples.

Beta Diversity

PCoA analysis could not be performed on beta diversity to compare pre-trial samples to the samples taken during the trials due to the limited dataset as explained previously. PCoA analysis was performed for Trial 3 - the only trial with both water and soil sequencing data available. Figure 74 shows a clear separation between the soil and water samples for trial 3. PERMANOVA analysis proved this difference was significant- $p < 0.01$. Water samples appear to cluster more tightly than soil samples. This indicates a wider range in diversity values over the sampling depths for soil, whereas the water samples, sampled only 60 minutes apart, could have more similar values. From the figure the water pre-trial sample can also be seen to be clearly separate to the other water samples, whereas the pre-trial soil sample's separation is not as clear.

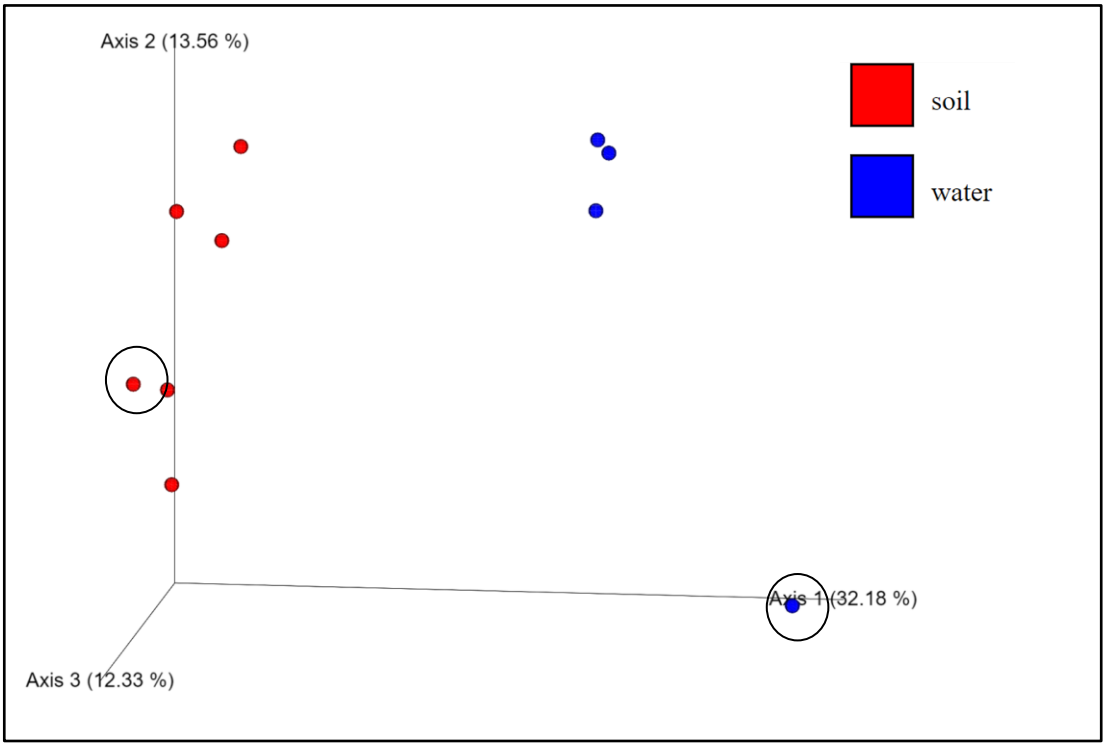


**Figure 74.** PCoA (Principle Coordinate Analysis) Emperor plot based on weighted unifrac distance analysis of beta diversity values from Trial 3. n= 10. Pre-trial samples circled in black.

### 6.5.6.2 Trial 2

Soil HPC and spatial distribution

To compare the spatial distribution of bacteria in the soil vertically through the column to that of an inert solute, HPC CFUs collected from the floodwater trial, from varying depths were plotted against predicted solute concentration based on the ADE model (Figure 75). The ADE curve was calculated as outlined by equation 6.5. Inputs for the ADE model were: diffusion coefficient from salt trial 3 (diffusion coefficient=2.44 × 10-6), and then flow rate (~0.18 litres per minute) and travel distance for the floodwater trial (Trial 2= 0.56m).

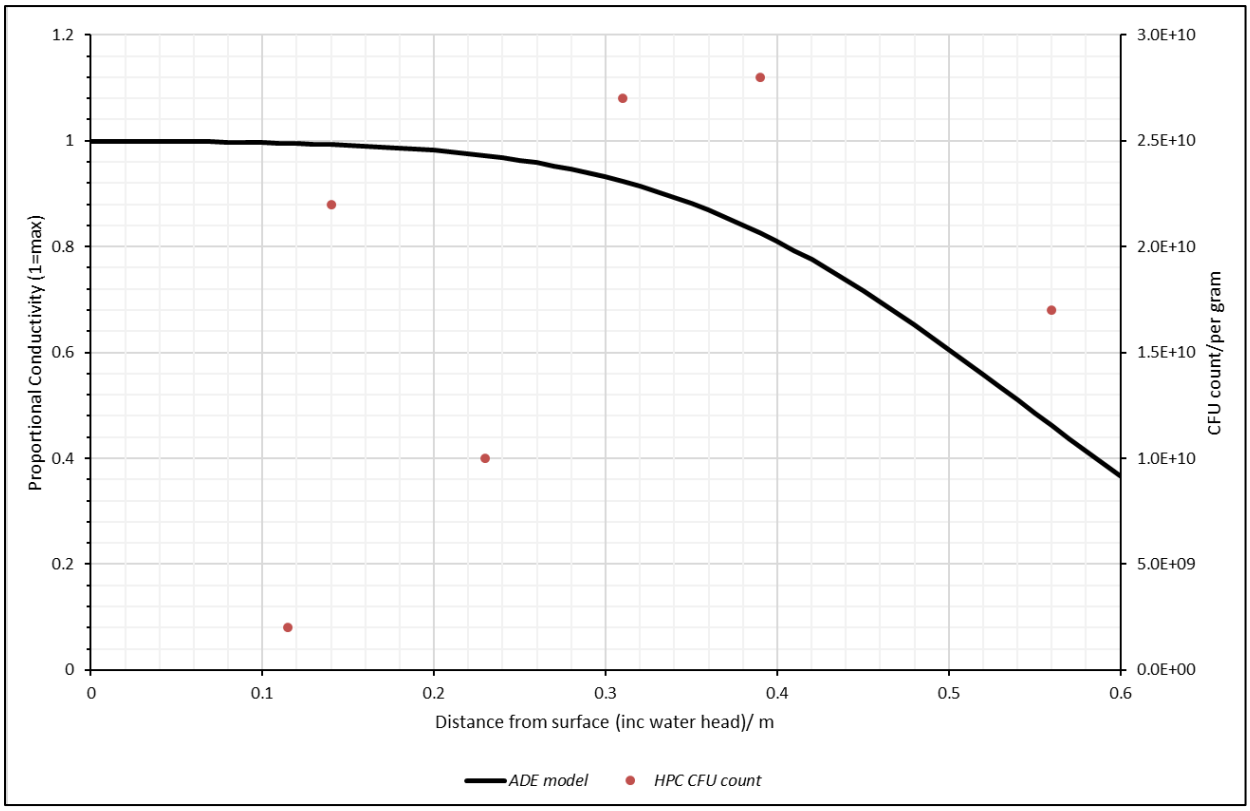


**Figure 75**. Proportional conductivity for spatial ADE model based on salt conductivity trials, plotted against HPC CFU count per gram for soil samples from Floodwater Trial 2 (n=6). CFU data plotted on secondary axis.

HPCs for soil samples from trial 2 were variable over the depths, with the highest number of HPCs - $2.8x10^{10}$ CFU/g at 0.31m from the column surface. All counts recorded over the soil depths were higher than the pre-trial soil HPCs count ($2x10^{9}$ CFU/g).

The bacteria (based on HPC CFU) in the soil follows a different trend to that which would be expected from an inert solute (ADE model). The model estimates that the solute would be at maximum concentration from the column surface to 0.1m from the surface- then the concentration decreased over the depths- reaching a concentration of 0.37 at the bottom of the column. In contrast, no clear trend was seen for bacteria, other than the highest HPCs were found between 0.2m and 0.32m from the column surface.

When compared to the bacterial distribution of Trial 3 (Figure 70 ), clear differences were seen. HPCs for Trial 3 were highest towards the soil surface, and lowest at 0.3m from the surface. Whereas in trial 2 (figure 75), HPCs were higher at 0.3m from the surface. Both trials had similar HPC CFU ranges over the sampling depths (Trial 3: $5x10^{9}$- $3.3x10^{10}$CFUs/g, Trial 2: $1x10^{10}$- $2.8x10^{10}$ CFUs/g).
In this trial, further samples, over a longer time period may have shown a clearer trend to compare bacterial distribution to the solute's distribution.

Taxonomic Classification

Relative abundance of different phyla present in soil samples taken from different depths was established based on assigned ASVs. The composition of each sample at phylum level, compared with the pre-trial soil sample is seen in Figure 76.

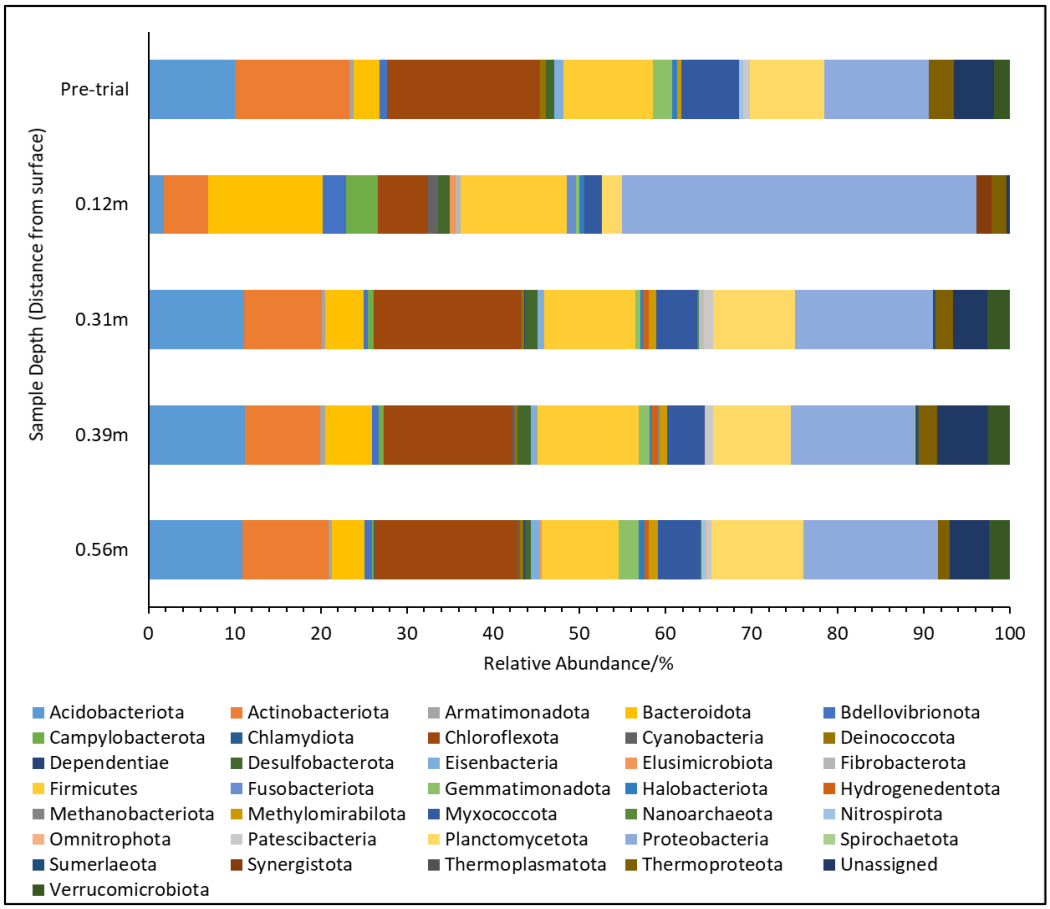


**Figure 76.** Comparison of bacterial communities present in Soil samples from floodwater trial 2 at phylum level (n=5). Unassigned taxa grouped into 'Unassigned' category.

For Trial 2, taxonomic data is available for samples taken from the following depths: pre-trial, 0.12m, 0.31m, 0.39m, and 0.56m . All 22 phyla detected in pre-trial soil were also detected in soil sampled from the column. Despite this, 13 phyla were unique to soil samples obtained at different depths. These phyla included: Nanoarchaeota, Dependentiae, and Spirochaetota. The abundance and number of different phyla varied over the sample depths but Proteobacteria was prominent in soil samples. Proteobacteria was present over the depths at 14-16% abundance, with a very high abundance at 0.12m of 41%- almost 30% higher than its abundance in the pre-trial soil sample (12%). Over the sample depths, some phyla appear once only, for example Chlamydiota was present at <0.1% at 0.56m only, and Spirochaetota at ~0.1% at 0.3m only.

At genus level, all 11 genera found in pre-trial samples, were detected in soil samples taken from different depths (Figure 77). Overall, the number of genera decreased with depth , highest – 24- at 0.12m and lowest – 13 – at 0.56m. *Povalibacter* was abundant throughout the soil depths ranging from 0.85% at 0.31m to 1.4% at 0.56m. However this genus had higher abundance in the pre-trial

soil sample at >2%. 3 genera were found at all soil depths (apart from pre-trial sample). These genera were: *Flavobacterium, Cloacibacterium*, and *Aliarcobacter.* All of which decreased down the depths of the column, i.e. highest abundance at 0.12m, and lowest at 0.56m. As these genera were not found in the pre-trial soil sample, they may have moved from the floodwater and into the soil, with higher abundance towards the soil surface.

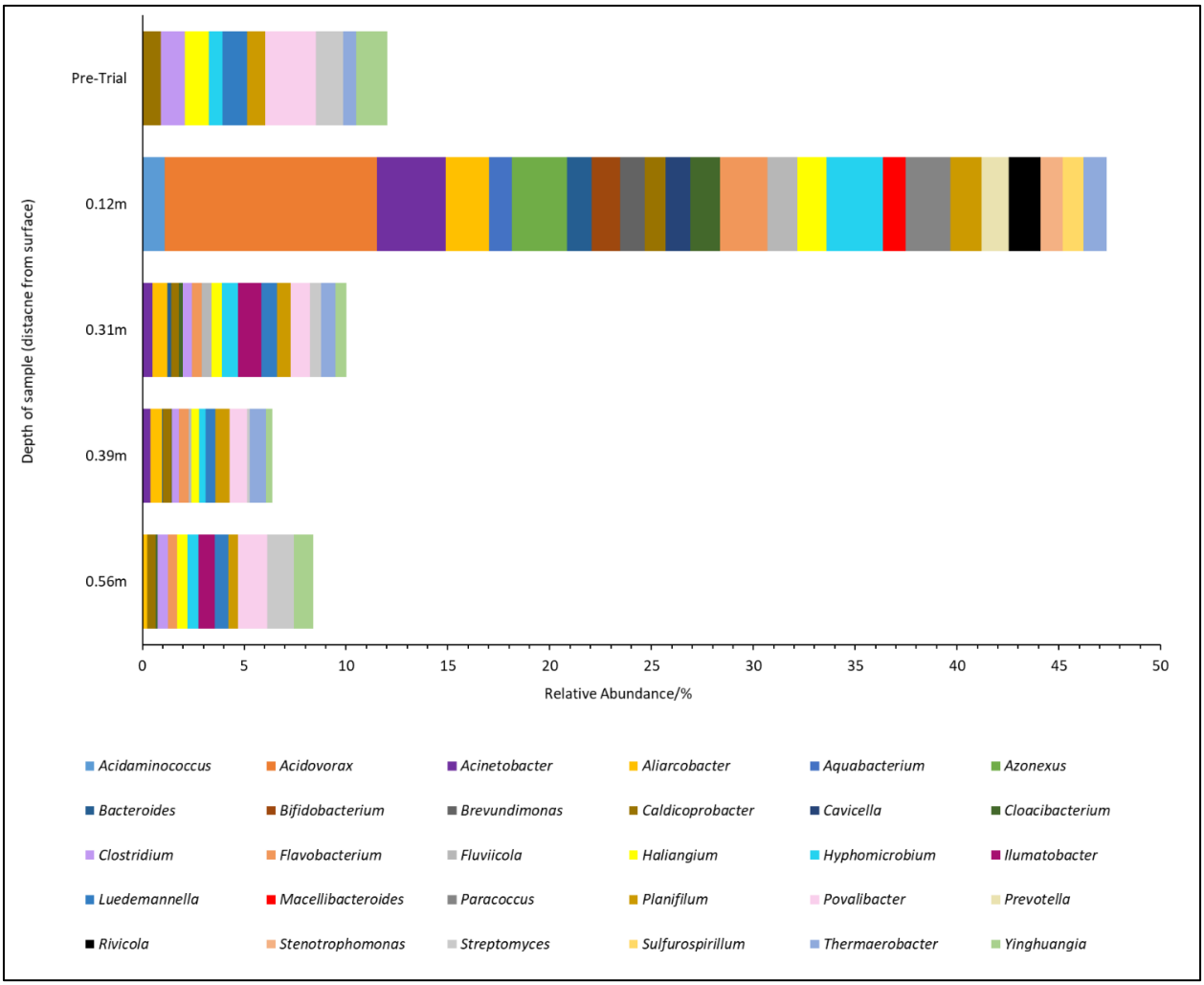


**Figure 77.** Comparison of bacterial communities present in Soil samples at different depths at Genus level (n=5). Filtered to include only genera with a relative abundance of >0.5%, in at least one of the samples taken.

At species level, sequences that were not assigned to a species were removed, along with any sequences that had a relative abundance of 0.5% at one depth sampled or more. For trial 2, in general, the number of species detected in soil decreased with depth. Starting with 30 species at 0.12m decreasing to 20 at 0.56m. 6 species were found at all sampling depths, and in the pre-trial sample. No clear trends were seen in relation to abundance of different species and sample depth. As well as this, no pathogenic or faecal related bacteria were found in soil samples at species level.

<u>Comparison of solute spatial distribution to specific soil species</u>

In order to accurately assess and compare distribution of a solute against bacteria, the relative abundance (%) of two species - *Caulifigura coniformis,* and *Methanosarcina thermophila* – were plotted against time on figure 78.

