Characterising and modelling the sources and drivers of antimicrobial resistance in rural river catchments in Northern England

Katie Robins

Submitted for a Doctor of Philosophy

School of Engineering

Newcastle University

January 2024

Abstract

Reducing Antimicrobial Resistance (AMR) in the environment requires adopting a One Health approach. However, most current river catchment monitoring is limited to focused rather than holistic research work, often in more contaminated river catchments. The sources and drivers of AMR in less-impacted rural catchments are not as well defined. Further, studying such catchments can be problematic because wide scale environmental monitoring is expensive, time consuming and currently not standardised. To address such problems, a comprehensive spatial assessment was conducted on AMR in two rural river catchments in the North of England. The work included the Coquet River in Northumberland and the Eden in Cumbria, which both have rural land-use, but different hydrometeorological characteristics. A subsequent focused study was performed on the Coquet catchment that used Fast-Expectation Maximization for Microbial Source Tracking (FEAST), a relatively new approach, for whole community microbial source tracking. The catchment comparison revealed that elevated river flows, rainfall and runoff in the Eden catchment led to a higher abundance and higher diversity of resistance genes (Kruskal Wallis, $p<0.05$) and clinically relevant ARGs such as bla_{KPC}. The FEAST analysis also revealed the influx of ARGs from wastewater effluent were not sustained down the course of the Coquet river. Microbial communities and resistance genes in these catchments were driven primarily by environmental factors, such as catchment hydrology and nutrient limitation (described through N:P ratios), rather than point sources. These studies highlighted the need for increased monitoring to support these findings, and to inform routine monitoring. Models can be important public health decision tools to support such surveillance. The Soil and Water Assessment Tool (SWAT) was used to simulate *E. coli* and ESBL *E. coli* for the Coquet catchment, demonstrating the potential of SWAT to predict *E. coli* and ESBL *E. coli* to support on-site monitoring of AMR.

Title photo by Katie Robins

Acknowledgements

This work was funded by the EPSRC Doctoral Training Partnership (EJU/170341185), the Environment Agency and the Department for Environment Food and Rural Affairs (DEFRA).

Firstly, a huge thank you goes to my supervisory team. To David Graham, for your support, enthusiasm and leadership throughout this project. To Anke Neumann for your guidance and excellent fine-tuning of my academic writing. And to Greg O'Donnell for your help with developing the model and your mentorship in this new and daunting task in my PhD work.

Thank you to my industrial sponsors for valuable insight into this work, including Alwyn Hart and Wiebke Schmidt from the Environment Agency, and Ruth Waite and Amelia Breeze from DEFRA. To Michael Donkin from the Environment Agency and Sean Burke from the British Geological Survey, thank you for providing on-site tours of the Coquet and Eden, and lending insight that helped inform the sampling for this project. Also thank you to Andrew Moore from Northumbrian Water for arranging access to relevant wastewater treatment sites.

I would also like to thank colleagues in Newcastle University for assisting with the field sampling, including Maggie White, Sreelakshmi Babu, Panagiota Adamou, Senan Almashta, Amelie Ott and Jakub Konieczynski. In addition, thank you to Marcos Quintela-Baluja and Kelly Jobling for their insight in the project and their assistance with laboratory work. Also thank you to Amelie Ott for bioinformatics guidance and assistance in planning the field sampling. Thank you also to the team at Resistomap, for performing the HT-qPCR.

A special thanks goes to my friends and colleagues at Newcastle University for coffee breaks, snack sharing and putting things into perspective. Thank you also to my friends and family for the support over the years. In particular, thank you to my parents for teaching me early on to remain curious and to persevere when the going gets tough. And last but definitely not least, to my fiancée Jakub for the much-needed distractions and trips away from work, cooking all the meals during the write up process and your support and kindness throughout.

Declaration

I thereby certify that the work presented in this thesis is my original research work with the exception of below analysis:

- High-throughput qPCR to measure antibiotic resistant genes and mobile genetic elements in Chapter 3 and 4 and selection of 96 genes, performed by Resistomap Oy, Helsinki Finland.
- Illumina DNA sequencing in Chapter 3 and 4, performed by Andrew Nelson at Northumbria University.