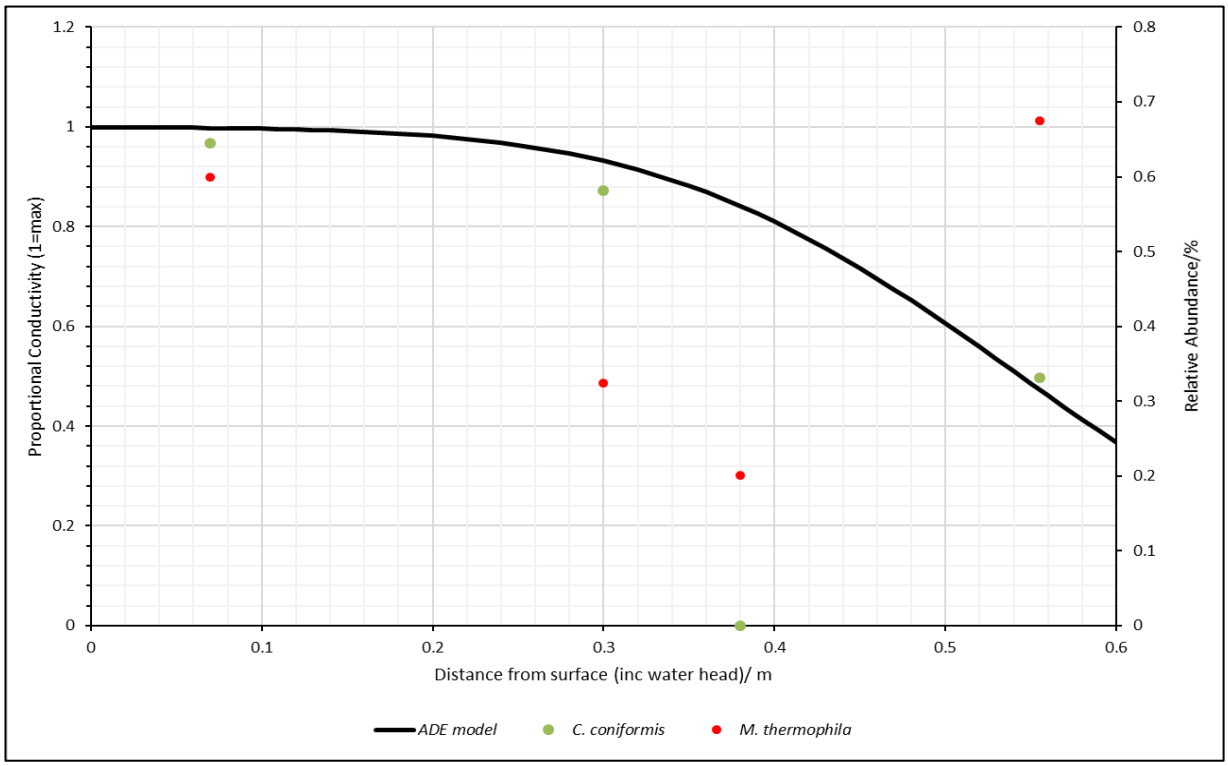


**Figure 78**. Proportional conductivity for spatial ADE model based on salt conductivity trials, plotted against relative abundance of 2 species from soil samples from Floodwater Trial 2. Relative Abundance plotted on secondary axis.

Both species follow the same trend in terms of abundance, having highest abundance towards the column surface, lowest at 0.38m, before increasing in abundance again at the bottom of the column. Although a different spatial distribution was seen when comparing the solute to the species- they are similar in that both the species abundance and solute conductivity are highest towards the column surface, before decreasing with depth. However, the solute conductivity continued to decrease after 0.5m, whereas the species abundance increased slightly.

Alpha Diversity Metrics

To determine alpha diversity, 3 unique measures were calculated at species level using all ASV's, for all soil samples (Table 20).

**Table 20.** Alpha diversity indices for soil samples taken from Floodwater trial 2 (n=5).

| Trial | Depth of sample (distance from surface) | ACE Species Richness | Faith's Diversity Index | Simpson's Evenness |
|---|---|---|---|---|
| Trial 2 | Pre-trial | 174 | 25.55 | 0.71 |
| | 0.12m | 104 | 17.74 | 0.68 |
| | 0.31m | 279 | 35.46 | 0.73 |
| | 0.39m | 399 | 44.66 | 0.73 |
| | 0.56m | 284 | 34.94 | 0.73 |

For trial 2, all alpha diversity metrics were lowest at 0.12m for this trial, and highest at 0.39m. Simpsons Evenness specifically was lowest towards the column surface (0.12m), with a higher value of 0.73 (higher than pre-trial) at 0.31m, which remained unchanged as soil depth increased. Pre-trial alpha diversity metrics were lower than that for all samples apart from at 0.12m.

### 6.5.6.3 Trial 1

Soil HPC CFU counts and spatial distribution

To compare the spatial distribution of bacteria in the soil vertically through the column to that of an inert solute, HPC CFU counts collected from the floodwater trial, from varying depths down the column were plotted against solute concentration based on the ADE model (Figure 79). The ADE curve was calculated as outlined by equation 6.5. Inputs for the ADE model were: diffusion coefficient from salt trial 3 (diffusion coefficient=2.44 × 10-6), and then flow rate (~0.18 litres per minute) and travel distance for the floodwater trial (Trial 1= 0.57m).

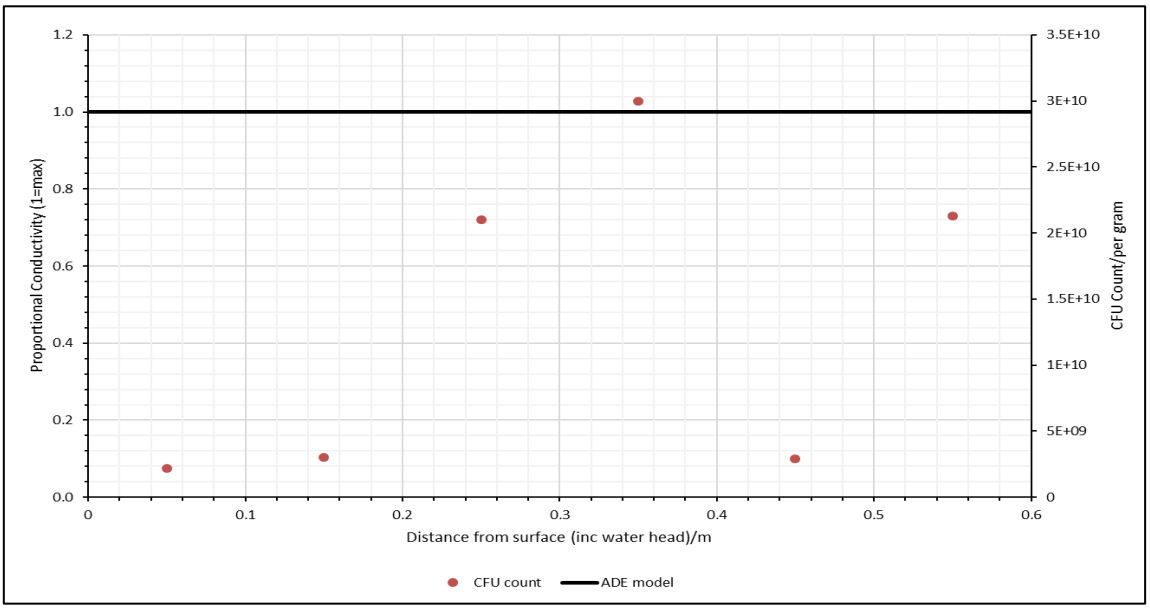


**Figure 79**. Proportional conductivity for spatial ADE model based on salt conductivity trials, plotted against HPC CFU count ,per gram ,for soil samples from Floodwater Trial 1 (n=6). CFU data plotted on secondary axis.

HPCs at all depths were higher than that of the pre-trial sample – $2.65x10^8$ CFUs/g. HPCs from trial 1 were lower towards the column surface ($2.2x10^9$ CFUs/g at 0.13m), but increased with depth until the highest CFU count was reached at 0.32m ($3x10^{10}$ CFUs/g). CFU count then decreased at 0.44m, before increasing again at the bottom of the column to $2.13x10^{10}$ CFUs/g.
The spatial distribution of bacteria in soil for trial 1 was clearly different from that of the solute. The solute conductivity was at maximum (1) over all of the soil depths, whereas HPC CFU counts were variable.

Taxonomic Classification

Relative abundance of different phyla present in soil samples taken from different depths down the column was established based on assigned ASVs. The composition of each sample at phylum level, compared with the pre-trial soil sample as seen in Figure 80. DNA data was available for all depths apart from pre-trial, and at 0.13m.

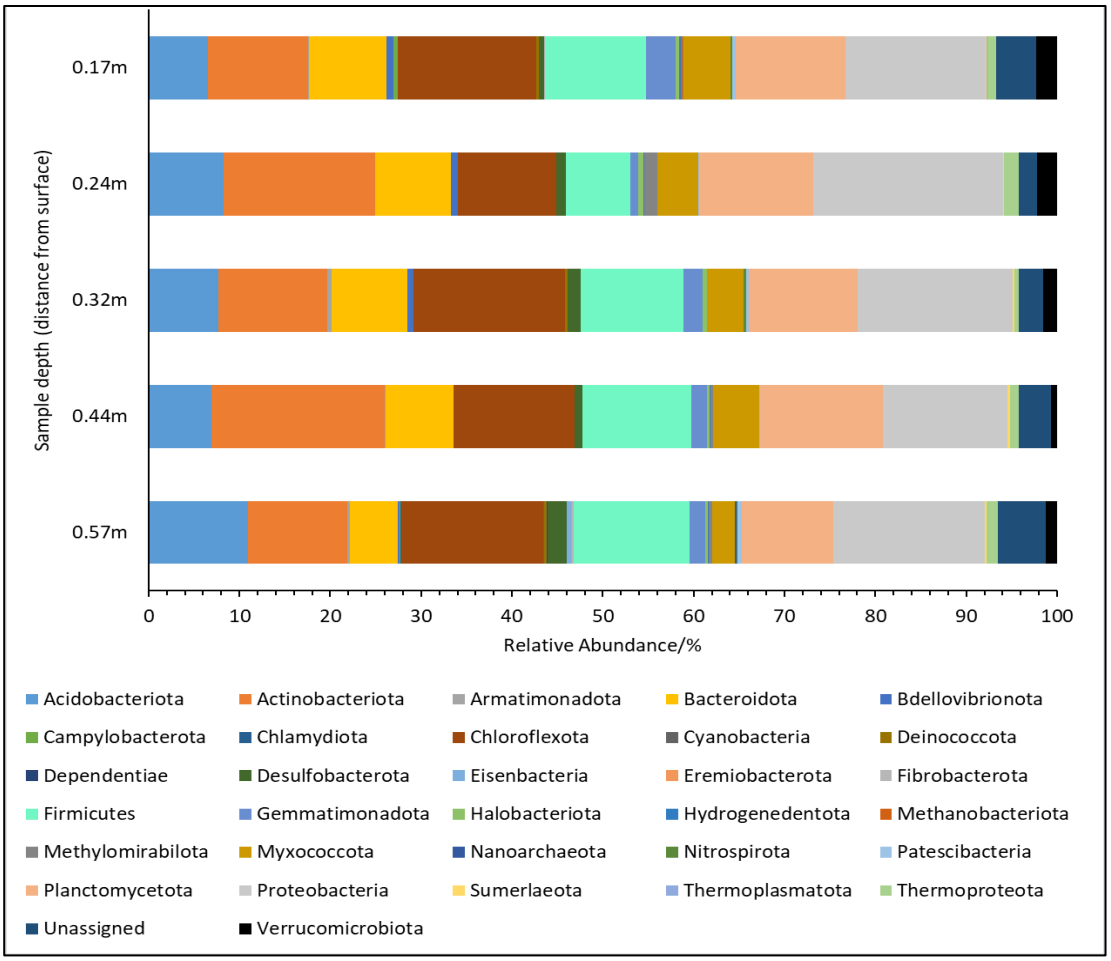


**Figure 80**. Comparison of bacterial communities present in Soil samples from floodwater trial 1 at phylum level (n=5). Unassigned taxa grouped into 'Unassigned' category.

For Trial 1 the number of phyla present at each depth was variable. The highest amount of phyla- 31 - were present at 0.57m, and the lowest number of phyla- 17, at 0.44m. 15 phyla were consistently

present in all samples, regardless of the depth. These phyla included Acidobacteriota, Bacteroidota, and Firmicutes. Proteobacteria was the most abundant phylum in all samples apart from that taken at 0.44m. Some phyla were present at one depth only, such as Chlamydiota, and Dependentiae, present at very low abundance (0.05-0.1%) at 0.57m only.

11 genera were found in all soil samples taken from different depths (Figure 81). Some such genera included: *Flavobacterium, Rhodococcus,* and *Caulifigura.* The number of genera decreased over the soil depths with 24 genera present at 0.17m decreasing to just 17 at 0.57m. Similarly, the abundance of specific genera decreased over sample depths. For example *Phenylobacterium,* had abundance of 1.33% at 0.17m , which decreased as soil depth increased, reaching <0.1% at 0.57m. Conversely, some genera followed the opposite trend, increasing in abundance as sample depth increased. One such genera was *Terrimonas,* present at <1% at 0.17m and increasing to 1.85% at 0.32m.

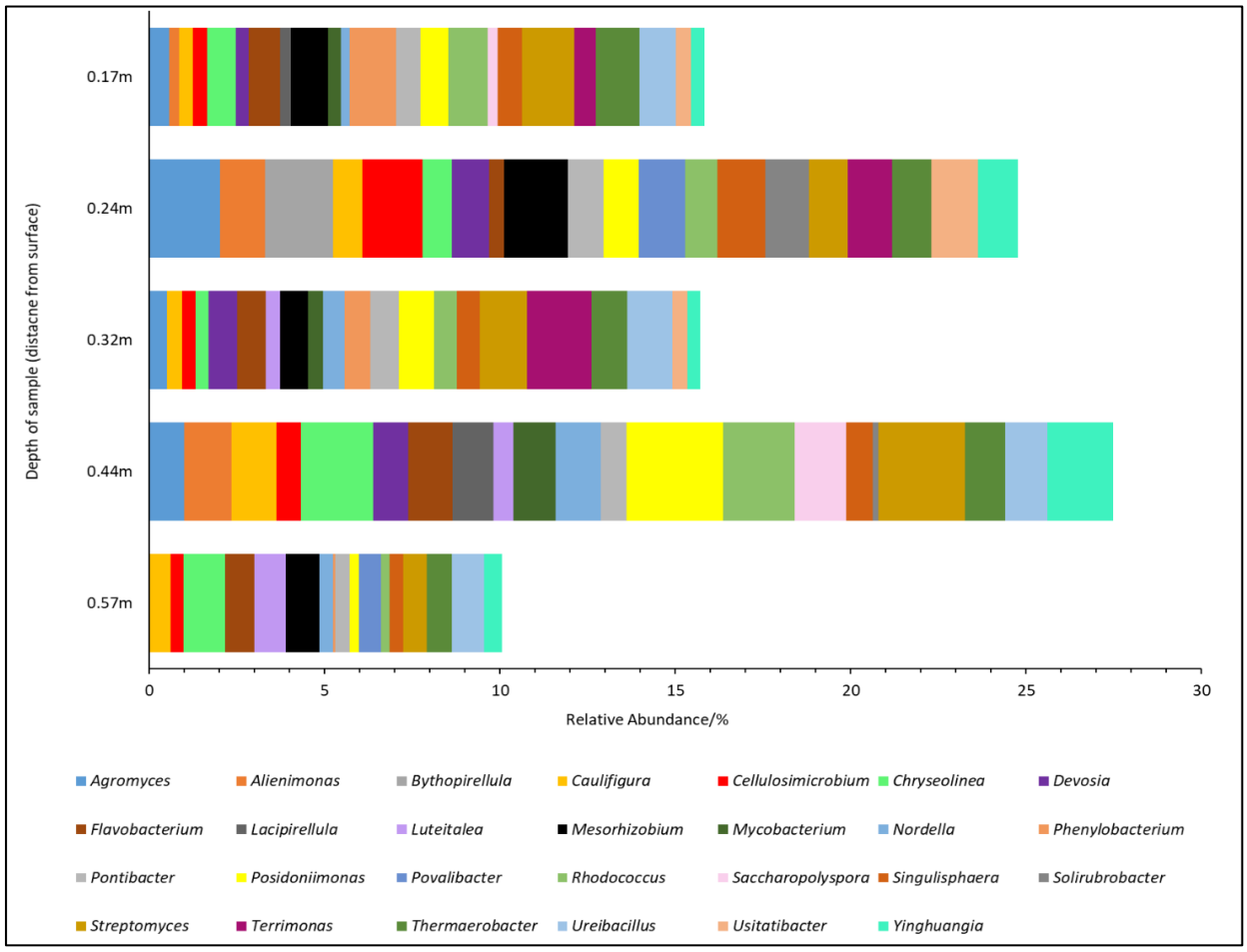


**Figure 81**. Comparison of bacterial communities present in Soil samples at different depths at Genus level (n=5). Filtered to include only genera with a relative abundance of >0.5%, in at least one of the samples taken.

The number of species present in soil varied over the depths with no clear trend. However the highest number of species- 28- was seen at 0.32m with the lowest, - 22- at 0.57m. 10 species were present in all samples, but no faecal associated or pathogenic species were detected in any soil samples

Comparison of solute spatial distribution to specific soil species

In order to accurately assess the distribution, two species were chosen to be plotted on Figure 82, that were present in soil at multiple depths in all three floodwater trials (not in any background soil samples). The relative abundance (%) of each species was plotted against time. These species were, *Caulifigura coniformis,* and *Methanosarcina thermophila*.

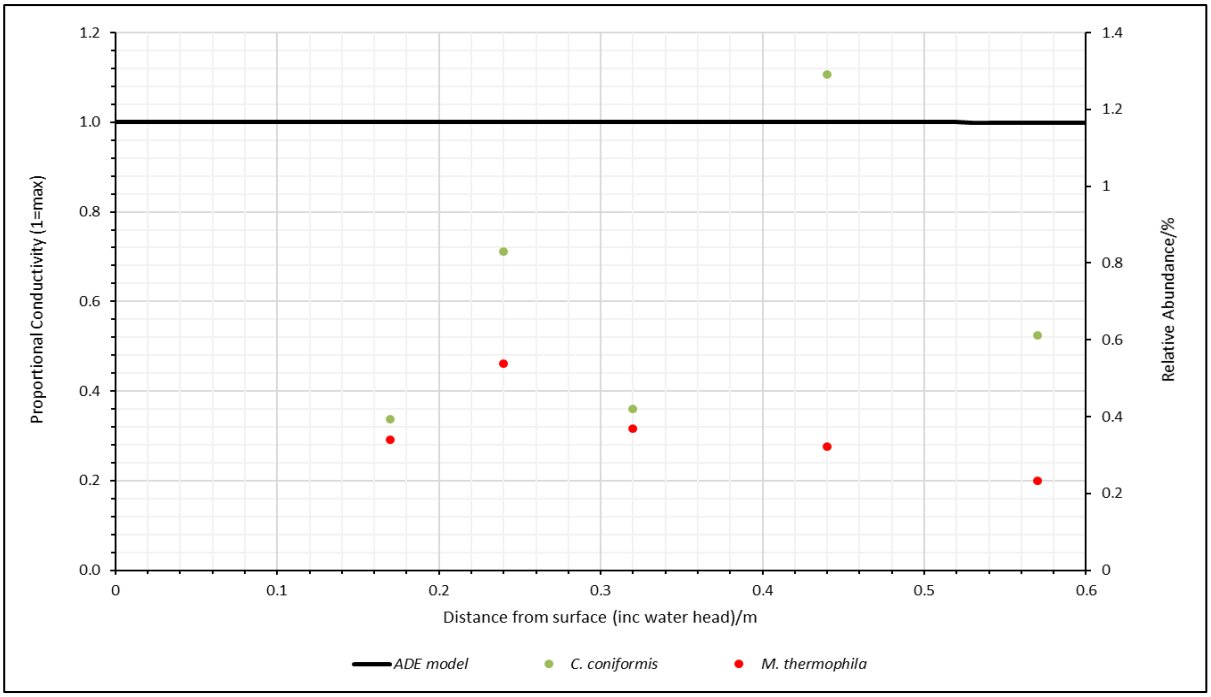


**Figure 82**. Proportional conductivity for spatial ADE model based on salt conductivity trials, plotted against relative abundance of 2 species from soil samples from Floodwater Trial 1. Relative Abundance plotted on secondary axis.

Relative abundance of *C. coniformis* was variable over the soil depths, with no clear trend of increase /decrease with depth. On the other hand, *M. thermophila* had highest abundance at 0.24m from the column surface, and then abundance decreased as depth increased, reaching lowest abundance at the bottom of the column (0.57m). This could indicate that *M. thermophila* tends to transfer out of the floodwater and into the soil towards the soil surface, with the transfer not as evident as depth increases. The spatial distribution of the different species in soil for trial 1 is clearly very different from that of the solute. The solute conductivity is at maximum (1) over all of the soil depths, whereas the species abundance is variable over the depths.

Alpha Diversity Metrics

To determine alpha diversity, 3 unique measures were calculated at species level using all ASV's, for all soil samples (Table 21).

**Table 21.** Alpha diversity indices for soil samples taken from Floodwater trial 1 (n=5).

| Trial | Depth of sample (distance from surface) | ACE Species Richness | Faith's Diversity Index | Simpson's Evenness |
|---|---|---|---|---|
| Trial 1 | 0.17m | 332 | 42.46 | 0.76 |
| | 0.24m | 155 | 22.44 | 0.76 |
| | 0.32m | 263 | 32.50 | 0.76 |
| | 0.44m | 182 | 24.32 | 0.72 |
| | 0.57m | 418 | 46.09 | 0.63 |

For trial 1, species richness and faiths diversity was variable, but both metrics had highest values at 0.57m, and lowest at 0.24m. Simpson's Evenness wa higher towards the column surface wit the value remaining at 0.76 until a depth of 0.44m. The evenness value then decreased to the lowest value of 0.63 at 0.57m.

**6.6 Discussion**

From this study, it is clear that over the depths of the soil column, differences in the distribution of bacteria were found, including differences in species type, number, and diversity. As well as this, it can be assumed that transfer of bacteria occurred between soil and water, however no faecal or pathogenic bacteria detected in water samples were found in any soil samples. To understand the parameters driving these differences and the implications that this has for assessing bacterial risk under urban flood conditions, the transfer of bacteria from water to the soil surface and also into the soil column should be considered. Multiple factors influence these processes, including bacterial cell properties (such as size and morphology), solution chemistry (pH, temperature, solute concentration), soil characteristics (grain size, type, porosity), and hydrodynamics (flow rate, velocity) (Gerlach and Cunningham, 2010). This discussion will delve into specific aspects such as; the impact of soil type and grain size on bacterial adsorption and transport, the role of bacterial aggregation including biofilm formation, and the utilisation of tracer studies in investigating bacterial transport dynamics. Furthermore, insights are provided into potential enhancements that could be applied to future studies to improve the understanding of bacterial behaviour in these environments.

6.6.1 Distribution of bacteria vertically through soil column in relation to soil characteristics

After allowing urban flood water collected from the field to infiltrate through a soil column of sandy loam soil, this study found a change in bacterial species and number at different soil depths, compared to pre-trial soil samples. Notably, HPCs were highest in soil at 15cm-30cm from the soil surface. This is similar to previous studies, that have found that when tracking bacterial distribution

through soil columns, higher bacterial cell concentrations are found towards the column surface. One study measuring *Cryptosporidium* in a loam soil column found ~62% of the oocysts were recovered at the top 10cm of the soil column, with only ~5% recovered from the bottom 5cm (Petersen e*t al*., 2012). Additionally, a soil column study investigating how application of sewage sludge affects vertical microbe transport found the highest numbers of human pathogen *Salmonella enterica* were found at the top 0-5cm of the soil column (Horswell et al., 2010). This is also the case for 'natural soils'. A study looking at bacterial community composition in soil columns from a forested montane landscape, finding highest bacterial diversity within the top 10cm of the soil (Eilers et al., 2012) . In relation to our study, it was found that the number of different species identified was highest at soil surface level and up to 40cm below the soil surface. However, (alpha) diversity metrics were variable and did not show a consistent trend or correlation throughout the trials.

It is thought that the higher number of HPCs and higher number of species found in soil samples towards the surface is due to the soil characteristics- namely grain and pore size, contributing to a 'straining' effect. Sand-based soils , like the one used in this study (sandy loam) exhibit a filtering effect on many bacteria due to their granular composition and pore structure (Bahgat et al., 1999). One column study using sand and *E.coli,* found that the level of straining (bacteria 'trapped' in the column') increased with the use of sand types with smaller grain size and smaller pore size (Bai et al., 2016). As bacteria pass through the column, the level of 'straining' would increase with column depth/length, with more pores available for bacteria to be 'trapped' in (Du et al., 2021).

However, our study did also find high numbers of HPCs at the bottom of the column in one of the three trials. It is thought that this could , be related to pore size (macropore formation or uneven pore distribution), but could more likely be due to preferential flow. Macropores and preferential flow paths are commonplace in 'field' soils- created by factors such as: plant roots disrupting the soil profile, invertebrate activity -i.e. earthworms- and natural weather cycles (swell-shrink cycles from temperature and moisture changes) (Jarvis and Larsbo, 2023; Kirkham, 2014; Wang and Wei, 2015). In our study macropores may have been created due to the human aspect of packing the column- although soil was sieved and compacted, it was difficult to measure the level of compaction used and to replicate this with the addition of each soil layer. This could have increased the likelihood of macropores or uneven pore distribution. Preferential flow paths could have occurred by the column walls- with some other column studies coating the walls with material similar to that of the soil , or 'scuffing' the walls, to create a rougher texture to try to avoid preferential flow routes (Gilbert *et al.,* 2014; Smajstrla, 1985). If this was the case, and the floodwater found a preferential route, or moved through the column through larger macropores, this could explain the high levels of bacteria found at the bottom of the column. The water, transporting bacteria - without the bacteria becoming trapped by smaller pore sizes - could have reached the mesh at the bottom of the column, with bacteria then becoming trapped on the fine mesh. These flow paths are essentially what could also cause the spreading of an inert substance-such as the salt used in the saltwater trials. With this being the case, although tests show a concentration of bacteria towards the surface, this could explain how some bacteria travel more quickly through the column (like the solute).This could also have been enhanced by the translocation and accumulation of fine soil material observed at the bottom of the column (sitting on the mesh). The bacteria may have adsorbed to the finer material - most likely clay- due to its adsorptive capacity ( Burford *et al*., 2003; Cuadros, 2017).

6.6.2 Biofilm formation and bacterial aggregation affecting infiltration into soil column

Another process that could explain the higher levels of bacteria present towards the column surface is biofilm formation and bacterial aggregation. Bacterial cells are rarely present as singular cells, and tend to exist as aggregations of multiple cells, suspended, in this case, in water (O'Toole et al., 2000). These aggregations can adsorb to each other ,as well as to particles and surfaces , such as soil (Kostakioti et al., 2013). Studies have even indicated that the preferred lifestyle for bacterial cells in soil is biofilm formation (Fenchel, 2002; Young and Crawford, 2004). When the adsorption and growth rate increases more than that of the rate of desorption, biofilms form. Biofilm formation aids bacterial survival in a number of ways- mainly due to the self-produced extracellular matrix that coats the biofilm, giving stability to the biofilm structure, as well as trapping nutrients and water , protecting from stressors such as desiccation, and allowing cell to cell interactions, including gene transfer (Flemming and Wingender, 2010; Jefferson et al. 2005; Lewis, 2005).
Biofilm formation affects permeability and porosity of soil , with a study using *Pseudomonas aeroginusa* as the inoculum, finding biofilm formation in sand decreased flow rate and reduced porosity by up to 96% (Cunningham et al., 1999). Another study found biofilm formation caused permeability in sandy soil to decrease from 9.9 to 4.9 (Darcy) (Dunsmore et al., 2003). This is due to the biofilms 'clogging' the pore spaces, leaving fewer or smaller pores in which the water (transporting other bacterial cells) can travel through (Nevo and Mitchell, 1967). Interestingly, it has also been found that bacteria can 'clog' soils towards the soil surface , especially in nutrient rich soil, due to overproduction of excreted extracellular polymeric substances (EPS- 'coating ' of the biofilm) (More et al., 2014). This creates a layer of soil at the surface with reduced hydraulic conductivity (Hill and Palange, 1972).

This bacterial 'layer' that could have formed at the surface of the soil column could be thought to be similar to the 'schmutzdecke' (German for 'dirt layer') that forms at the surface of sand filters during water treatment (Hammes et al., 2011). This layer accumulates organic and inorganic debris, as well as bacteria,algae, fungi and minerals which once formed ,can actually help with further biofiltration of water passing through, straining bacteria from the water, thus reducing the bacterial level in the water (Brandt et al., 2017; Moran, 2018; Yildiz, 2012). At the end of the floodwater trials, a layer of 'sludge' did appear to be present at the surface of the soil column, and so this layer could very well have been due to a build up of bacterial cells and debris, forming a 'biofilter'. It is possible that this could also happen at a field flood site- with the flood water infiltrating through the soil, with a higher abundance of bacteria towards the surface of the soil . This would mean that there is a higher risk of a member of the public coming into contact with said bacteria, even after the visible floodwater recedes. This could especially be worrying, for example in the context of a childrens play area (such as that sampled in chapters 4 and 5) - that often contain 'sand pits' for recreation. Children digging into the sand could come into contact with bacteria originating from urban flood water, accumulated at depths of up to 40cm. This is only a concern if the bacteria have the potential to cause disease, which also includes being alive at the time of contact. This study did not find any of the pathogenic species found in flood water (such as *Dialister invisu*s and *Phoecicola vulgatus*), had been transported into the sampled column soil, but it is known that such species can in fact be deposited from contaminated water to soils- as discussed in Chapters 4 and 5. Having said this, *D. invisus* , present in the pre-trial floodwater sample, was detected in effluent collected at the bottom of the column- indicating that some of this species had remained suspended in the water. This

indicates that this species is more likely to be transported through this soil type, rather than become trapped/sorbed to the soil. In a 'field' flood situation, if *D. invisus* did not become attached to other bacteria or soil particles it could pose a risk further into the soil profile if reaching groundwater.

### 6.6.3 Temporal and spatial distribution of bacteria though the soil column in comparison to a solute

Traditionally, inert tracers, such as salt, , have been used for the purpose of hydraulically characterising the soil, determining factors such as: infiltration rate, velocity, and diffusion coefficients (Divine and McDonnell, 2005; Hautojärvi, et al., 1997). Theoretical models are often used to validate data from tracer experiments, as seen in our study (using an ADE model) but also in many others, such as that of Masipan *et al* 2016., using a convection-dispersion model to validate data from a bromide tracer experiment- amongst others (Jiang et al., 2005; Masipan et al., 2016; Sinton et al., 2000). Naturally, inert tracers will not behave in the same way as bacterial cells- which have the ability to absorb and desorb to particles and surfaces, which inert tracers cannot. Microorganisms have even been used as tracers to resolve this- with a review of bacterial groundwater tracers finding bacteriophages, actually make good tracers for tracking underground water movement due to their size, lack of pathogenicity, and ability to behave as bacterial cells (sorption behaviour) (Keswick et al., 1982). Bacteriophages were also used successfully as a tracer to determine the efficiency of sand-filters in removing pathogens from wastewater (Vega et al., 2003).

During this study the temporal and spatial distribution of bacterial HPC CFUs, as well as abundance of specific species, in water and soil were compared to that of a calculated ADE model based on solute (salt) transport experiments. Our study found a difference in the travel times of salt, as well as the spatial distribution of salt, compared to bacteria suspended in water, through the soil column. Plate counts showed that bacteria suspended in water generally moved through the soil column more slowly than the solute. Not only that but when looking at the movement of specific species compared against the solute, both *Prevotella copri* and *Flavobacterium album* abundance decreased over time in the water samples- indicating deposition of the bacteria out of the water and into the soil. As mentioned in the previous section, it is thought that the slower travel time of bacteria compared to salt is due to bacterial aggregation, soil pore size, and sorption capabilities of the bacteria (Bai et al., 2016; Gannon et al., 1990; Hong and Brown, 2006; Santore, 2022). These same factors would impact the spatial distribution, with our study finding that distribution of HPCs was variable over the soil depths, with a higher number of HPC at the top 0.3m of the soil column. In 2 out of the three trials, *C. conformis* also showed highest abundance between 0.3-0.4m from the column surface. The bacteria may not have travelled as far into the soil column as the solute due to being trapped by the small pore sizes of the soil or adsorbed to the soil particles. It could also be due to the floodwater potentially not having enough time to travel through the entire length of the soil column. Without the flow and hydraulic pressure, the bacteria may not have been able to travel further down the column depths. It is thought that it is more likely a combination of multiple factors that influence this, including those discussed in this section, as well as some factors not discussed here. Such factors include: cell surface chemistry, cell shape, and ionic interactions between the cell surface and soil particles (Bai et al., 2016; Gannon et al., 1990; Hong and Brown, 2006; Santore, 2022).

In summary, it has been discussed that there are a plethora of factors governing both the spatial and temporal distribution of bacteria through a sandy loam soil column. This study found that bacterial transport happens at a much slower rate than that of an inert tracer, due to factors such as soil structure, and bacterial aggregation behaviour. Not only this, but it was found that after floodwater was ran through the soil column, bacteria appeared to be concentrated at the top 0.4m of the soil column- with the distribution of different species governed by the aforementioned factors. This study illustrates how bacterial contaminants found in floodwater can transfer into urban soils, and may do so in the field. Although no pathogenic bacteria were found to have remained in the soil column, other bacteria originating from floodwater were found, at the soil surface, but also at all depths throughout the column. This also demonstrated how the soil type and grain size of a sandy loam soil, can act as a filter for bacteria- trapping them, rather than the bacteria infiltrating through and into water sources such as groundwater. However, one pathogenic species- *Dialister invisus* was found to be present in water effluent from the column-indicating that the soil did not 'filter' out this pathogen-posing a potential contamination risk to groundwater if this did happen in-situ.

### 6.7 Conclusions

From this research it can be concluded that:

- After application of urban floodwater data from HPCs and taxonomic analysis indicated bacteria are generally concentrated within the top 0.4m of the soil column following application of urban floodwater.
- Faecal associated and pathogenic species- *Phocaeicola vulgatus* and *Dialister invisus* were detected in floodwater but were not found to have transferred into the soil profile.
- Bacterial transport through the soil column is slower than that of a solute, and distribution is also different to that of a solute. The transport and distribution of bacteria is thought to be influenced by soil structure, bacterial agglomeration and/or biofilm formation.

Data from this study contributes to the existing literature on not only, bacterial structure of urban floodwater, but also the dynamics governing bacterial transport-specifically in an urban flood scenario, using flood water from the field and a soil representative of urban field conditions. Taxonomic data indicates that following a flood event, although the top 0.4m of soil is abundant with various bacterial species- none of these bacteria were found to be dangerous or a risk to public health. In fact the soil type used- loam with a high percentage of sand (70%) appears to work as a 'filter' for floodwater originating bacteria. Despite this, it is concerning that pathogenic *D. invisus* was not completely 'filtered' out of the floodwater by the soil, giving opportunity for the species to be transported further into the soil profile and posing a risk via contamination of groundwater. Having said this, further research would be beneficial utilising this 'column' style experiment setup. This includes: further column tests with different soil types to determine the influence of soil characteristics such as pore size. Tests controlling and changing specific flow rates to determine hydraulic influences of bacterial transport and distribution, and also potentially 'reverse' tests, to determine whether bacteria in already contaminated soil has the potential to move back out of the soil and into clean water (replicating the dynamic of a frequently flooded field site).

# Chapter 7
# Concluding remarks and Future work

**7.1 General Overview**

The overall research aim of this project was to investigate the public health risk of urban flooding events. To fully, investigate this, two field sites in Sheffield, UK, - subject to frequent urban flooding, were sampled over the long term (several months) and the short term (3 days consecutively following heavy rainfall). By sampling flood affected soil, and floodwater at the sites, along with measuring physicochemical factors (pH, temperature, rainfall depth), the bacterial communities present, along with their behaviour and response to changes could be tracked. Pathogenic or disease causing bacteria could also be identified, along with their persistence and movement over the soil-water interface over different timescales. This field study was complemented with a laboratory based soil column study. By running floodwater from a field site through a sandy loam soil column, travel times and distribution of bacteria typically found at a flood site could be investigated. Changes in the bacterial community were measured over varied depths, to determine whether bacteria adsorb to the soil or remain suspended in the floodwater, with consequences for public health risk in either scenario. Whilst other studies have attempted to quantify the health risks of flooding before, the methods used have been more basic in nature- testing for only specific species, or investigating the health risks after one singular flood event, rather than over different timescales. Traditional methods, such as heterotrophic and coliform plate counts were used to estimate bacterial abundance, with a more sensitive molecular approach- sequencing of the 16s rRNA gene - used to determine which bacteria were present down to species level.

The long term field study aimed to develop a new understanding of the bacterial composition of soil and water in urban flood areas and investigate how this varies over seasonal timescales. The study focused focus on the specific aspects of : identifying the bacterial composition of urban flood water, and evaluating whether frequent flooding has an effect on the soil bacterial community, understanding whether the bacteria from the floodwater move between the soil and water, and assessing how the composition of the bacterial communities at a flood site change over seasonal timescales. After analysis and discussion, this study found a diverse community of bacteria existed within urban floodwater, including species that may pose a risk to public health (*Arcobacter* and *Bacteroides* species). The soil affected by floodwater at this site showed a long term altered bacterial community structure, with bacteria thought to have transferred out of the floodwater and into the soil, and vice versa. Changes in the bacterial community were observed seasonally, including changes in diversity, and bacterial abundance, which may be driven by bacterial preference to different moisture and temperature conditions that fluctuate over the seasons. Interestingly this long term study highlighted the need for a consequent short term study, with bacterial abundance showing changes following shorter, intense rainfall events over the months sampled.

Following the findings of the long term study, the short term study aimed to identify the bacterial communities present in soil and water in the short term- 3 days following an urban flood event, as well as investigating the behaviour over the short term-determining movement of bacteria across the soil-water interface. This study also looked at how potential faecal contamination of floodwater

influences the bacterial community structure of receiving soils, and whether traditional culture-based methods (Heterotrophic plate counts and faecal indicator bacteria) are most appropriate for assessing contamination at an urban flood site. Results of this study showed that, faecal contamination was thought to be present in floodwater at one site-Tongue Gutter, with the presence of potentially pathogenic species present, including *Faecalibacterium prasunitzii* and *Dialister invisus*. The presence of faecal contamination was found to alter the soil bacterial community , with alpha diversity lower at Tongue Gutter (site thought to have faecal contamination). However this study did find that potentially pathogenic species did not appear to have transferred to the flood affected soil in the short term. Despite this, some evidence of bacterial transfer over the soil-water interface was found at both sites, with *Acinetobacter* and *Peredibacter,* found in floodwater and affected soil but not in soil not subject to flooding. This study also concluded that for identifying health risks at urban flood sites, a combination of traditional methods- HPCs, and more advanced methods- next generation sequencing, used together would be most appropriate. HPCs can indicate whether faecal contamination may be present, with sequencing methods being able to identify the species, and source of contamination (whether from human sewage, or animal etc.) .The results of this study indicate faecal contamination of floodwater, inferring the floodwater may be the main vector of risk to public health in the 3 days following an urban flood. Especially since no pathogenic or faecal associated bacteria were found in flood affected soil over the time period.