Due reference is given to literature and any other research collaborations where appropriate. The research presented in this thesis has not been submitted for any other degree or professional qualification.

Statement of publications

Chapter 3 is based on a publication with some edits and further details added.

Robins, K., O'Donnell, G., Neumann, A., Schmidt, W., Hart, A. and Graham, D.W., 2024. Antimicrobial resistance in rural rivers: Comparative study of the Coquet (Northumberland) and Eden (Cumbria) River catchments. *Science of the Total Environment*, *928*, p.172348. https://doi.org/10.1016/j.scitotenv.2024.172348

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

The effect of the recent COVID-19 pandemic on human health and global economies provided a stark insight into the potential impacts of antimicrobial resistance (AMR), should the problem continue to proliferate. Antibiotics are essential for modern healthcare, treating enumerable infections such as pneumonia, bacteraemia and surgical site infections. Due to increased antibiotic use, secondary infections and the development of drug-resistant nosocomial infections, the pandemic has only enhanced the existing problem of AMR (Ukuhor 2021). Additional external factors are also reported to influence the development of resistance, such as population density and local temperature, meaning the burden of AMR could be further enhanced with global population increase and climate change (MacFadden et al. 2018).

The introduction of numerous types of antibiotics for human and animal healthcare during the 20th century, resulted in a significant advancement of healthcare and agricultural practices. However, with limited novel antibiotics being discovered (Fair and Tor 2014), and novel forms of resistance emerging, we are running out of options to treat emerging forms of resistance and resistant infections. In addition, some of the resistance mechanisms are so complex, that our current antimicrobials cannot overcome them (Ventola 2015). Whilst alternative treatments to bacterial infections exist, including phage therapy and nanoparticles, these are still in their infancy (Alaoui Mdarhri et al. 2022).

Globally, significant economic and health costs related to AMR have been projected and estimates for 2050 predict that AMR will incur costs of around US$100 trillion as well as ~10 million deaths, a figure which exceeds that of cancer (O'Neill 2016). Therefore, AMR has become a global priority, particularly in the response to the World Health Organisation (WHO) action plan on AMR (WHO 2016). Part of this plan includes an increase in surveillance and research, and has led to initiatives such as the WHO Global Antimicrobial Resistance and Use Surveillance System (GLASS). Launched in 2015 GLASS aims to standardise the approach for monitoring and surveillance of AMR across countries (WHO 2020, 2021a).

AMR is naturally occurring, with antibiotic resistance genes (ARGs) found in ~30,000 yr-old glacial sediments (Dcosta et al. 2011). However, the naturally present resistance in the environment has been enhanced and expanded by human activity.

Antibiotics are extensively used in animal, environmental as well as human settings. In agriculture, antibiotics are supplied to animals for therapeutic purposes, but in some countries and settings, also for growth promotion or to compensate for poor hygiene practices (Wegener 2003). Resistance can enter the environment through faecal contamination, or through spreading of farm slurry on to agricultural land. In the UK, around 50% of antibiotics are used for agricultural processes, and this is predicted to rise to 67% in 2030 (Manyi-Loh et al. 2018). Whilst the UK government are introducing measures to limit agricultural antibiotic use, resistance to commonly used antibiotics remains high in livestock (Hennessey et al. 2020). Due to the potential intersection of human, animal and environmental antimicrobial resistance, there is a need to understand the potential transmission pathways. Therefore, AMR should be studied within the context of 'One Health', which concerns the interconnection of human, animal and environmental health.