Findings from the field studies gave an indication of the wide range of bacteria found in urban flood affected soils and in urban floodwater itself. With indications of faecal contamination and also movement of bacteria between the soil and water, influencing changes in bacterial community composition and abundance. To further understand the risk urban flood events may pose to public health, it was important to investigate bacterial behaviour in a more controlled setting, in which bacterial transport in floodwater could be investigated over time, and the effect of floodwater on soil communities could be observed spatially. Thus, the column based study aimed to determine the spatial and temporal distribution of bacterial communities present in floodwater when ran through a soil column, and understand how this effected the soil bacterial community. This study also investigated whether soil characteristics and composition of a typical urban soil, along with bacterial biological behaviour (adsorption and biofilm formation), were likely to have an effect on the distribution of bacteria in an urban flood scenario. This study found that bacteria seemed to be most abundant towards the surface and down to 0.4m of the soil column, after floodwater had been applied. This was thought to be due to the 'filtering' capabilities of the sand present in the soil, with small pore sizes trapping bacteria, meaning they did not travel as far into the soil profile. It was also thought that biofilm formation and bacterial aggregation may have affected this further, meaning over time, bacteria would adhere to biofilms towards the soil surface rather than remain suspended in the floodwater. Not only this but the faecal associated and pathogenic species detected in floodwater- *D. invisus* and *Phoecicola vulagtaus*, did not seem to transfer into the soil, with *D. invisus* being present in floodwater that had infiltrated through the entire depth of the column.

**7.2 Summary of conclusions**

- Faecal contamination and potentially disease causing bacteria were found in floodwater at one site- Tongue Gutter- during this study. Such bacteria included *Faecalibacterium prausnitzii, Dialister invisus*, and *Phoecicola vulagatus*.
- Frequent urban flooding affects the bacterial community structure of the receiving soil, reducing diversity, and introducing species thought to have originated from floodwater.

- Based on laboratory testing, after the application of urban floodwater bacteria tend to be concentrated from the surface down to 0.4m in the soil profile.
- The bacterial communities at flood sites are affected in the long term by seasonal changes - with changes in soil moisture content and temperature causing fluctuations in bacterial abundance and diversity.
- In the short term (days after heavy rainfall and thus flood event), floodwater appears to be the vector of risk to public health, with no potentially disease causing bacteria detected in receiving soils.
- In the long term, potentially disease causing bacteria, though to have originated from floodwater was found in flood affected soils, raising concerns of the soil also posing a public health risk.
- To accurately determine bacterial risks at an urban flood site a combination of traditional coliform counts and a more specific molecular method such as next generation sequencing is required.

### 7.3 Future Work

Although this study has highlighted that potentially disease causing bacteria can be present at urban flood sites, in both floodwater , but also in affected soils, there are a number of subsequent complimentary studs that could be performed to enhance the knowledge presented here, and further understand the risk to public health. Naturally an epidemiological study is suggested, with amount of bacteria required to actually cause illness in a human quantified to determine whether bacteria found in this study exist in large enough amounts to cause disease. Along with this, behavioural or psychological studies could be carried out around human behaviour in urban flood settings, which could link to the likelihood of contaminated water or soil being ingested by a member of the public. A study could also be carried out looking at survival rates of bacteria found in urban floodwater and soils, perhaps using methodology such as flow cytometer analysis, as only living bacteria can infect and cause disease. In regards to further soil column studies that could be carried out, it would be interesting for similar experiments to be performed using various soil types to understand how differing soil structures and characteristics affect infiltration and adsorption of bacteria in floodwater. Column studies could also be ran 'in reverse', with soil already saturated with contaminated floodwater, having a 'clean water' (mimicking rainwater) head sat on the soil surface. It could then be determined whether bacteria transfer from the soil, and back into the standing water, creating another route and vector of risk. As well as this, as it was found in the column study that pathogenic bacteria, including *D. invisus*, was able to pass through the soil and remain in the effluent, raising a new concern for the fate of such species, potentially being transported into and contaminating groundwater. As such, a study could be performed, similarly using a column set-up to investigate the likelihood of potentially dangerous bacteria transferring through soils at flood sites and into groundwater.

This study contributes new information on the bacterial communities present in urban floodwater and also affected soils, along with examining the dynamics affecting bacterial transfer and travel times, and distribution in urban soil. Data from this study ,consolidated with information from aforementioned suggested studies, could be used by authorities to determine the safety of areas subject to frequent urban flood events, especially in terms of bacterial risk, and risk posed to public health.

# References


Abdellatif, M., Atherton, W., Alkhaddar, R.M., Osman, Y.Z., (2015), Quantitative assessment of sewer overflow performance with climate change in northwest England, Hydrological Sciences Journal, 60, 636-650

Abdelsalam, N.A., Hegazy, S.M. and Aziz, R.K., 2023. The curious case of Prevotella copri. Gut Microbes, 15(2), p.2249152.

Abu-Ashour, J., Joy, D.M., Lee, H. et al. Transport of microorganisms through soil. Water Air Soil Pollut 75, 141–158 (1994). https://doi.org/10.1007/BF01100406

Acharya, K., Khanal, S., Pantha, K. et al. A comparative assessment of conventional and molecular methods, including MinION nanopore sequencing, for surveying water quality. Sci Rep 9, 15726 (2019). https://doi.org/10.1038/s41598-019-51997-x

Allen, M.J., Edberg, S.C. and Reasoner, D.J., 2004. Heterotrophic plate count bacteria—what is their significance in drinking water?. International journal of food microbiology, 92(3), pp.265-274.

Andersen, T.K., Marshall Shepherd, J., (2013) Floods in a changing climate, Geography Compass, 7, 95-115.

Anderson, M.J., 2014. Permutational multivariate analysis of variance (PERMANOVA). Wiley statsref: statistics reference online, pp.1-15.

Andrade, L., O'Dwyer, J., O'Neill, E. and Hynds, P., 2018. Surface water flooding, groundwater contamination, and enteric disease in developed countries: A scoping review of connections and consequences. Environmental pollution, 236, pp.540-549

Ansari, S., Dhital, R., Shrestha, S., Thapa, S., Puri, R., Chaudhary, N., Khatiwada, S. and Gautam, R., 2016. Growing menace of antibacterial resistance in clinical isolates of Pseudomonas aeruginosa in Nepal: an insight of beta-lactamase production. BioMed Research International, 2016.

Arnone, R.D. and Perdek Walling, J., 2007. Waterborne pathogens in urban watersheds. Journal of Water and Health, 5(1), pp.149-162.

Baek, C., Shin, S.K. and Yi, H., 2018. Flavobacterium magnum sp. nov., Flavobacterium pallidum sp. nov., Flavobacterium crocinum sp. nov. and Flavobacterium album sp. nov. International Journal of Systematic and Evolutionary Microbiology, 68(12), pp.3837-3843.

Bahgat, M., Dewedar, A. and Zayed, A., 1999. Sand-filters used for wastewater treatment: buildup and distribution of microorganisms. Water Research, 33(8), pp.1949-1955.

Bai, H., Cochet, N., Drelich, A., Pauss, A., Lamy, E., (2016). Comparison of transport between two bacteria in saturated porous media with distinct pore size distribution. Research Advances, 6,14602-14614.

Bai, H., Cochet, N., Pauss, A. and Lamy, E., 2016. Bacteria cell properties and grain size impact on bacteria transport and deposition in porous media. Colloids and Surfaces B: Biointerfaces, 139, pp.148-155.

Baker-Austin, C., Oliver, J.D., Alam, M., Ali, A., Waldor, M.K., Qadri, F., Martinez-Urtaza, J., (2018). Vibrio species. Infections. Nature Reviews Disease Primers, 4, 1-19.

Bakker, J.D., 2024. PCoA. Applied Multivariate Statistics in R., https://uw.pressbooks.pub/appliedmultivariatestatistics/.

Barboza, K., Cubillo, Z., Castro, E., Redondo-Solano, M., Fernández-Jaramillo, H. and Echandi, M.L.A., 2017. First isolation report of Arcobacter cryaerophilus from a human diarrhea sample in Costa Rica. Revista do Instituto de Medicina Tropical de São Paulo, 59, p.e72.

Barrell, R.A., Hunter, P.R. and Nichols, G.C.D.P.H., 2000. Microbiological standards for water and their relationship to health risk. Commun Dis Public Health, 3(1), pp.8-13.

Basri, Z.D.M., Othman, Z. and Wahid, M.A., 2015. Detection of pathogenic bacteria in flood water. In ISFRAM 2014: Proceedings of the International Symposium on Flood Research and Management (pp. 185-193). Springer Singapore.

Becker Y. Chlamydia. In: Baron S, editor. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996. Chapter 39. Available from: https://www.ncbi.nlm.nih.gov/books/NBK8091/

Belova, S.E., Pankratov, T.A. and Dedysh, S.N., 2006. Bacteria of the genus Burkholderia as a typical component of the microbial community of Sphagnum peat bogs. Microbiology, 75, pp.90-96.

Bergström, L., (2000). Leaching of agrochemicals in field lysimeters—A method to test mobility of chemicals in soil. Pesticide/Soil Interactions, 279-285, Paris, France.

Bett B., Tumusiime D., Lindahl J., Roesel K., Delia G. (2021) The Role of Floods on Pathogen Dispersion. In: . The Handbook of Environmental Chemistry. Springer, Berlin, Heidelberg. https://doi.org/10.1007/698_2021_761

Bitton, G., Henis, Y., Lahav, N., (1972). Effect of several clay minerals and humic acid on the survival of Klebsiella aerogenes exposed to ultraviolet irradiation. Applied microbiology, 23,870-874.

Bolster, C.H., Walker, S.L. and Cook, K.L., 2006. Comparison of Escherichia coli and Campylobacter jejuni transport in saturated porous media. Journal of environmental quality, 35(4), pp.1018-1025.

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, and Caporaso JG. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology 37: 852–857. https://doi.org/10.1038/s41587-019-0209-9.

Booth, D.B, (1991), Urbanization and the natural drainage system: Impacts, solutions and prognosis, Northwest Environmental Journal., 7, 93-118.

Borowik, A., and J. Wyszkowska. "Soil moisture as a factor affecting the microbiological and biochemical activity of soil." Plant, Soil and Environment 62, no. 6 (2016): 250-255.7
Boyer, D.G., Kuczynska, E., Fayer, R., (2009). Transport, fate, and infectivity of Cryptosporidium parvum oocysts released from manure and leached through macroporous soil. Environmental geology, 58, 1011-1019.

Brandt, M.J., Johnson, K.M., Elphinston, A.J. and Ratnayaka, D.D., 2017. Chapter 9: Water Filtration. Twort's Water Supply, pp.367-406.

Brennan, F.P., Moynihan, E., Griffiths, B.S., Hillier, S., Owen, J., Pendlowski, H. and Avery, L.M., 2014. Clay mineral type effect on bacterial enteropathogen survival in soil. Science of the Total Environment, 468, pp.302-305.
British Geological Survey, United Kingdom Soil Observatory, Map Viewer, Available online: http://mapapps2.bgs.ac.uk/ukso/home.html [Accessed: 20/11/2020].
British Standards Institute (1990), BS 1377 PART 2, Chapter 3.2, Moisture Content Determination, Oven Method.
Bromly, M., Hinz, C., Aylmore, L.A.G., (2007). Relation of dispersivity to properties of homogeneous saturated repacked soil columns. European journal of soil science, 58(1), pp.293-301.
Brumfield, K.D., Cotruvo, J.A., Shanks, O.C., Sivaganesan, M., Hey, J., Hasan, N.A., Huq, A., Colwell, R.R. and Leddy, M.B., 2021. Metagenomic sequencing and quantitative real-time PCR for fecal pollution assessment in an urban watershed. Frontiers in water, 3, p.626849.

Burford, E.P., Fomina, M.,Gadd, G.M., (2003). Fungal involvement in bioweathering and biotransformation of rocks and minerals. Mineralogical Magazine, 67,1127-1155.

Burford, E.P., Fomina, M.,Gadd, G.M., (2003). Fungal involvement in bioweathering and biotransformation of rocks and minerals. Mineralogical Magazine, 67,1127-1155.

Burghardt, W., Morel, J.L., Zhang, G.L., (2015). Development of the soil research about urban, industrial, traffic, mining and military areas (SUITMA). Soil Science and Plant Nutrition, 61, 3-21.

Caillon F, Besemer K, Peduzzi P, Schelker J. Soil microbial inoculation during flood events shapes headwater stream microbial communities and diversity. Microb Ecol. 2021 Oct;82(3):591-601. doi: 10.1007/s00248-021-01700-3. Epub 2021 Feb 2. PMID: 33532913; PMCID: PMC8463373.

Cambra-López M., Aarnink A.J.A., Zhao Y., Calvet S., Torres A.G. Airborne particulate matter from livestock production systems: A review of an air pollution problem. Environ. Pollut. 2010;158:1–17. doi: 10.1016/j.envpol.2009.07.011

Camper,A.K, J.T. Hayes, P.J. Sturman, W.L. Jones, A.B. Cunningham. Effects of motility and adsorption rate coefficient on transport of bacteria through saturated porous media. Applied and Environmental Microbiology, 59 (10) (1993), pp. 3455-3462.

Campos, L.C. and Darch, G., 2015. Adaptation of UK wastewater infrastructure to climate change. Infrastructure Asset Management, 2(3), pp.97-106.

Cao Y., Shen J., Ran Z.H. Association between Faecalibacterium prausnitzii reduction and inflammatory bowel disease: a meta-analysis and systematic review of the literature. Gastroenterol Res Pract. 2014:2014.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M. and Gormley, N., 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME journal, 6(8), pp.1621-1624.

Casanovas-Massana, A., Pedra, G.G., Wunder Jr, E.A., Diggle, P.J., Begon, M. and Ko, A.I., 2018. Quantification of Leptospira interrogans survival in soil and water microcosms. Applied and environmental microbiology, 84(13), pp.e00507-18.

Casteel, M.J., Sobsey, M.D. and Mueller, J.P., 2006. Fecal contamination of agricultural soils before and after hurricane-associated flooding in North Carolina. Journal of Environmental Science and Health, Part A, 41(2), pp.173-184.

Castro-Ibáñez, I., Gil, M.I., Tudela, J.A. and Allende, A., 2015. Microbial safety considerations of flooding in primary production of leafy greens: A case study. Food Research International, 68, pp.62-69.

Chandler, D.M. and Lerner, D.N., 2015. A low cost method to detect polluted surface water outfalls and misconnected drainage. Water and Environment Journal, 29(2), pp.202-206.

Chao, A., and Lee, S.-M. “Estimating the number of classes via sample coverage.” Journal of the American statistical Association 87.417 (1992): 210-217.

Chen J, Du Y, Zhu W, Pang X, Wang Z. Effects of organic materials on soil bacterial community structure in long-term continuous cropping of tomato in greenhouse. Open Life Sci. 2022 Apr 22;17(1):381-392. doi: 10.1515/biol-2022-0048. PMID: 35528277; PMCID: PMC9034347.

Chernogor, L., Bakhvalova, K., Belikova, A. and Belikov, S., 2022. Isolation and properties of the bacterial strain Janthinobacterium sp. SLB01. Microorganisms, 10(5), p.1071.

Chu, Y., Salles, C., Tournoud, M.G., Got, P., Troussellier, M., Rodier, C. and Caro, A., 2011. Faecal bacterial loads during flood events in Northwestern Mediterranean coastal rivers. Journal of Hydrology, 405(3-4), pp.501-511.

Colin, R., Drescher, K., Sourjik, V., (2019). Chemotactic behaviour of Escherichia coli at high cell density. Nature communications, 10, 1-11.

Collado, L. and Figueras, M.J., 2011. Taxonomy, epidemiology, and clinical relevance of the genus Arcobacter. Clinical microbiology reviews, 24(1), pp.174-192.

Collender, P.A., Cooke, O.C., Bryant, L.D., Kjeldsen, T.R. , Remais, J.V., (2016). Estimating the microbiological risks associated with inland flood events: Bridging theory and models of pathogen transport. Critical reviews in environmental science and technology, 46(23-24), pp.1787-1833.

Colley, T, 2022, ‘National scandal’: Failing monitoring systems found at 50% of England’s sewer overflows, ENDS Report, 22nd August 2022.https://www.endsreport.com/article/1796320/national-scandal-failing-monitoring-systems-found-50-englands-sewer-overflows

Colley, T, 2022, ‘National scandal’: Failing monitoring systems found at 50% of England’s sewer overflows, ENDS Report, 22nd August 2022.https://www.endsreport.com/article/1796320/national-scandal-failing-monitoring-systems-found-50-englands-sewer-overflows

Collingro A, Köstlbacher S, Horn M. Chlamydiae in the Environment. Trends Microbiol. 2020 Nov;28(11):877-888. doi: 10.1016/j.tim.2020.05.020. Epub 2020 Jun 23. PMID: 32591108.

Córdova, P., Rivera-González, J.P., Rojas-Martínez, V., Villarreal, P., Zamorano, A., Fiore, N., San Martín, D., Vera, F., Gálvez, E., Romero, J. and Barrueto, J., 2022. Antimicrobial Multiresistant Phenotypes of Genetically Diverse Pseudomonas spp. Isolates Associated with Tomato Plants in Chilean Orchards. Horticulturae, 8(8), p.750.

Crabtree, B., 2003. CSO Pollution Control Policies and Procedures in the UK. In Urban Water Management: Science Technology and Service Delivery (pp. 191-202). Dordrecht: Springer Netherlands.

Crost, E.H., Coletto, E., Bell, A. and Juge, N., 2023. Ruminococcus gnavus: friend or foe for human health. FEMS Microbiology Reviews, 47(2), p.fuad014.

Cruz-Martínez KRosling AZhang Y, Song M, Andersen GL, Banfield JF 2012. Effect of Rainfall-Induced Soil Geochemistry Dynamics on Grassland Soil Microbial Communities. Appl Environ Microbiol 78:.https://doi.org/10.1128/AEM.00203-12

Cuadros, J.,(2017). Clay minerals interaction with microorganisms: a review. Clay Minerals, 52, 235-261.

Cuadros, J.,(2017). Clay minerals interaction with microorganisms: a review. Clay Minerals, 52, 235-261.

Cunningham, A.B., Sharp, R.R., Caccavo Jr, F. and Gerlach, R., 2007. Effects of starvation on bacterial transport through porous media. Advances in water resources, 30(6-7), pp.1583-1592.

D.E. Hill, J.-Y. Parlange, Wetting front instability in layered soils, Soil Sci. Soc. Am. Proc., 36 (1972), pp. 692-702.

De Man, H.H.H.J.L., Van Den Berg, H.H.J.L., Leenen, E.J.T.M., Schijven, J.F., Schets, F.M., Van Der Vliet, J.C., Van Knapen, F. and de Roda Husman, A.M., 2014. Quantitative assessment of infection risk from exposure to waterborne pathogens in urban floodwater. Water research, 48, pp.90-99.

DEFRA (2014), Department for Environment, Food, and Rural Affairs, Water Framework Directive implementation in England and Wales: new and updated standards to protect the water environment, Available online: www.gov.uk/government/publications , [Accessed:16/08/2021].

Deng, J., Yin, Y., Zhu, W. and Zhou, Y., 2018. Variations in soil bacterial community diversity and structures among different revegetation types in the Baishilazi Nature Reserve. Frontiers in Microbiology, 9, p.2874.

Department for Energy, Security and Net Zero, 2024, National Statistics Energy Trends, 2021 Data, 2022 Data, Energy trends - GOV.UK (www.gov.uk).

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P. and Andersen, G.L., 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Applied and environmental microbiology, 72(7), pp.5069-5072.

Detectronic, (2020), REMOTE RAINFALL MONITORINGMSFM-RAIN/GPRS data sheet. Available online: https://detectronic.org/wp-content/uploads/2020/01/MMS-D005-raingauge_data_sheet-iss2.pdf [Accessed: 18/10/2022].

Devane, M.L., Moriarty, E., Weaver, L., Cookson, A. and Gilpin, B., 2020. Fecal indicator bacteria from environmental sources; strategies for identification to improve water quality monitoring. Water Research, 185, p.116204.

Diaper, J.P., Tither, K. and Edwards, C., 1992. Rapid assessment of bacterial viability by flow cytometry. Applied Microbiology and Biotechnology, 38, pp.268-272.

Divine, C.E., McDonnell, J.J. The future of applied tracers in hydrogeology. Hydrogeol J 13, 255–258 (2005). https://doi.org/10.1007/s10040-004-0416-3

Dolata, M., Moore, R. and Orman, N., 2013. Sewer Misconnections: What is the True Non-agricultural Diffuse Water Pollution Impact?. UK Water Industry Research Limited.

Downes, J., Munson, M. and Wade, W.G., 2003. Dialister invisus sp. nov., isolated from the human oral cavity. International journal of systematic and evolutionary microbiology, 53(6), pp.1937-1940.

Drinking Water Inspectorate, The Water Supply (Water Quality) (Amendment) Regulations 2018 (SI No. 706), https://www.dwi.gov.uk/water-companies/guidance-and-codes-of-practice/guidance-on-implementing-the-water-supply-water-quality-regulations/ .

Du, M., Wang, L., Ebrahimi, A., Chen, G., Shu, S., Zhu, K., Shen, C., Li, B. and Wang, G., 2021. Extracellular polymeric substances induced cell-surface interactions facilitate bacteria transport in saturated porous media. Ecotoxicology and Environmental Safety, 218, p.112291.

Dunsmore, B.C., Bass, C.J. and Lappin-Scott, H.M., 2004. A novel approach to investigate biofilm accumulation and bacterial transport in porous matrices. Environmental Microbiology, 6(2), pp.183-187.

Duranti, S., Ruiz, L., Lugli, G.A., Tames, H., Milani, C., Mancabelli, L., Mancino, W., Longhi, G., Carnevali, L., Sgoifo, A. and Margolles, A., 2020. Bifidobacterium adolescentis as a key member of the human gut microbiota in the production of GABA. Scientific reports, 10(1), p.14112.

Dzodzomenyo, M., Asamoah, M., Li, C., Kichana, E. and Wright, J., 2022. Impact of flooding on microbiological contamination of domestic water sources: a longitudinal study in northern Ghana. Applied Water Science, 12(10), p.235.

Economy, L.M., Wiegner, T.N., Strauch, A.M., Awaya, J.D. and Gerken, T., 2019. Rainfall and streamflow effects on estuarine Staphylococcus aureus and fecal indicator bacteria concentrations. Journal of Environmental Quality, 48(6), pp.1711-1721.

Edberg, S.C.L., Rice, E.W., Karlin, R.J. and Allen, M.J., 2000. Escherichia coli: the best biological drinking water indicator for public health protection. Journal of applied microbiology, 88(S1), pp.106S-116S.
Eilers, K.G., Debenport, S., Anderson, S. and Fierer, N., 2012. Digging deeper to find unique microbial communities: the strong effect of depth on the structure of bacterial and archaeal communities in soil. Soil Biology and Biochemistry, 50, pp.58-65.

Ekhlas, D., Kurisu, F., Kasuga, I., Cernava, T., Berg, G., Liu, M. and Furumai, H., 2021. Identification of new eligible indicator organisms for combined sewer overflow via 16S rRNA gene amplicon sequencing in Kanda River, Tokyo. Journal of Environmental Management, 284, p.112059.

Ellis, J.B. and Butler, D., 2015. Surface water sewer misconnections in England and Wales: Pollution sources and impacts. Science of the Total Environment, 526, pp.98-109.

Elsaghir, H. and Reddivari, A.K.R., 2022. Bacteroides fragilis. In StatPearls [Internet]. StatPearls Publishing.

England, L.S., Lee, H., Trevors. J.T., (1993), Bacterial survival in soil: effect of clays and protozoa, Soil Biology and Biochemistry, 25, 525-531.
Environment Agency, (2019) Surface-water long term flood risk map, Available online: https://flood-warning-information.service.gov.uk/long-term-flood-risk/map ,[Accessed: 18/05/2021].
Environment Agency, 2018, Storm Overflow Assessment Framework, June 2018, Version 1.6, https://www.water.org.uk/wp-content/uploads/2018/12/SOAF.pdf

Environment Agency, 2023, Press Release:Environment Agency publishes Event Duration Monitoring data for 2022, https://www.gov.uk/government/news/environment-agency-publishes-event-duration-monitoring-data-for-2022

European Commission, Council Directive 91/271/EEC of 21 May 1991 concerning urban waste-water treatment, Article 130 S, L 135 , 30 May 1991 P. 40 – 52.

European Union (2006), Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC (OJ L 64, 4.3.2006, pp. 37-51).

Evans, A.E., Mateo-Sagasta, J., Qadir, M., Boelee, E. and Ippolito, A., 2019. Agricultural water pollution: key knowledge gaps and research needs. Current opinion in environmental sustainability, 36, pp.20-27.

Faith DP (1992) Conservation evaluation and phylogenetic diversity. Biol Cons 61(1):1–10

Fang, H., Vergara, G.G.R., Goh, S.G., Ang, C.Y.L., Gu, X. and Gin, K.Y.H., 2018. Effect of rainfall on the microbial water quality of a tropical urban catchment. Journal of environmental quality, 47(5), pp.1242-1248.

Fennelly K.P., Martyny J.W., Fulton K.E., Orme I.M., Cave D.M., Heifets L.B. Cough generated aerosols of Mycobacterium tuberculosis: A new method to study infectiousness. Am. J. Respir. Crit. Care Med. 2004;169:604–609. doi: 10.1164/rccm.200308-1101OC.

Fera, M.T., Maugeri, T.L., Gugliandolo, C., Beninati, C., Giannone, M., La Camera, E. and Carbone, M., 2004. Detection of Arcobacter spp. in the coastal environment of the Mediterranean Sea. Applied and Environmental Microbiology, 70(3), pp.1271-1276.

Ferguson, A.S., Layton, A.C., Mailloux, B.J., Culligan, P.J., Williams, D.E., Smartt, A.E., Sayler, G.S., Feighery, J., McKay, L.D., Knappett, P.S. and Alexandrova, E., 2012. Comparison of fecal indicators with pathogenic bacteria and rotavirus in groundwater. Science of the Total Environment, 431, pp.314-322.

Fewtrell, L., Kay, D., Watkins, J., Davies, C., Francis, C., (2011). The microbiology of urban UK floodwaters and a quantitative microbial risk assessment of flooding and gastrointestinal illness. Journal of flood risk management, 4, 77-87.

Fierer, N., Schimel, J.P. and Holden, P.A., 2003. Influence of drying–rewetting frequency on soil bacterial community structure. Microbial ecology, 45, pp.63-71.

Flemming HC, Wingender J 2010. The biofilm matrix. Nat Rev Microbiol 8: 623–633.

Fomina, M., Skorochod, I., (2020). Microbial Interaction with Clay Minerals and Its Environmental and Biotechnological Implications. Minerals, 10,861.

Förster, S., Kuhlmann, B., Lindenschmidt, K.-E., and Bronstert, A.: Assessing flood risk for a rural detention area, Nat. Hazards Earth Syst. Sci., 8, 311–322, https://doi.org/10.5194/nhess-8-311-2008, 2008.

Foulquier, A., Volat, B., Neyra, M., Bornette, G. and Montuelle, B., 2013. Long-term impact of hydrological regime on structure and functions of microbial communities in riverine wetland sediments. FEMS Microbiology Ecology, 85(2), pp.211-226.

Francioli, D., Cid, G., Kanukollu, S., Ulrich, A., Hajirezaei, M.R. and Kolb, S., 2021. Flooding causes dramatic compositional shifts and depletion of putative beneficial bacteria on the spring wheat microbiota. Frontiers in Microbiology, 12, p.773116.

Fudou, R., Jojima, Y., Iizuka, T. and Yamanaka, S., 2002. Haliangium ochraceum gen. nov., sp. nov. and Haliangium tepidum sp. nov.: novel moderately halophilic myxobacteria isolated from coastal saline environments. The Journal of general and applied microbiology, 48(2), pp.109-115.

Fukui M, Iwai S, Sakamoto R, Takahashi H, Hayashi T, Kenzaka T. Clostridium paraputrificum Bacteremia in an Older Patient with No Predisposing Medical Condition. Intern Med. 2017 Dec 15;56(24):3395-3397. doi: 10.2169/internalmedicine.8164-16. Epub 2017 Sep 25. PMID: 28943535; PMCID: PMC5790735.
Fukui M, Iwai S, Sakamoto R, Takahashi H, Hayashi T, Kenzaka T. Clostridium paraputrificum Bacteremia in an Older Patient with No Predisposing Medical Condition. Intern Med. 2017 Dec 15;56(24):3395-3397. doi: 10.2169/internalmedicine.8164-16. Epub 2017 Sep 25. PMID: 28943535; PMCID: PMC5790735.

Furtak, K., Grządziel, J., Gałązka, A. and Niedźwiecki, J., 2020. Prevalence of unclassified bacteria in the soil bacterial community from floodplain meadows (fluvisols) under simulated flood conditions revealed by a metataxonomic approachss. Catena, 188, p.104448.
Gang Wang and Xing Wei. 2015. Modeling swelling–shrinkage behavior of compacted expansive soils during wetting–drying cycles. Canadian Geotechnical Journal. 52(6): 783-794. https://doi.org/10.1139/cgj-2014-0059.

Gannon, J.T., Manilal, V.B. and Alexander, M., 1991. Relationship between cell surface properties and transport of bacteria through soil. Applied and environmental microbiology, 57(1), pp.190-193.

Gelsomino, A., Badalucco, L., Ambrosoli, R., Crecchio, C., Puglisi, E. and Meli, S.M., 2006. Changes in chemical and biological soil properties as induced by anthropogenic disturbance: A case study of an agricultural soil under recurrent flooding by wastewaters. Soil Biology and Biochemistry, 38(8), pp.2069-2080.

Gerba, C.P., (1983). Methods for recovering viruses from the water environment, 19, 20-22, CRC Press, Boca Raton.
Gerlach, R. and Cunningham, A.B., 2010. Influence of biofilms on porous media hydrodynamics. Porous media: applications in biological systems and biotechnology, pp.173-230.
Ghane, E., Ranaivoson, A.Z., Feyereisen, G.W., Rosen, C.J. and Moncrief, J.F., 2016. Comparison of contaminant transport in agricultural drainage water and urban stormwater runoff. Plos one, 11(12), p.e0167834.
Gholami, V., Mohseni Saravi, M., Ahmadi, H., (2010). Effects of impervious surfaces and urban development on runoff generation and flood hazard in the Hajighoshan watershed. Caspian journal of environmental sciences, 8, 1-12.

Giakoumis, T. and Voulvoulis, N., 2023. Combined sewer overflows: relating event duration monitoring data to wastewater systems' capacity in England. Environmental Science: Water Research & Technology, 9(3), pp.707-722.

Gilbert, O., Hernández, M., Vilanova, E., Cornellà, O., (2014). Guidelining protocol for soil-column experiments assessing fate and transport of trace organics. Demeau, European Union: Brussels, Belgium.
Gill, J.S., Sivasithamparam, K., Smettem, K.R.J., (2000). Soil types with different texture affects development of Rhizoctonia root rot of wheat seedlings. Plant and Soil, 221, 113-120.

Goñi-Urriza, M., Capdepuy, M., Raymond, N., Quentin, C. and Caumette, P., 1999. Impact of an urban effluent on the bacterial community structure in the Arga River (Spain), with special reference to culturable Gram-negative rods. Canadian journal of microbiology, 45(10), pp.826-832.

González Macé, O., Steinauer, K., Jousset, A., Eisenhauer, N. and Scheu, S., 2016. Flood-induced changes in soil microbial functions as modified by plant diversity. PLoS One, 11(11), p.e0166349.

Good C, Davidson J, Wiens GD, Welch TJ, Summerfelt S. Flavobacterium branchiophilum and F. succinicans associated with bacterial gill disease in rainbow trout Oncorhynchus mykiss (Walbaum) in water recirculation aquaculture systems. J Fish Dis. 2015 Apr;38(4):409-13. doi: 10.1111/jfd.12249. Epub 2014 Apr 10. PMID: 24720801; PMCID: PMC4406145.

Goodfellow, I., Taube, S., (2016). Calicivirus replication and reverse genetics, Viral Gastroenteritis, Academic Press, 355-378.

Gordon, H., Haygarth, P.M. and Bardgett, R.D., 2008. Drying and rewetting effects on soil microbial community composition and nutrient leaching. Soil Biology and Biochemistry, 40(2), pp.302-311.

Gordon, N.S., Valenzuela, A., Adams, S.M., Ramsey, P.W., Pollock, J.L., Holben, W.E. and Gannon, J.E., 2009. Pedobacter nyackensis sp. nov., Pedobacter alluvionis sp. nov. and Pedobacter borealis sp. nov., isolated from Montana flood-plain sediment and forest soil. International journal of systematic and evolutionary microbiology, 59(7), pp.1720-1726.
Gray, S., (2008). Long-term health effects of flooding, Journal of Public Health, 30, 353-354.

Gronewold, A.D., Borsuk, M.E., Wolpert, R.L. and Reckhow, K.H., 2008. An assessment of fecal indicator bacteria-based water quality standards.

Gross A, Azulai N, Oron G, Ronen Z, Arnold M, Nejidat A. Environmental impact and health risks associated with greywater irrigation: a case study. Water Sci Technol. 2005;52(8):161-9. PMID: 16312963.

Gschwend, F., Hartmann, M., Mayerhofer, J., Hug, A.S., Enkerli, J., Gubler, A., Meuli, R.G., Frey, B. and Widmer, F., 2021. Site and land-use associations of soil bacteria and fungi define core and indicative taxa. FEMS Microbiology Ecology, 97(12), p.fiab165.

Guarino, A., Capano, G., Malamisura, B., Alessio, M., Guandalini, S. and Rubino, A., 1987. Production of Escherichia coli STa-like heat-stable enterotoxin by Citrobacter freundii isolated from humans. Journal of clinical microbiology, 25(1), pp.110-114.
Haake, D.A., Levett, P.N., (2015). Leptospirosis in humans, Leptospira and leptospirosis, Springer, Berlin, Heidelberg, 65-97.

Hachad, M., Lanoue, M., Duy, S.V., Villemur, R., Sauvé, S., Prévost, M. and Dorner, S., 2022. 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. Science of the Total Environment, 811, p.152060.

Hamilton, G.W., Waddington, D.V., (1999) Infiltration rates on residential lawns in central Pennsylvania, Journal of soil and water conservation, 54, 564-568.

Hammes, F., Velten, S., Egli, T. and Juhna, T., 2011. 6.41 Biotreatment of Drinking Water. Reference Module in Earth Systems and Environmental Sciences-Comprehensive Biotechnology, pp.517-530.

Han YW. Fusobacterium nucleatum: a commensal-turned pathogen. Curr Opin Microbiol. 2015 Feb;23:141-7. doi: 10.1016/j.mib.2014.11.013. Epub 2015 Jan 8. PMID: 25576662; PMCID: PMC4323942

Han, X. Shen, H. Yang, H. Kim, M. Tong. Influence of nutrient conditions on the transport of bacteria in saturated porous media. Colloids Surf. B Biointerfaces, 102 (1) (2013), pp. 752-758

Haskoning, R., 2010. Non-agricultural diffuse pollution: evidence base update. Report 9V6524 for the Environment Agency. Haywards Heath, W Sussex, UK: Royal Haskoning (Environment).

Hautojärvi, A., Andersson, P. and Volckaert, G., 1997. What Has Been Learned from Field Tracer Transport Experiments–A Critical Overview. NEA/AEN EC/CE, p.63.

Hayashi, H., Shibata, K., Sakamoto, M., Tomita, S. and Benno, Y., 2007. Prevotella copri sp. nov. and Prevotella stercorea sp. nov., isolated from human faeces. International journal of systematic and evolutionary microbiology, 57(5), pp.941-946.

Hedlund, Brian P., John J. Gosink, and James T. Staley. "Phylogeny of Prosthecobacter, the Fusiform Caulobacters: Members of a Recently Discovered Division of the Bacteria." International Journal of Systematic Bacteriology. October 1996;46(4):960-966

Hedrick, D. B., A. Peacock, J. R. Stephen, S. J. MacNaughton, J. Brüggemann, and D. C. White. 2000.“Measuring Soil Microbial Community Diversity Using Polar Lipid Fatty Acid and Denatured Gradient Gel Electrophoresis Data,” Journal of Microbiological Methods 41: 235–48. PMID 10958969.www.ncbi.nlm.nih.gov/pubmed/10958969.

Hill, D.D., Owens, W.E. and Tchounwou, P.B., 2006. The impact of rainfall on fecal coliform bacteria in Bayou Dorcheat (North Louisiana). International Journal of Environmental Research and Public Health, 3(1), pp.114-117.
Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H. and Kanae, S., 2013. Global flood risk under climate change. Nature climate change, 3(9), pp.816-821.

Höfte, M. and De Vos, P., 2007. Plant pathogenic Pseudomonas species. In Plant-associated bacteria (pp. 507-533). Dordrecht: Springer Netherlands.

Holcomb DA, Stewart JR. Microbial Indicators of Fecal Pollution: Recent Progress and Challenges in Assessing Water Quality. Curr Environ Health Rep. 2020 Sep;7(3):311-324. doi: 10.1007/s40572-020-00278-1. PMID: 32542574; PMCID: PMC7458903.

Hong, Y. and Brown, D.G., 2008. Electrostatic behavior of the charge-regulated bacterial cell surface. Langmuir, 24(9), pp.5003-5009.

Hooper, L. V., T. Midtvedt, and J. I. Gordon. 2002. How host-microbial interactions shape the nutrient environment of the mammalian intestine. Annu. Rev. Nutr. 22:283-307.

Hosokawa, M., Endoh, T., Kamata, K., Arikawa, K., Nishikawa, Y., Kogawa, M., Saeki, T., Yoda, T. and Takeyama, H., 2022. Strain-level profiling of viable microbial community by selective single-cell genome sequencing. Scientific reports, 12(1), p.4443.

Howard, A., O'Donoghue, M., Feeney, A. and Sleator, R.D., 2012. Acinetobacter baumannii: an emerging opportunistic pathogen. Virulence, 3(3), pp.243-250.
Hsu, B.-M., Huang, C. P., Pan, J. R. (2001) Filtration behaviors of Giardia And Cryptosporidiums ionic strength and pH effects, Water Research, 35, 3777-3782

Hu, Y.O.O., Ndegwa, N., Alneberg, J. et al. Stationary and portable sequencing-based approaches for tracing wastewater contamination in urban stormwater systems. Sci Rep 8, 11907 (2018). https://doi.org/10.1038/s41598-018-29920-7

Huy, H., Jin, L., Lee, Y.K., Lee, K.C., Lee, J.S., Yoon, J.H., Ahn, C.Y. and Oh, H.M., 2013. Arenimonas daechungensis sp. nov., isolated from the sediment of a eutrophic reservoir. International journal of systematic and evolutionary microbiology, 63(Pt_2), pp.484-489.

Huysman, F. and Verstraete, W., 1993. Water-facilitated transport of bacteria in unsaturated soil columns: influence of cell surface hydrophobicity and soil properties. Soil Biology and Biochemistry, 25(1), pp.83-90.

Illumina, (2020), 16S Metagenomic Sequencing Library Preparation, Part # 15044223 Rev. B, Available online:
https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf [Accessed: 12/08/2021].
Ishii, S. and Sadowsky, M.J., 2008. Escherichia coli in the environment: implications for water quality and human health. Microbes and environments, 23(2), pp.101-108.
Islam MA, Ikeguchi A, Naide T. Concentrations of Aerosol Numbers and Airborne Bacteria, and Temperature and Relative Humidity, and Their Interrelationships in a Tie-Stall Dairy Barn. Animals (Basel). 2019 Nov 24;9(12):1023. doi: 10.3390/ani9121023. PMID: 31771315; PMCID: PMC6941166.