Assessing the anthropogenic impact on AMR within an environment can be determined through analysis of a river system at catchment scale, as rivers act as a conduit for environmental contamination. Most studies to date have focused on evidently "contaminated" catchments (particularly in the United Kingdom) and studies investigating rural AMR are lacking. In addition, studies investigating environmental AMR primarily focus on isolated and single river catchments. Due to the lack of standardisation of techniques for investigating AMR and microbial communities, comparison across different studies is often impossible. Therefore, external drivers for AMR, such as hydrology and climate cannot easily be identified or distinguished. Factors like precipitation and runoff are likely to be important for the dissemination of AMR in a river catchment (Almakki et al. 2019), particularly in environments where there is agricultural contamination.

In rural river catchments, there are many potential sources which can enhance the natural resistome, including small wastewater treatment works, septic tanks, agricultural contamination, and-or co-selective agents, such as metals and biocides (Stepanauskas et al. 2006; Seiler and Berendonk 2012; Robins et al. 2022). Recently, Fast Expectation Maximization Microbial Source Tracking- or 'FEAST', was developed as an approach for source tracking using whole microbial communities and resistance genes from sources and sinks (Shenhav et al. 2019). Whilst methods such as these can be used to estimate the sources contributing to the microbial

communities and resistance genes in a sample, the factors which control the survival need to be understood. Nutrient availability could be an important factor in influencing this (Hibbing et al. 2010), but has yet to be investigated in the context of AMR within rural river systems.

Environmental surveillance for the purposes of AMR risk assessment is still not common practice, and prioritising the locations and tools to conduct this surveillance is still ongoing (Bengtsson-Palme et al. 2023). For a more targeted surveillance approach, there is a need to determine where areas of interest may be (e.g., AMR hotspots), where in-depth analysis may be conducted (e.g. genomic and metagenomic analysis). For this, predictive models, such as AMR models in river systems may help to identify AMR hotspots and have received increased research interest over the last decade (Hellweger et al. 2011; Hellweger 2013; Gothwal and Thatikonda 2018, 2020; Van Heijnsbergen et al. 2022; Jampani et al. 2023; Niebaum et al. 2023) and those which enable the use of Geographical Information Systems (GIS) can incorporate spatial information into the AMR model. This includes the Soil and Water Assessment tool (SWAT), which is open source, and has been applied to many disciplines, including predicting pathogen concentrations in a river catchment (Coffey et al. 2010a, 2013; Frey et al. 2013; Niazi et al. 2015), but has yet to be used for AMR prediction.

1.1 Thesis aims and objectives

This PhD thesis aims to understand the importance of various drivers and sources of rural AMR in Northern England and to suggest the best tools and approaches for continued surveillance. This aim will be met through addressing the following objectives:

1. Perform a thorough spatial assessment of AMR and microbial communities in the Coquet (Northumberland) and Eden (Cumbria) Rivers.
2. Compare the microbial and hydrological signatures of the Coquet and Eden River catchments.
3. Focusing on the Coquet catchment, use FEAST as an approach for community source tracking and determine the most important sources in the catchment.
4. Identify the sources and drivers for AMR in rural river catchments in the UK

5. Demonstrate the ability of the open-source SWAT modelling tool to predict AMR hotspots in a river catchment.
6. Suggest tools and approaches to continue monitoring environmental AMR in rural river catchments

1.2 Thesis outline

This thesis consists of six chapters:

Chapter 1 introduces the thesis, the aims and tasks, and the structure.

Chapter 2 reviews the literature on the sources and drivers of environmental AMR and modelling AMR in the environment.

Chapter 3 compares the Coquet and Eden River catchments, which both have similar land-uses but different hydrometeorological characteristics. Through comparison of catchment hydrology and microbial signatures, drivers for elevated AMR could be identified and interpreted.

Chapter 4 focuses on the Coquet River catchment and assesses the contribution of 'source' samples to the downstream river composition. Microbial communities and resistance genes from point source samples and river samples were used in combination with the FEAST approach for microbial source tracking. This study also investigates the impact of nutrient availability on the persistence of resistance genes.

Chapter 5: develops a SWAT hydrological model for the Coquet catchment and integrates modelling of *E. coli* and extended spectrum beta-lactamase (ESBL) *E. coli* as indicators for AMR.

Chapter 6 summarises the findings from the studies, the application of the research findings to evaluate and map AMR risk in river catchments and recommends directions for future research.

Chapter 2 Literature review

2.1 Origin and dissemination of antimicrobial resistance

The discovery of antibiotics in 1928 has transformed our modern healthcare system, making previously life-threatening infections easily treatable, as well as allowing for major surgeries and organ transplantation. Alexander Fleming's discovery of penicillin, triggered the 'golden age of antibiotic discovery' in the 1930s and 1960s (Dantas and Sommer 2011), with the development of broad-spectrum streptomycin, tetracyclines, chloramphenicol, erythromycin and other life-saving antibiotics (Keen and Montforts 2011). However, as stated by Alexander Fleming in his Nobel Prize acceptance speech in 1945: *"It is not difficult to make microbes resistant to penicillin in the laboratory by exposing them to concentrations not sufficient to kill them, and the same thing has occasionally happened in the body"* (Fleming 1945). This effect was observed with the discovery of novel antimicrobials, quickly followed by the development of resistance (Dantas and Sommer 2011) (Figure 2-1).

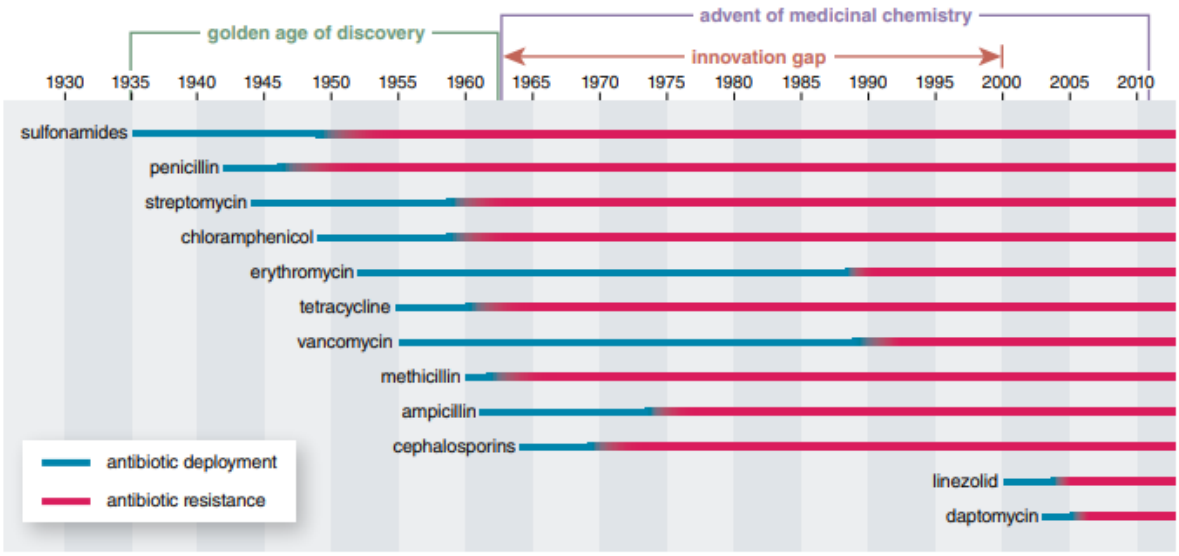


Figure 2-1 The clinical deployment of antibiotics (blue bars) was quickly followed by resistant bacteria (red bars) from (Dantas and Sommer, 2014)

In the 1970s and 80s, multidrug resistant (MDR) infectious organisms were emerging, and the problem of resistance had extended beyond clinical settings (Keen and Montforts 2011). The widespread use of antimicrobials in agriculture changed the course of livestock production, towards intensive industrialised production (Caneschi et al. 2023). Antimicrobials in agriculture were used for therapeutic use, but also intensively for growth promotion purposes (Davies and Davies 2010). The long-term

development of resistance in agriculture is evident from archival soils, which showed an increase in resistance from 1940 to 2010 (Knapp et al. 2010).