Jamieson, R.C., Gordon, R.J., Sharples, K.E., Stratton, G.W. and Madani, A., 2002. Movement and persistence of fecal bacteria in agricultural soils and subsurface drainage water: a review. Canadian biosystems engineering, 44(1), pp.1-9.

Jan Lorenz Soliman, Alex Dekhtyar, Jennifer Vanderkellen, Aldrin Montana, Michael Black, Emily Neal, Kevin Webb, Chris Kitts, and Anya Goodman. 2012. Microbial source tracking by molecular fingerprinting. In Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine (BCB '12). Association for Computing Machinery, New York, NY, USA, 617–619. https://doi.org/10.1145/2382936.2383042

Janna Pietikäinen, Marie Pettersson, Erland Bååth, Comparison of temperature effects on soil respiration and bacterial and fungal growth rates, FEMS Microbiology Ecology, Volume 52, Issue 1, March 2005, Pages 49–58, https://doi.org/10.1016/j.femsec.2004.10.002

Janssen, P.H., 2015. Opitutus. Bergey's Manual of Systematics of Archaea and Bacteria, pp.1-3.

Jarvis, N. and Larsbo, M., 2023. Macropores and macropore flow.Encyclopedia of soils in the environment, 5, 306-313.

Jayaswal, R.K., Fernandez, M.A. and Schroeder III, R.G., 1990. Isolation and characterization of a Pseudomonas strain that restricts growth of various phytopathogenic fungi. Applied and environmental microbiology, 56(4), pp.1053-1058.

Jefferson KK, Goldmann DA, Pier GB 2005. Use of confocal microscopy to analyze the rate of vancomycin penetration through Staphylococcus aureus biofilms. Antimicrob Agents Chemother 49: 2467–2473.

Jeffery, S., Van der Putten, W.H., (2011). Soil borne human diseases. Luxembourg: publications office of the European Union, 37199, 1-10.
Jeong, J.Y., Park, H.D., Lee, K.H., Weon, H.Y. and Ka, J.O., 2011. Microbial community analysis and identification of alternative host-specific fecal indicators in fecal and river water samples using pyrosequencing. The Journal of Microbiology, 49, pp.585-594.

Jiang, G., 2005. Bacterial transport and deposition in an intact soil lysimeter and packed sand columns (Doctoral dissertation, Lincoln University).

Johnson, J.S., Spakowicz, D.J., Hong, B.Y., Petersen, L.M., Demkowicz, P., Chen, L., Leopold, S.R., Hanson, B.M., Agresta, H.O., Gerstein, M. and Sodergren, E., 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nature communications, 10(1), p.5029.

Jongman, B., Kreibich, H., Apel, H., Barredo, J.I., Bates, P.D., Feyen, L., Gericke, A., Neal, J., Aerts, J.C.J.H., Ward, P.J., (2012). Comparative flood damage model assessment: towards a European approach. Natural Hazards and Earth System Sciences, 12(12), pp.3733-3752.
Joshi NA, Fass JN. (2011). Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.200) [Software]. Available at https://github.com/najoshi/sickle.

Jun Chen, Kyle Bittinger, Emily S. Charlson, Christian Hoffmann, James Lewis, Gary D. Wu, Ronald G. Collman, Frederic D. Bushman, Hongzhe Li, Associating microbiome composition with environmental covariates using generalized Unifrac distances, Bioinformatics, Volume 28, Issue 16, August 2012, Pages 2106–2113, https://doi.org/10.1093/bioinformatics/bts342.

Jung, J. and Park, W., 2015. Acinetobacter species as model microorganisms in environmental microbiology: current state and perspectives. Applied microbiology and biotechnology, 99, pp.2533-2548.

Jurkevitch, E., 2012. Isolation and classification of Bdellovibrio and like organisms. Current protocols in microbiology, 26(1), pp.7B-1.

Kallscheuer, N., Wiegand, S., Boedeker, C., Peeters, S.H., Jogler, M., Heuer, A., Jetten, M.S., Rohde, M. and Jogler, C., 2020. Caulifigura coniformis gen. nov., sp. nov., a novel member of the family Planctomycetaceae isolated from a red biofilm sampled in a hydrothermal area. Antonie Van Leeuwenhoek, 113, pp.1927-1937.

Kalmbach, S., Manz, W., Wecke, J. and Szewzyk, U., 1999. Aquabacterium gen. nov., with description of Aquabacterium citratiphilum sp. nov., Aquabacterium parvum sp. nov. and Aquabacterium commune sp. nov., three in situ dominant bacterial species from the Berlin drinking water system. International Journal of Systematic and Evolutionary Microbiology, 49(2), pp.769-777.

Kalyuzhnaya, M.G., Beck, D.A., Suciu, D., Pozhitkov, A., Lidstrom, M.E. and Chistoserdova, L., 2010. Functioning in situ: gene expression in Methylotenera mobilis in its native environment as assessed through transcriptomics. The ISME Journal, 4(3), pp.388-398.

Kämpfer, P., Lodders, N., Huber, B., Falsen, E. and Busse, H.J., 2008. Deinococcus aquatilis sp. nov., isolated from water. International journal of systematic and evolutionary microbiology, 58(12), pp.2803-2806.

Karimi, B., Terrat, S., Dequiedt, S., Saby, N.P., Horrigue, W., Lelièvre, M., Nowak, V., Jolivet, C., Arrouays, D., Wincker, P. and Cruaud, C., 2018. Biogeography of soil bacteria and archaea across France. Science advances, 4(7), p.eaat1808.

Karunakaran, E., Battarbee, R., Tait, S., Brentan, B.M., Berney, C., Grinham, J., Herrero, M.A., Omolo, R. and Douterelo, I., 2024. Integrating molecular microbial methods to improve faecal pollution management in rivers with designated bathing waters. Science of The Total Environment, 912, p.168565.

Karunakaran, E., Vernon, D., Biggs, C.A., Saul, A., Crawford, D. and Jensen, H., (2016). Enumeration of sulphate-reducing bacteria for assessing potential for hydrogen sulphide production in urban drainage systems. Water Science and Technology, 73(12), pp.3087-3094.
Kasel, D., Bradford, S.A., Šimůnek, J., Heggen, M., Vereecken, H., Klumpp, E., (2013). Transport and retention of multi-walled carbon nanotubes in saturated porous media: Effects of input concentration and grain size. Water research, 47, 933-944.

Kayikcioglu, H.H., 2012. Short-term effects of irrigation with treated domestic wastewater on microbiological activity of a Vertic xerofluvent soil under Mediterranean conditions. Journal of environmental management, 102, pp.108-114.

Keswick, B.H., Wang, D.S. and Gerba, C.P., 1982. The use of microorganisms as ground-water tracers: A review. Groundwater, 20(2), pp.142-149.

Kim, Y. Hong, I. Lee, S.A. Bradford, S.L. Walker. Surface characteristics and adhesion behavior of Escherichia coli O157:H7: role of extracellular macromolecules. Biomacromolecules, 10 (9) (2009), pp. 2556-2564

Kirkham, M.B., 2023. Principles of soil and plant water relations. Elsevier.

Kritzberg, E. and Bååth, E., 2022. Seasonal variation in temperature sensitivity of bacterial growth in a temperate soil and lake. FEMS Microbiology Ecology, 98(10), p.fiac111.

Kumar, M.S., Philominathan, P., (2010). The physics of flagellar motion of E. coli during chemotaxis. Biophysical reviews, 2,13-20.

Kumru, S., Tekedar, H.C., Blom, J., Lawrence, M.L. and Karsi, A., 2020. Genomic diversity in flavobacterial pathogens of aquatic origin. Microbial pathogenesis, 142, p.104053.
L. W. Sinton , M. J. Noonan , R. K. Finlay , L. Pang & M. E. Close (2000) Transport and attenuation of bacteria and bacteriophages in an alluvial gravel aquifer, New Zealand Journal of Marine and Freshwater Research, 34:1, 175-186, DOI: 10.1080/00288330.2000.9516924

La Reau, A.J. and Suen, G., 2018. The Ruminococci: key symbionts of the gut ecosystem. Journal of microbiology, 56, pp.199-208.
La Rosa, G. and Muscillo, M., 2013. Molecular detection of viruses in water and sewage. In Viruses in Food and Water (pp. 97-125). Woodhead Publishing.

Labbe, R.G., 1991. Clostridium perfringens. Journal of the Association of Official Analytical Chemists, 74(4), pp.711-714.

Lamy, E., Lassabatere, L., Bechet, B.,Andrieu, H., (2013). Effect of a nonwoven geotextile on solute and colloid transport in porous media under both saturated and unsaturated conditions. Geotextiles and Geomembranes, 36, 55-65.
Lance, J.C., Gerba, C.P., (1984). Virus movement in soil during saturated and unsaturated flow. Applied and Environmental Microbiology, 47, 335-337.
Landry, E.F., Vaughn, J.M., Thomas, M.Z., Beckwith, C.A., (1979). Adsorption of enteroviruses to soil cores and their subsequent elution by artificial rainwater. Applied and environmental microbiology, 38, 680-687.

La-Sa, A., Gao, G.L., Ding, G.D., Zhang, Y., Cao, H.Y. and Yu-Jia, D.U., 2022. Eco-physiological functions of soil microbial biofilms: A review. Ying yong sheng tai xue bao= The journal of applied ecology, 33(7), pp.1885-1892.

Leal J, Gregson DB, Ross T, Church DL, Laupland KB. Epidemiology of Clostridium species bacteremia in Calgary, Canada, 2000-2006. J Infect 57: 198-203, 2008.

Lee, S.A., Kim, T.W., Sang, M.K., Song, J., Kwon, S.W. and Weon, H.Y., 2021. Arachidicoccus soli sp. nov., a bacterium isolated from soil. International Journal of Systematic and Evolutionary Microbiology, 71(1), p.004566.
Leight, A.K. and Hood, R.R., 2018. Precipitation thresholds for fecal bacterial indicators in the Chesapeake Bay. Water research, 139, pp.252-262.
Levin, P.A., Angert, E.R., (2015). Small but mighty: cell size and bacteria. Cold Spring Harbor perspectives in biology, 7(7), p.a019216.

Lewis K 2005. Persister cells and the riddle of biofilm survival. Biochemistry (Mosc) 70: 267–274

Lewis, J., Sjöstrom, J., (2010). Optimizing the experimental design of soil columns in saturated and unsaturated transport experiments. Journal of contaminant hydrology, 115,1-13.

Li, E., Saleem, F., Edge, T.A. and Schellhorn, H.E., 2021. Biological indicators for fecal pollution detection and source tracking: A review. Processes, 9(11), p.2058.

Li, H.Q., Li, H., Zhou, X.Y., Shen, Y.J. and Su, J.Q., 2021. Distinct patterns of abundant and rare subcommunities in paddy soil during wetting–drying cycles. Science of the Total Environment, 785, p.147298.

Lipps, S.M. and Samac, D.A., 2022. Pseudomonas viridiflava: An internal outsider of the Pseudomonas syringae species complex. Molecular plant pathology, 23(1), pp.3-15.

Liu, X., Gao, C., Ji, D., Walker, S.L., Huang, Q., Cai, P., (2017). Survival of Escherichia coli O157: H7 in various soil particles: importance of the attached bacterial phenotype. Biology and Fertility of Soils, 53, 209-219.

Liu, Y., Cheng, J., Liu, X., Zhong, H., Wang, B., Kong, Z. and Wu, L., 2021. Tracking the changes of wetland soil bacterial community and metabolic potentials under drought and flooding conditions in experimental microcosms. Journal of Soils and Sediments, 21, pp.2404-2417.

Loch, T.P. and Faisal, M., 2015. Emerging flavobacterial infections in fish: a review. Journal of advanced research, 6(3), pp.283-300.

Lopez-Siles, M., Duncan, S.H., Garcia-Gil, L.J. and Martinez-Medina, M., 2017. Faecalibacterium prausnitzii: from microbiology to diagnostics and prognostics. The ISME journal, 11(4), pp.841-852.

Lovley, D.R., Ueki, T., Zhang, T., Malvankar, N.S., Shrestha, P.M., Flanagan, K.A., Aklujkar, M., Butler, J.E., Giloteaux, L., Rotaru, A.E. and Holmes, D.E., 2011. Geobacter: the microbe electric's physiology, ecology, and practical applications. Advances in microbial physiology, 59, pp.1-100.

Lozupone C, Knight R . (2005). Unifrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol 71: 8228–823

Luo, X., Wang, M.K., Hu, G. and Weng, B., 2019. Seasonal change in microbial diversity and its relationship with soil chemical properties in an orchard. PLoS One, 14(12), p.e0215556.

MADIGAN, M.T., 2005. Brock biology of microorganisms, 11th edn.

Mahto, K.U., Priyadarshanee, M., Samantaray, D.P. and Das, S., 2022. Bacterial biofilm and extracellular polymeric substances in the treatment of environmental pollutants: Beyond the protective role in survivability. Journal of Cleaner Production, 379, p.134759

Makk, J., Homonnay, Z.G., Keki, Z., Nemes-Barnas, K., Marialigeti, K., Schumann, P. and Toth, E.M., 2015. Arenimonas subflava sp. nov., isolated from a drinking water network, and emended description of the genus Arenimonas. International Journal of Systematic and Evolutionary Microbiology, 65(Pt_6), pp.1915-1921.

Marchandin H, Damay A, Roudière L, Teyssier C, Zorgniotti I, Dechaud H, Jean-Pierre H, Jumas-Bilak E. Phylogeny, diversity and host specialization in the phylum Synergistetes with emphasis on strains and clones of human origin. Res Microbiol. 2010 Mar;161(2):91-100. doi: 10.1016/j.resmic.2009.12.008. Epub 2010 Jan 15. PMID: 20079831.

Mark F. Ditmar,CHAPTER 11 - Infectious Diseases,Editor(s): Richard A. Polin, Mark F. Ditmar,Pediatric Secrets (Fifth Edition),Mosby,2011,Pages 354-422,ISBN 9780323065610,https://doi.org/10.1016/B978-0-323-06561-0.00011-2.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal, 17(1), pp.10-12.

Masipan, T., Chotpantarat, S. and Boonkaewwan, S., 2016. Experimental and modelling investigations of tracer transport in variably saturated agricultural soil of Thailand: Column study. Sustainable Environment Research, 26(2), pp.97-101.

Matches, J.R., Liston, J. and Curran, D., 1974. Clostridium perfringens in the environment. Applied microbiology, 28(4), pp.655-660.

Mattingly AE, Weaver AA, Dimkovikj A, Shrout JD. Assessing Travel Conditions: Environmental and Host Influences On Bacterial Surface Motility. J Bacteriol. 2018 Mar 19;200(11):e00014-18. doi: 10.1128/JB.00014-18. Epub ahead of print. PMID: 29555698; PMCID: PMC5952383.

Mawdsley, J.L., Brooks, A.E., Merry, R.J., Pain, B.F., (1996). Use of a novel soil tilting table apparatus to demonstrate the horizontal and vertical movement of the protozoan pathogen Cryptosporidium parvum in soil. Biology and Fertility of Soils, 23, 215-220.

McClaine, J.W. Ford., R.M. Characterizing the adhesion of motile and nonmotile Escherichia coli to a glass surface using a parallel-plate flow chamber.Biotechnology and Bioengineering, 78 (2) (2002), pp. 179-189.

McDonald JH. Handbook of Biolological Statistics. Third Edition. Baltimore, Maryland, U.S.A: Sparky House Publishing; 2014. University of Delaware.

Merz, B., Aerts, J.C.J.H., Arnbjerg-Nielsen, K., Baldi, M., Becker, A., Bichet, A., Blöschl, G., Bouwer, L.M., Brauer, A., Cioffi, F. and Delgado, J.M., 2014. Floods and climate: emerging perspectives for flood risk assessment and management. Natural Hazards and Earth System Sciences, 14(7), pp.1921-1942.

Mesibov, R., Adler, J., (1972). Chemotaxis toward amino acids in Escherichia coli. Journal of Bacteriology, 112, 315-326.

Mhuantong, W., Wongwilaiwalin, S., Laothanachareon, T., Eurwilaichitr, L., Tangphatsornruang, S., Boonchayaanant, B., Limpiyakorn, T., Pattaragulwanit, K., Punmatharith, T., McEvoy, J. and Khan, E., 2015. Survey of microbial diversity in flood areas during Thailand 2011 flood crisis using high-throughput tagged amplicon pyrosequencing. PloS one, 10(5), p.e0128043.

Miquel, S., Martin, R., Rossi, O., Bermúdez-Humarán, L.G., Chatel, J.M., Sokol, H., Thomas, M., Wells, J.M. and Langella, P., 2013. Faecalibacterium prausnitzii and human intestinal health. Current opinion in microbiology, 16(3), pp.255-261.

Moran, A.P., Thieken, A.H., Schöbel, A., Rachoy, C., (2010). Documentation of flood damage on railway infrastructure. In Data and Mobility: Transforming Information into Intelligent Traffic and Transportation Services Proceedings of the Lakeside Conference 2010 (pp. 61-70). Springer Berlin Heidelberg.

Moran, S., 2018. An applied guide to water and effluent treatment plant design. Butterworth-Heinemann.

More, T.T., Yadav, J.S.S., Yan, S., Tyagi, R.D. and Surampalli, R.Y., 2014. Extracellular polymeric substances of bacteria and their potential environmental applications. Journal of environmental management, 144, pp.1-25.

Moriarty, E.M., Mackenzie, M.L., Karki, N., Sinton, L.W., (2011). Survival of Escherichia coli, Enterococci, and Campylobacter species. in sheep feces on pastures. Applied and environmental microbiology, 77, 1797-1803.

Morita, Y., Tomida, J. and Kawamura, Y., 2014. Responses of Pseudomonas aeruginosa to antimicrobials. Frontiers in microbiology, 4, p.422.

Morosyuk SV, Cunningham PR, SantaLucia J., Jr Structure and function of the conserved 690 hairpin in Escherichia coli 16 S ribosomal RNA. II. NMR solution structure. J. Mol. Biol. 2001;307(1):197. doi: 10.1006/jmbi.2000.4431.

Muirhead, R.W., Davies-Colley, R.J., Donnison, A.M., Nagels, J.W., (2004). Faecal bacteria yields in artificial flood events: quantifying in-stream stores. Water research, 38, 1215-1224.

Muniesa, M., Jofre, J., García-Aljaro, C. and Blanch, A.R., 2006. Occurrence of Escherichia coli O157: H7 and other enterohemorrhagic Escherichia coli in the environment. Environmental science & technology, 40(23), pp.7141-7149.

Muyzer, G. E., E. C. de Waal, and A. G. Uitterlinden. 1993. “Profiling of Complex Microbial Populations by Denaturing Gradient Gel Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes Coding for 16S rRNA,” Applied and Environmental Microbiology 59: 695–700. PMID 7683183. www.ncbi.nlm.nih.gov/pubmed/7683183.

National Academies of Sciences, Engineering, and Medicine, (2019). Framing the challenge of urban flooding in the United States, National Academies Press.

National Infrastructure Commission, 2022. Reducing the risk of surface water flooding, https://nic.org.uk/studies-reports/reducing-the-risks-of-surface-water-flooding/surface-water-flooding-final-report/

NEOF,2024, https://neof.org.uk/

Nie, K., Ma, K., Luo, W., Shen, Z., Yang, Z., Xiao, M., Tong, T., Yang, Y. and Wang, X., 2021. Roseburia intestinalis: a beneficial gut organism from the discoveries in genus and species. Frontiers in Cellular and Infection Microbiology, 11, p.757718.