Whilst resistance to antibiotics aligns with widespread antibiotic consumption (Dantas and Sommer 2014), the evolution and acquisition of antibiotic resistance genes (ARGs) is a natural phenomenon and occurs in the absence of antibiotic use. It has been suggested that the role of natural (or 'intrinsic') resistance, is to resist antibiotic compounds produced by a microorganism to inhibit the growth for other members of the microbial community and reduce competition for resources (Kümmerer 2009). ARGs have been recovered from DNA extracted from ancient Beringian permafrost sediments (Dcosta et al. 2011) and multiresistant *Springomonas* sp. was isolated from an Antarctic hypolith (Gunnigle et al. 2015). Chen et al., (2019a) compared the metagenomics profiling of the Chaobai river in China and two pristine environments: Antarctic soils and deep-sea sediments. In pristine environments, it was found that ARGs encoding resistance to the antibiotic bacitracin were the most dominant in abundance. Additional studies have also shown that the gene *bacA* is intrinsic (Hu et al. 2013). Overall, the concentration of intrinsic resistance has been estimated to be very low (Kümmerer 2009; Dcosta et al. 2011).

Due to the low concentration of intrinsic environmental AMR, most of the discussion related to antibiotic resistance in environmental settings focuses on acquired resistance, where stressors from human activity exacerbate and select resistance in strains, including pathogens, which were not originally resistant. For acquired resistance, there are four ways in which bacteria can resist antibiotics: (1) exporting substance from cells using the upregulation of efflux pumps (2) modifying the antibiotic through genetic mutation or hydrolysis, (3) expressing degradation enzymes which make the antibiotic ineffective, and (4) protecting the molecular target of the antibiotic (e.g. such as the ribosome) (Arzanlou et al. 2017; Abushaheen et al. 2020). Mechanisms which are characteristics of intrinsic resistance include modifying the permeability of the outer membrane, which can influence the ability of an antibiotic to enter the bacteria cells (Cox and Wright 2013). Efflux pumps can also be an intrinsic mechanism of resistance, as they have been shown to have evolved for purposes other than resisting antibiotics (Cox and Wright 2013). Resistance can be transferred between bacteria through horizontal (HGT) and vertical gene transfer (VGT) (Arnold et al. 2022). VGT occurs through progeny of original hosts (Lawrence 2005). However, compared to VGT, HGT has more diverse methods of acquiring

resistance, including (1) natural transformation, where a cell assimilates DNA from its natural environment (2) transduction, where a virus moves DNA from one bacterial cell to another, and (3) conjugation, where one bacterium transfers genetic material to another bacterium through physical connection (von Wintersdorff et al. 2016).

Acquired resistance can also be facilitated by mobile genetic elements (MGEs). MGEs, including (1) plasmids, small circular double stranded extrachromosomal DNA found within a bacteria and capable of self-replication (Bennett 2008), (2) transposons, genetic elements referred to as 'jumping genes' due to their ability to 'jump' to different locations in the genome (Pray 2008) and (3) bacteriophages, viruses that infect bacteria and facilitate the exchange of genetic material (Calero-Cáceres et al. 2019; Strange et al. 2021). Prior to the development of antibiotics, ARGs were located on the chromosome of non-pathogenic bacteria: however, since the antibiotic era, ARGs have been increasingly found on MGEs in pathogenic and faecal bacteria (Datta and Hughes 1983).

Integrons have been particularly associated with the dissemination of AMR, as they are genetic elements with the ability to integrate and express ARGs. Class 1 integrons (*intI1*) were proposed as a proxy for anthropogenic pollution (Gillings et al. 2015). Integrons facilitate the integration of ARGs, typically as part of gene cassettes, which are inserted into the integron structure by recombination between the *attC* of the cassette and *attI* site of an integron (Figure 2-2) (Gillings 2014). Through repeated recombination events, integrons can collect as well as disperse gene cassettes as free circular DNA elements (Hocquet et al. 2012; Quintela-Baluja et al. 2021).