Nieder, R., Benbi, D.K., Reichl, F.X., Nieder, R., Benbi, D.K. and Reichl, F.X., 2018. Soil as a transmitter of human pathogens. Soil components and human health, pp.723-827.

O’Toole G, Kaplan HB, Kolter R 2000. Biofilm formation as microbial development. Annu Rev Microbiol 54: 49–79

Odonkor, S.T. and Ampofo, J.K., 2013. Escherichia coli as an indicator of bacteriological quality of water: an overview. Microbiology research, 4(1), p.e2.
O'Donnell, E.C. and Thorne, C.R., 2020. Drivers of future urban flood risk. Philosophical Transactions of the Royal Society A, 378(2168), p.20190216.

Ogwu, M.C., Srinivasan, S., Dong, K., Ramasamy, D., Waldman, B. and Adams, J.M., 2019. Community ecology of Deinococcus in irradiated soil. Microbial ecology, 78, pp.855-872.
Ohl C., Tapsell S.(2000), Flooding and human health. British Medical Journal, 321, 1167–1168.

Okaka FO, Odhiambo BDO. Relationship between Flooding and Out Break of Infectious Diseasesin Kenya: A Review of the Literature. J Environ Public Health. 2018 Oct 17;2018:5452938. doi: 10.1155/2018/5452938. PMID: 30416526; PMCID: PMC6207902.

Olds, H.T., Corsi, S.R., Dila, D.K., Halmo, K.M., Bootsma, M.J. and McLellan, S.L., 2018. High levels of sewage contamination released from urban areas after storm events: A quantitative survey with sewage specific bacterial indicators. PLoS medicine, 15(7), p.e1002614.
Osborn, T. and Maraun, D., (2008). Changing intensity of rainfall over Britain. Climatic Research Unit Information Sheet, 15, 2.
Osborn, T.J., Hulme, M., (2002). Evidence for trends in heavy rainfall events over the UK. Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 360(1796), pp.1313-1325.

Paix, B. and Jacquet, S., 2017. Bdellovibrio and like organisms bacterial predators are not equally distributed in peri-alpine lakes. bioRxiv, p.194688.

Pall, P., Aina, T., Stone, D.A., Stott, P.A., Nozawa, T., Hilberts, A.G., Lohmann, D., Allen, M.R., (2011), Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000, Nature, 470, 382-385

Pang, Z., Raudonis, R., Glick, B.R., Lin, T.J. and Cheng, Z., 2019. Antibiotic resistance in Pseudomonas aeruginosa: mechanisms and alternative therapeutic strategies. Biotechnology advances, 37(1), pp.177-192.

Parsaei, M., Sarafraz, N., Moaddab, S.Y. and Leylabadlo, H.E., 2021. The importance of Faecalibacterium prausnitzii in human health and diseases. New Microbes and New Infections, 43.

Paruch, L., Paruch, A.M., Eiken, H.G. and Sørheim, R., 2019. Faecal pollution affects abundance and diversity of aquatic microbial community in anthropo-zoogenically influenced lotic ecosystems. Scientific reports, 9(1), p.19469.

Passerat, J., Ouattara, N.K., Mouchel, J.M., Rocher, V. and Servais, P., 2011. Impact of an intense combined sewer overflow event on the microbiological water quality of the Seine River. Water research, 45(2), pp.893-903.

Patle, G.T., Sikar, T.T., Rawat, K.S., Singh, S.K., (2019). Estimation of infiltration rate from soil properties using regression model for cultivated land, Geology, Ecology, and Landscapes, 3, 1-13.

Pearson's Correlation Coefficient (2008).. In: Kirch, W. (eds) Encyclopedia of Public Health. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-5614-7_2569.

Peer Timmers , Joost van den Bulk , Leo Heijnen, Edwin Kardinaal , Susan Sollie , Gertjan Medema, 2021, Microbial fingerprint for source tracking in surface water management, Best Practices, Blue Green Deals with Integrated Solutions, Clean Water and Ecosystem Restoration, https://www.amsterdamiww.com/best-practices/microbial-fingerprint-for-source-tracking-in-surface-water-management/

Petersen, H.H., Enemark, H.L., Olsen, A., Amin, M.M., Dalsgaard, A., (2012). Transport of Cryptosporidium parvum oocysts in soil columns following applications of raw and separated liquid slurries. Applied and environmental microbiology, 78, 5994-6000.
Peterson CT, Ward RC (1989), Development of a bacterial transport model for coarse soils. Water Resources Bulletin, 25, 349-357

Petti, C.A., Polage, C.R. and Schreckenberger, P., 2005. The role of 16S rRNA gene sequencing in identification of microorganisms misidentified by conventional methods. Journal of clinical microbiology, 43(12), pp.6123-6125.

Phanuwan, C., Takizawa, S., Oguma, K., Katayama, H., Yunika, A. and Ohgaki, S., 2006. Monitoring of human enteric viruses and coliform bacteria in waters after urban flood in Jakarta, Indonesia. Water Science and Technology, 54(3), pp.203-210.

Phillips, C.A., 2001. Arcobacters as emerging human foodborne pathogens. Food Control, 12(1), pp.1-6.
Phillips, P.J, Chalmers, A.T, Gray, J.L, Kolpin, D.W, Foreman, W.T, Wall, G.R (2012), Combined Sewer Overflows: An Environmental Source of Hormones and Wastewater Micropollutants, Environmental Science & Technology, 46 ,5336-5343.
Pouyat, R.V., Yesilonis, I.D., Dombos, M., Szlavecz, K., Setälä, H., Cilliers, S., Hornung, E., Kotze, D.J., Yarwood, S., (2015). A global comparison of surface soil characteristics across five cities: a test of the urban ecosystem convergence hypothesis. Soil Science, 180, 136-145.
Powers, N.C., Wallgren, H.R., Marbach, S. and Turner, J.W., 2020. Relationship between rainfall, fecal pollution, antimicrobial resistance, and microbial diversity in an urbanized subtropical bay. Applied and environmental microbiology, 86(19), pp.e01229-20.

Prakash, B., Veeregowda, B.M. and Krishnappa, G., 2003. Biofilms: a survival strategy of bacteria. Current science, pp.1299-1307.

Preston, D.R., Farrah, S.R., (1988). Activation thermodynamics of virus adsorption to solids. Applied and environmental microbiology, 54, 2650-2654.
Price, R.G. and Wildeboer, D., 2017. E. coli as an indicator of contamination and health risk in environmental waters. Escherichia coli-Recent Advances on Physiology, Pathogenesis and Biotechnological Applications, p.3.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O., (2012). The SILVA ribosomal RNA gene database project: improved data processing and

Quintana, E.T., Wierzbicka, K., Mackiewicz, P., Osman, A., Fahal, A.H., Hamid, M.E., Zakrzewska-Czerwinska, J., Maldonado, L.A. and Goodfellow, M., 2008. Streptomyces sudanensis sp. nov., a new pathogen isolated from patients with actinomycetoma. Antonie Van Leeuwenhoek, 93, pp.305-313.

R.F. Mueller. Bacterial transport and colonization in low nutrient environments. Water Research, 30 (11) (1996), pp. 2681-2690.

Randle-Boggis, R.J., Ashton, P.D. and Helgason, T., 2018. Increasing flooding frequency alters soil microbial communities and functions under laboratory conditions. MicrobiologyOpen, 7(1), p.e00548.

Rees, G., Pond, K., Johal, K., Pedley, S. and Rickards, A., 1998. Microbiological analysis of selected coastal bathing waters in the UK, Greece, Italy and Spain. Water Research, 32(8), pp.2335-2340.

Revitt, D.M. and Ellis, J.B., 2016. Urban surface water pollution problems arising from misconnections. Science of the Total Environment, 551, pp.163-174.

Robertson, L.J., Campbell, A.T. and Smith, H.V., (1992). Survival of Cryptosporidium parvum oocysts under various environmental pressures. Applied and environmental microbiology, 58(11), pp.3494-3500.

Rodney M. Donlan, Biofilm Formation: A Clinically Relevant Microbiological Process, Clinical Infectious Diseases, Volume 33, Issue 8, 15 October 2001, Pages 1387–1392, https://doi.org/10.1086/322972

Rodzik, J., Furtak, T. and Zglobicki, W., 2009. The impact of snowmelt and heavy rainfall runoff on erosion rates in a gully system, Lublin Upland, Poland. Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group, 34(14), pp.1938-1950.

Roeber R, Bibby D, Bahmani M, Burghardt W (2000), Some features of the soils of old Constance, Proceedings of the First International Conference on Soils of Urban, Industrial, Traffic and Mining Areas,2, 21–26.
Roper, M.M., Marshall, K.C., (1974), Modification of the interaction between Escherichia coli and bacteriophage in saline sediment, Microbial Ecology, 1, 1-13.
Rose, J.B., (1997). Environmental ecology of Cryptosporidium and public health implications. Annual review of public health, 18, 135-161.
Rosenzweig, W.D, Stotzky, G. (1979) Influence of environmental factors on antagonism of fungi by bacteria in soil: clay minerals and pH, Applied Environmental Microbiology, 38,1120-1126.
Rowell, D.L., 2014. Soil science: Methods & applications. Routledge.
Rubinato, M., Nichols, A., Peng, Y., Zhang, J.M., Lashford, C., Cai, Y.P., Lin, P.Z., Tait, S., (2019) Urban and river flooding: Comparison of flood risk management approaches in the UK and China and an assessment of future knowledge needs, Water Science and Engineering, 12, 274-283.

Rui, Y., Fu, D., Do Minh, H., Radhakrishnan, M., Zevenbergen, C. and Pathirana, A., 2018. Urban surface water quality, flood water quality and human health impacts in Chinese cities. What do we know?. Water, 10(3), p.240.

Saingam, P., Li, B. and Yan, T., 2020. Fecal indicator bacteria, direct pathogen detection, and microbial community analysis provide different microbiological water quality assessment of a tropical urban marine estuary. Water Research, 185, p.116280.

Sakthivadivel, R., (1966). Theory and mechanism of filtration of non-paticleal fines through a porous medium, Hydraulic Engineering Laboratory, 110, 5-15, University. of California., Berkeley.

Salter, S; Cox, M; Turek, E; Calus, S; Cookson, W; Moffatt, M; Turner, P; Parkhill, J; Loman, N; Walker, A (2014). "Reagent contamination can critically impact sequence-based microbiome analyses".
Sanderson, R.A., Maas, J.A., Blain, A.P., Gorton, R., Ward, J., O'Brien, S.J., Hunter, P.R., Rushton, S.P., (2018). Spatio-temporal models to determine association between Campylobacter cases and environment. International journal of epidemiology, 47, 202-216.

Santore, M.M., 2022. Interplay of physico-chemical and mechanical bacteria-surface interactions with transport processes controls early biofilm growth: A review. Advances in Colloid and Interface Science, 304, p.102665.

Sao, S., Ann, V., Nishiyama, M., Praise, S. and Watanabe, T., 2023. Tracing the pathways by which flood duration impacts soil bacteria through soil properties and water-extractable dissolved organic matter: A soil column experiment. Science of The Total Environment, 902, p.166524.

Sasha Babicki, David Arndt, Ana Marcu, Yongjie Liang, Jason R. Grant, Adam Maciejewski, and David S. Wishart. Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res. 2016 May 17 (epub ahead of print). doi:10.1093/nar/gkw419

Sauer, E.P., VandeWalle, J.L., Bootsma, M.J. and McLellan, S.L., 2011. Detection of the human specific Bacteroides genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. Water research, 45(14), pp.4081-4091.

Schets, F.M., Van Wijnen, J.H., Schijven, J.F., Schoon, H., de Roda Husman, A.M., (2008). Monitoring of waterborne pathogens in surface waters in Amsterdam, The Netherlands, and the potential health risk associated with exposure to Cryptosporidium and Giardia in these waters. Applied and environmental microbiology, 74, 2069-2078.
Schiffenbauer, M., Stotzky, G., (1982). Adsorption of coliphages T1 and T7 to clay minerals. Applied and Environmental Microbiology, 43, 590-596.
Schirmer, M., Ijaz, U.Z., D'Amore, R., Hall, N., Sloan, W.T. and Quince, C., 2015. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. Nucleic acids research, 43(6), pp.e37-e37.
Schultz, B., 2006. Flood management under rapid urbanisation and industrialisation in flood-prone areas: a need for serious consideration. Irrigation and Drainage: The journal of the International Commission on Irrigation and Drainage, 55(S1), pp.S3-S8.

Schwartz, E., Van Horn, D.J., Buelow, H.N., Okie, J.G., Gooseff, M.N., Barrett, J.E. and Takacs-Vesbach, C.D., 2014. Characterization of growing bacterial populations in McMurdo Dry Valley soils through stable isotope probing with 18O-water. FEMS microbiology ecology, 89(2), pp.415-425.
Scoullos, I.M., Adhikari, S., Vazquez, C.M.L., van de Vossenberg, J., Brdjanovic, D., (2020). Inactivation of indicator organisms on different surfaces after urban floods. Science of the Total Environment, 704, 135456.

Scoullos, I.M., Vazquez, C.M.L., van de Vossenberg, J. and Brdjanovic, D., 2019. Die-off of E. coli as fecal indicator organism on different surfaces after urban floods. Journal of environmental management, 250, p.109516.

Serinaldi, F., Loecker, F., Kilsby, C.G., Bast, H., (2018). Flood propagation and duration in large river basins: a data-driven analysis for reinsurance purposes. Natural Hazards, 94(1), pp.71-92.

Sgorbati, B., Biavati, B. and Palenzona, D., 1995. The genus Bifidobacterium. In The genera of lactic acid bacteria (pp. 279-306). Boston, MA: Springer US.

Shah, A., Shah, S. and Shah, V., 2021. Impact of flooding on the soil microbiota. Environmental Challenges, 4, p.100134

Shanks OC, Newton RJ, Kelty CA, Huse SMSogin ML, McLellan SL. 2013. Comparison of the Microbial Community Structures of Untreated Wastewaters from Different Geographic Locales. Appl Environ Microbiol 79:https://doi.org/10.1128/AEM.03448-12

Shapiro, S.S. and Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). Biometrika, 52(3/4), pp.591-611.

Shen, R., Lan, Z., Rinklebe, J., Nie, M., Hu, Q., Yan, Z., Fang, C., Jin, B. and Chen, J., 2021. Flooding variations affect soil bacterial communities at the spatial and inter-annual scales. Science of the Total Environment, 759, p.143471.
Shiraz, S., Djebbi-Simmons, D., Alhejaili, M., Danos, K., Janes, M., Fontenot, K. and Xu, W., 2020. Evaluation of the microbial safety and quality of Louisiana strawberries after flooding. Food Control, 110, p.106970.

Simpson, E.H. (1949) Measurement of Diversity. Nature, 163, 688.

Singh, K.P., Mohan, D., Sinha, S. and Dalwani, R., 2004. Impact assessment of treated/untreated wastewater toxicants discharged by sewage treatment plants on health, agricultural, and environmental quality in the wastewater disposal area. Chemosphere, 55(2), pp.227-255.

Smajstrla, A.G., (1985). A field lysimeter system for crop water use and water stress studies in humid regions, Fundamental and Applied Biological Sciences, 53-59.

Sockett, R.E., 2009. Predatory lifestyle of Bdellovibrio bacteriovorus. Annual review of microbiology, 63, pp.523-539.

Song, J., Choo, Y.J. and Cho, J.C., 2008. Perlucidibaca piscinae gen. nov., sp. nov., a freshwater bacterium belonging to the family Moraxellaceae. International journal of systematic and evolutionary microbiology, 58(1), pp.97-102.
Šotić, A., Rajić, R., (2015). The review of the definition of risk. Online Journal of Applied Knowledge Management, 3, 17-26.

Stackebrandt, E., Schumann, P. and Prauser, H.E.L.M.U.T., 2006. The family cellulomonadaceae. Prokaryotes, 3, pp.983-1001.

Steinberg, J.P. and Burd, E.M., 2010. Other gram-negative and gram-variable bacilli. Mandell, Douglas, and Bennett's principles and practice of infectious diseases, pp.3015-3033.

Stoeva, M.K., Garcia-So, J., Justice, N., Myers, J., Tyagi, S., Nemchek, M., McMurdie, P.J., Kolterman, O. and Eid, J., 2021. Butyrate-producing human gut symbiont, Clostridium butyricum, and its role in health and disease. Gut microbes, 13(1), p.1907272.

Stokell, J.R., Hamp, T.J. and Steck, T.R., 2016. Examining changes in bacterial abundance in complex communities using next-generation sequencing is enhanced with quantitative PCR. Antonie van Leeuwenhoek, 109, pp.1161-1166.

Strauch, G., Bittkau, A., Schlosser, D., Petzoldt, H. and Sbjeschnid, A., 2005. Assessing the pollution of water protection areas and river banks along the Elbe and Mulde Rivers after the flood event in August 2002. Acta hydrochimica et hydrobiologica, 33(5), pp.418-429.

Streeter, K. and Katouli, M., 2016. Pseudomonas aeruginosa: a review of their pathogenesis and prevalence in clinical settings and the environment.
T. Fenchel, Microbial behavior in a heterogeneous world, Science, 296 (2002), pp. 1068-1071.

Tan, B., Ng, C.M., Nshimyimana, J.P., Loh, L.L., Gin, K.Y. and Thompson, J.R., 2015. Next-generation sequencing (NGS) for assessment of microbial water quality: current progress, challenges, and future opportunities. Frontiers in microbiology, 6, p.158998.

Tandlich, R., Ncube, M., SMM, K. and Zuma, B.M., 2016. A case study on the health risks related to flood disasters in South Africa. Journal of Disaster Research, 11(4), pp.732-741.

Taylor, J., Davies, M., Canales, M. and man Lai, K., 2013. The persistence of flood-borne pathogens on building surfaces under drying conditions. International journal of hygiene and environmental health, 216(1), pp.91-99.

Taylor, J., Davies, M., Canales, M. and man Lai, K., 2013. The persistence of flood-borne pathogens on building surfaces under drying conditions. International journal of hygiene and environmental health, 216(1), pp.91-99.

Taylor, J., man Lai, K., Davies, M., Clifton, D., Ridley, I. and Biddulph, P., 2011. Flood management: prediction of microbial contamination in large-scale floods in urban environments. Environment international, 37(5), pp.1019-1029.

Ten Veldhuis, J.A.E., Clemens, F.H.L.R., Sterk, G. and Berends, B.R., 2010. Microbial risks associated with exposure to pathogens in contaminated urban flood water. Water research, 44(9), pp.2910-2918.

Tett, A. et al. The Prevotella copri complex comprises four distinct clades underrepresented in westernized populations. Cell Host Microbe 26, 666–679 (2019).
The Rivers Trust (2021), Is my river fit to play in? Map Tool, Available online: https://experience.arcgis.com/experience/e834e261b53740eba2fe6736e37bbc7b/page/home/ [Accessed: 20/05/2021].
Thermo Fisher Scientific (2021), Qubit™ 4 Fluorometer, 3-71, Available online: https://www.fishersci.co.uk/shop/products/qubit-4-fluorometer-wifi/16223001#?keyword=qubit [Accessed: 16/06/2022].

Thursby, E. and Juge, N., 2017. Introduction to the human gut microbiota. Biochemical journal, 474(11), pp.1823-1836.

Timms, S., Slade, J.S, Fricker, C.R., (1995). Removal of Cryptosporidium by slow sand filtration. Water science and technology, 31,81-84.

Tong, G. Long, X. Jiang, H.N. Kim. Contribution of extracellular polymeric substances on representative gram negative and gram-positive bacterial deposition in porous media. Environ. Sci. Technol., 44 (7) (2010), pp. 2393-2399

Towner, K.J., 2009. Acinetobacter: an old friend, but a new enemy. Journal of Hospital Infection, 73(4), pp.355-363.
Tufenkji, N.; Dixon, D. R.; Considine, R.; Drummond, C. J. (2006), Multi-scale Cryptosporidium/sand interactions in water treatment. Water Research, 40, 3315−3331.

UKHSA, 2024, P . aeruginosa bacteraemia: monthly data by location of onset and NHS organisation, from November 2022 to November 2023, https://www.gov.uk/government/statistics/p-aeruginosa-bacteraemia-monthly-data-by-location-of-onset

Van de Peer Y, Chapelle S, De Wachter R. A quantitative map of nucleotide substitution rates in bacterial rRNA. Nucleic Acids Res. 1996;24:3381–3391.
Van Elsas, J.D., Semenov, A.V., Costa, R. and Trevors, J.T., 2011. Survival of Escherichia coli in the environment: fundamental and public health aspects. The ISME journal, 5(2), pp.173-183.

Van Gestel M, Merckx VR, and Vlassak K. 1993. Microbial biomass responses to soil drying and rewetting: the fate of fast- and slow-growing microorganisms in soils from different climates. Soil Biol. Biochem. 25:109–123.

Vandekerckhove, T.T., Willems, A., Gillis, M. and Coomans, A., 2000. Occurrence of novel verrucomicrobial species, endosymbiotic and associated with parthenogenesis in Xiphinema americanum-group species (Nematoda, Longidoridae). International Journal of Systematic and Evolutionary Microbiology, 50(6), pp.2197-2205.

Vandenberg, O., Dediste, A., Houf, K., Ibekwem, S., Souayah, H., Cadranel, S., Douat, N., Zissis, G., Butzler, J.P. and Vandamme, P., 2004. Arcobacter species in humans. Emerging infectious diseases, 10(10), p.1863.
Vaughn, J.M., Landry, E.F., (1983). Viruses in soils and groundwaters. Viral pollution of the environment, 163-210.
Vega, E., Lesikar, B. and Pillai, S.D., 2003. Transport and survival of bacterial and viral tracers through submerged-flow constructed wetland and sand-filter system. Bioresource technology, 89(1), pp.49-56
Veronesi, M., Chawla, F., Maurer, M., Lienert, J., 2014. Climate change and the willingness to pay to reduce ecological and health risks from wastewater flooding in urban centers and the environment. Ecological Economics, 98, pp.1-10.
Vítězová, M.; Kohoutová, A.; Vítěz, T.; Hanišáková, N.; Kushkevych, I. Methanogenic Microorganisms in Industrial Wastewater Anaerobic Treatment. Processes 2020, 8, 1546. https://doi.org/10.3390/pr8121546

Wang, C., Li, S., Zhang, Z., Yu, Z., Yu, L., Tian, F., Chen, W. and Zhai, Q., 2021. Phocaeicola faecalis sp. nov., a strictly anaerobic bacterial strain adapted to the human gut ecosystem. Antonie Van Leeuwenhoek, 114(8), pp.1225-1235.

Wang, D.S., Gerba, C.P., Lance, J.C., (1981). Effect of soil permeability on virus removal through soil columns. Applied and environmental microbiology, 42, 83-88.
Wang, J., (2014). Combined sewer overflows (CSOs) impact on water quality and environmental ecosystem in the Harlem River. Journal of Environmental Protection, 5(13), p.1373.

Waśkiewicz, A. and Irzykowska, L., 2014. Flavobacterium spp.–characteristics, occurrence, and toxicity.
Water Framework Directive., 2000. EU Water framework directive. EC Directive, 60.

Waters, D., Watt, W.E., Marsalek, J., Anderson, B.C., (2003), Adaptation of a storm drainage system to accommodate increased rainfall resulting from climate change, Journal of Environmental Planning and Management, 46, 755-770.
Webber, J.L., Gibson, M.J., Chen, A.S., Savic, D., Fu, G., Butler, D., (2018). Rapid assessment of surface-water flood-management options in urban catchments. Urban Water Journal, 15(3), pp.210-217.
Webre, D.J., Wolanin, P.M.,Stock, J.B., (2003). Bacterial chemotaxis. Current Biology, 13, 47-49.

Wei, Y., Wang, B., Zhang, L., Zhang, J. and Chen, S., 2018. Pedobacter yulinensis sp. nov., isolated from sandy soil, and emended description of the genus Pedobacter. International Journal of Systematic and Evolutionary Microbiology, 68(8), pp.2523-2529.
Weiss, T.H., Mills, A.L., Hornberger, G.M. ,Herman, J.S., (1995). Effect of bacterial cell shape on transport of bacteria in porous media. Environmental science & technology, 29, 1737-1740

Werner, D.; Acharya, K.; Blackburn, A.; Zan, R.; Plaimart, J.; Allen, B.; Mgana, S.M.; Sabai, S.M.; Halla, F.F.; Massawa, S.M.; et al. MinION Nanopore Sequencing Accelerates Progress towards Ubiquitous Genetics in Water Research. Water 2022, 14, 2491. https://doi.org/10.3390/w14162491

West, P.A., (1991). Human pathogenic viruses and parasites: emerging pathogens in the water cycle. Journal of applied bacteriology, 70,107-114.

Whittaker RH (1960). "Vegetation of the Siskiyou Mountains, Oregon and California". Ecological Monographs. 30 (3): 279–338.

Wilson, J.S., Baldwin, D.S., Rees, G.N. and Wilson, B.P., 2011. The effects of short-term inundation on carbon dynamics, microbial community structure and microbial activity in floodplain soil. River Research and Applications, 27(2), pp.213-225.

Wingender, J. and Flemming, H.C., 2011. Biofilms in drinking water and their role as reservoir for pathogens. International journal of hygiene and environmental health, 214(6), pp.417-423.

World Meteorological Organisation, 2020, Guide to Instruments and Methods of Observation Number 8, Volume I –Measurement of Meteorological Variables, 2020 edition , Chapter 6; Measurement of precipitation, 223-237. Available online: https://library.wmo.int/doc_num.php?explnum_id=11337 [Accessed: 18/10/2022].
Wu, J., Xu, Z., Li, H., Li, P., Wang, M., Xiong, L., Zhang, J., (2019). Long-term effect of water diversion and CSOs on the remediation of heavy metals and microbial community in river sediments. Water Science and Technology, 79(12), pp.2395-2406.

Wu, W., Jin, Y., Bai, F. and Jin, S., 2015. Pseudomonas aeruginosa. In Molecular medical microbiology (pp. 753-767). Academic Press.

Xia Y, Li AD, Deng Y, Jiang XT, Li LG, Zhang T. MinION Nanopore Sequencing Enables Correlation between Resistome Phenotype and Genotype of Coliform Bacteria in Municipal Sewage. Front Microbiol. 2017 Oct 31;8:2105. doi: 10.3389/fmicb.2017.02105. PMID: 29163399; PMCID: PMC5671560.

Xiao, S., Hu, S., Zhang, Y., Zhao, X. and Pan, W., 2018. Influence of sewage treatment plant effluent discharge into multipurpose river on its water quality: A quantitative health risk assessment of Cryptosporidium and Giardia. Environmental Pollution, 233, pp.797-805.

Xie, C.H. and Yokota, A., 2006. Sphingomonas azotifigens sp. nov., a nitrogen-fixing bacterium isolated from the roots of Oryza sativa. International journal of systematic and evolutionary microbiology, 56(4), pp.889-893.

Xu, Z., Xiong, L., Li, H., Liao, Z., Yin, H., Wu, J., Xu, J. and Chen, H., 2017. Influences of rainfall variables and antecedent discharge on urban effluent concentrations and loads in wet weather. Water Science and Technology, 75(7), pp.1584-1598.

Yadav, N. and Sharma, S., 2020. Pollution shapes the bacterial community of a river: a case study. International journal of environmental science and technology, 17, pp.2003-2016.

Yadav, R.K., Goyal, B., Sharma, R.K., Dubey, S.K. and Minhas, P.S., 2002. Post-irrigation impact of domestic sewage effluent on composition of soils, crops and ground water—a case study. Environment international, 28(6), pp.481-486.

Yang, J-L. Zhang, G-L. (2015), Formation, characteristics and eco-environmental implications of urban soils – A review, Soil Science and Plant Nutrition, 61:1, 30-46.

Yang, P. and van Elsas, J.D., 2018. Mechanisms and ecological implications of the movement of bacteria in soil. Applied Soil Ecology, 129, pp.112-120.

Yard, E.E., Murphy, M.W., Schneeberger, C., Narayanan, J., Hoo, E., Freiman, A., Lewis, L.S., Hill, V.R., (2014). Microbial and chemical contamination during and after flooding in the Ohio River—Kentucky, 2011. Journal of Environmental Science and Health, Part A, 49, 1236-1243.

Yeoh, Y.K., Sun, Y., Ip, L.Y.T., Wang, L., Chan, F.K., Miao, Y. and Ng, S.C., 2022. Prevotella species in the human gut is primarily comprised of Prevotella copri, Prevotella stercorea and related lineages. Scientific Reports, 12(1), p.9055.

Yildiz, B.S., 2012. Water and wastewater treatment: Biological processes. In Metropolitan Sustainability (pp. 406-428). Woodhead Publishing.

Yoon, J.H. and Park, Y.H., 2006. The genus Nocardioides. The prokaryotes, 3, pp.1099-1113.

Young, I.M, Crawford,J.W. Interactions and self-organization in the soil–microbe complex, Science, 304 (2004), pp. 1634-1637.

Yu, Z., Wang, G. and Marschner, P., 2014. Drying and rewetting–effect of frequency of cycles and length of moist period on soil respiration and microbial biomass. European Journal of Soil Biology, 62, pp.132-137.

Yunes, R.A., Poluektova, E.U., Vasileva, E.V., Odorskaya, M.V., Marsova, M.V., Kovalev, G.I. and Danilenko, V.N., 2020. A multi-strain potential probiotic formulation of GABA-producing Lactobacillus plantarum 90sk and Bifidobacterium adolescentis 150 with antidepressant effects. Probiotics and antimicrobial proteins, 12, pp.973-979.

Z. Nevo, R. Mitchell, Factors affecting biological clogging of sand associated with ground water recharge, Water Res., 1 (1967), pp. 231-236

Ze, X., Duncan, S.H., Louis, P. and Flint, H.J., 2012. Ruminococcus bromii is a keystone species for the degradation of resistant starch in the human colon. The ISME journal, 6(8), pp.1535-1543.

Zerda, K.S., Gerba, C.P., Hou, K.C., Goyal, S.M., (1985). Adsorption of viruses to charge-modified silica. Applied and Environmental Microbiology, 49, 91-95.

Zhang DC, Wang HX, Liu HC, Dong XZ, Zhou PJ. 2006. Flavobacterium glaciei sp. nov., a novel psychrophilic bacterium isolated from the China No.1 glacier. Int J Syst Evol Microbiol 56:2921–2925. doi: 10.1099/ijs.0.64564-0.

Zhou, J., Holmes, D.E., Tang, H.Y. and Lovley, D.R., 2021. Correlation of key physiological properties of Methanosarcina isolates with environment of origin. Applied and Environmental Microbiology, 87(13), pp.e00731-21.

Haas, C.N., Rose, J.B. and Gerba, C.P., 2014. Quantitative microbial risk assessment. John Wiley & Sons.

World Health Organization, 2016. Quantitative microbial risk assessment: application for water safety management.

Horswell, J., Hewitt, J., Prosser, J., Van Schaik, A., Croucher, D., Macdonald, C., Burford, P., Susarla, P., Bickers, P. and Speir, T., 2010. Mobility and survival of Salmonella Typhimurium and human adenovirus from spiked sewage sludge applied to soil columns. Journal of applied microbiology, 108(1), pp.104-114.

# Appendices

**Appendix 1**

**Table S1**. Table of sizes of sieves used to size soil samples, according to British Standard 1377 Part 2 (British Standards, 1990).

| Sieve aperture size |
|---|
| 5mm |
| 3.35mm |
| 2mm |
| 1.18mm |
| 600µm |
| 425 µm |
| 300 µm |
| 212 µm |
| 160 µm |
| 63 µm |
| Base/Passing 63 µm |

**Appendix 2**

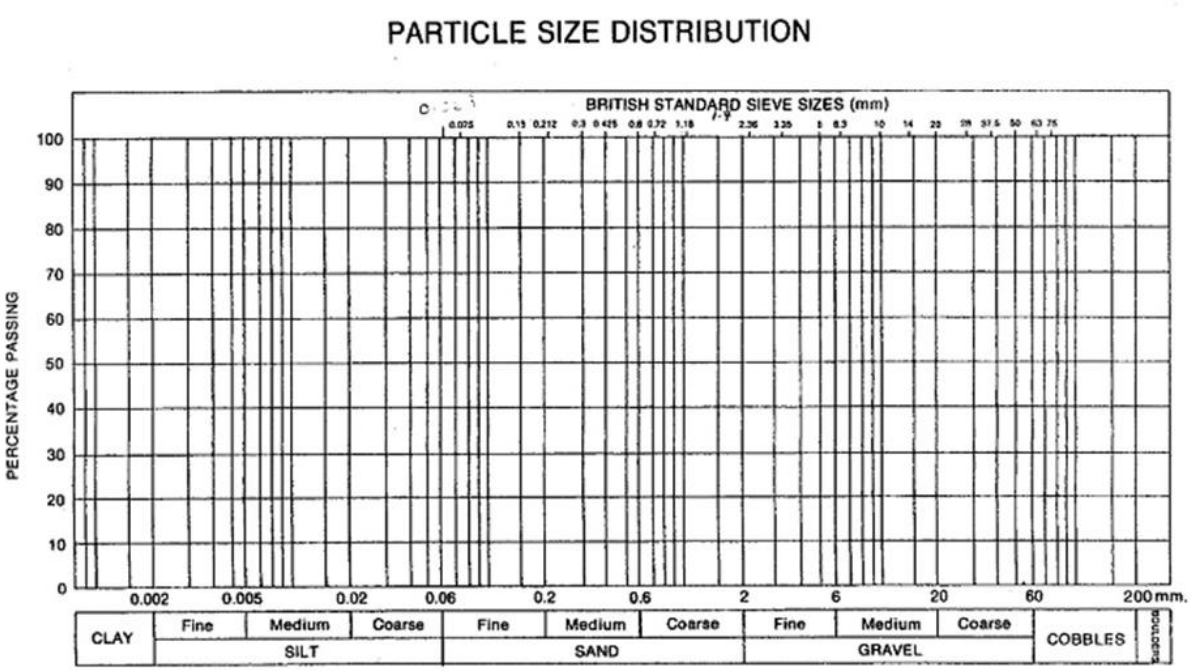


**Figure S1**. Particle size distribution chart, used to plot results from dry sieving. X axis is particle size in mm, secondary axis is the sieve aperture size (British Standards Institute, 1990).

## Appendix 3

**Table S2**. Soil type table (Gravels and Sands) used in conjunction with Appendix 2 to determine soil type according to British Standard 1377 Part 2 (British Standards, 1990).

| Soil Groups | | | | Description and Identification | | Sub-Groups | | Liquid Limit | Fines (% <0.06 mm) |
|---|---|---|---|---|---|---|---|---|---|
| COARSE SOILS (<35% fines) | GRAVELS (>50% of coarse material is of gravel size – >2mm) | Slightly silty or clayey GRAVEL | G | Well graded GRAVEL | GW | – | – | | 0–5 |
| | | | | Poorly graded GRAVEL | GP | Uniformly graded* | GPu | | |
| | | | | | | Gap graded | GPg | | |
| | | Silty or clayey GRAVEL | G-F | Silty GRAVEL | G-M | Well graded | GWM | | 5–15 |
| | | | | | | Poorly graded | GPM | | |
| | | | | Clayey GRAVEL | G-C | Well graded | GWC | | |
| | | | | | | Poorly graded | GPC | | |
| | | Very silty or clayey GRAVEL | GF | Very silty GRAVEL | GM | Subdivisions as for GC | GML etc. | As GC | 15–35 |
| | | | | Very clayey GRAVEL | GC | Low plasticity fines | GCL | <35 | |
| | | | | | | Intermediate plasticity fines | GCI | 35–50 | |
| | | | | | | High plasticity fines | GCH | 50–70 | |
| | | | | | | Very high plasticity fines | GCV | 70–90 | |
| | | | | | | Extremely high plasticity fines | GCE | >90 | |
| | SANDS (>50% of coarse material is of sand size – 0.06mm to 2.00mm) | Slightly silty or clayey SAND | S | Well graded SAND | SW | – | – | | 0–5 |
| | | | | Poorly graded SAND | SP | Uniformly graded* | SPu | | |
| | | | | | | Gap graded | SPg | | |
| | | Silty or clayey SAND | S-F | Silty SAND | S-M | Well graded | SWM | | 5–15 |
| | | | | | | Poorly graded | SPM | | |
| | | | | Clayey SAND | S-C | Well graded | SWC | | |
| | | | | | | Poorly graded | SPC | | |
| | | Very silty or clayey SAND | SF | Very silty SAND | SM | Subdivisions as for SC | SML etc. | As SC | 15–35 |
| | | | | Very clayey SAND | SC | Low plasticity fines | SCL | <35 | |
| | | | | | | Intermediate plasticity fines | SCI | 35–50 | |
| | | | | | | High plasticity fines | SCH | 50–70 | |
| | | | | | | Very high plasticity fines | SCV | 70–90 | |
| | | | | | | Extremely high plasticity fines | SCE | >90 | |

## Appendix 4

Scientific Publications

### *Appendix 4.1: Conference Proceedings Papers*

S Scutt, J Shucksmith, HS Jensen, J Diaz-Nieto, ID Soler, (2022) Urban floods: the microbial risks to human health, “Youth” in the forefront: before and after World Water Forum. Online Youth Water Congress: “Emerging water challenges since COVID-19”, BOOK OF ABSTRACTS, 12-13.

Scutt. S, J. Shucksmith, I.Douterelo, (2022) Investigating the microbiological risks associated with urban flooding in the UK, In: Access Microbiology. Microbiology Society Annual Conference 2021, 26-31.

### *Appendix 4.2: Conferences*

*Poster Presentation*: ‘Evaluating the Public Health Risks of Urban Flooding Events' April 2021 Microbiology Society Annual Conference, Online.

*Oral Presentation* : ‘Urban Floods: The Microbial Risk to Human Health' April 2022 The United Nations Educational, Scientific and Cultural Organization (UNESCO), Youth in the Forefront, Online Water Congress, ‘Emerging water challenges since COVID-19', Online.

*Poster Presentation*: ‘Evaluating the microbial risks of urban flooding events' April 2022 Microbiology Society Annual Conference, Belfast, United Kingdom.

*Oral Presentation*: ‘Analysis of soil and water from urban flood sites to identify seasonal changes in health risk from microbial communities'. August 2022. Conference on Sewer Networks and Processes (SPN 10), Graz, Austria.

*Oral Presentation*: ‘Floods in the city: The microbial risks' May 2023 Royal Society of Public Health- Thinking Outside the Box, planning plumbing systems for the future. Online.

*Poster Presentation*: ‘Using molecular methods to determine the risk posed by microorganisms present in urban flood water in the UK.' June 2023 Testing the Waters 6- Wastewater based epidemiology: current development and future perspectives, Oxford, United Kingdom.

*Appendix 4.3. Invited Talks*

Sophie Scutt, ‘Evaluating the Public health risks of urban flooding events', Urban Systems Studio, Rhode Island School of Design, Rhode Island, United States of America, online. March 2022.