2.2 Ecology of antimicrobial resistance in the environment

With the use of antibiotics in human, animal and environmental settings, AMR has been acknowledged as a One Health issue, which recognises the interconnection of the human, animal and environmental health sectors (Figure 2-3) (Robinson et al. 2016; Walsh 2018; Léger et al. 2021; Larsson et al. 2023). Through anthropogenic activity, bacteria from different origins move across sectors and exchange genetic information between compartments. The environment is a source for numerous non-

human and opportunistic pathogens, which have the potential to be transferred resistance (Bengtsson-Palme et al. 2018).

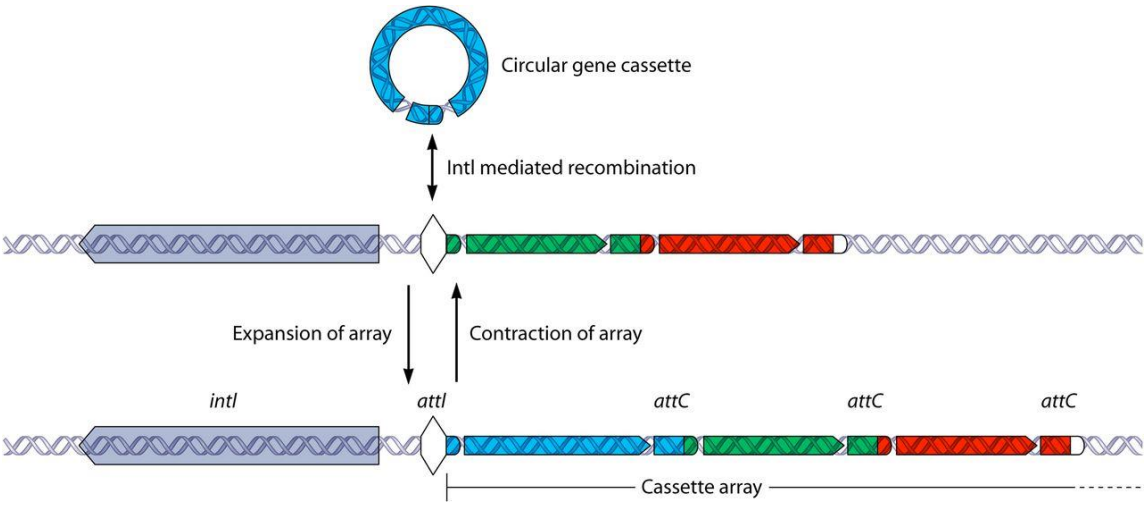


Figure 2-2 Integron gene acquisition process (Figure from Gillings, 2014).

Usually, the transfer of genetic material occurs when the bacteria are phylogenetically related and the species are in contact, as in, they are at least temporarily sharing the same habitat (Smillie et al. 2011; Wiedenbeck and Cohan 2011). Additionally, transfer of genetic material in the environment is usually induced by external stressors, such as a high concentration of antibiotics, metals or biocides (Baker-Austin et al. 2006; Wiedenbeck and Cohan 2011). High concentrations of metals and metalloids have been shown to increase ARG selection (Knapp et al. 2010, 2011; Robins et al. 2022), due to co-selection and co-resistance (Pal et al. 2015).

Selective pressure due to external factors in the environment has been found to be important for the persistence of AMR in environmental settings. ARG transfer has an associated fitness cost, therefore ARGs kept by bacteria in the absence of selection pressure are typically low fitness cost and high fitness cost ARGs are typically lost when there is no selection pressure (Bengtsson-Palme et al. 2018). Even if an ARG is lost in the absence of selective pressure, it can be quickly regained if the selection pressure returns (Levin et al. 1997). Successfully mobilised ARGs either begin with a low fitness cost or evolve to have a low fitness cost on an MGE (Gonzalez et al. 2013). The role of the environment as a source, recipient and sometimes amplifier of resistance is evident in the clinically relevant ARGs which have been shown to have environmental origin, including beta-lactam resistance genes *bla*$_{TEM}$ and *bla*$_{SHV}$ (Wright 2010).

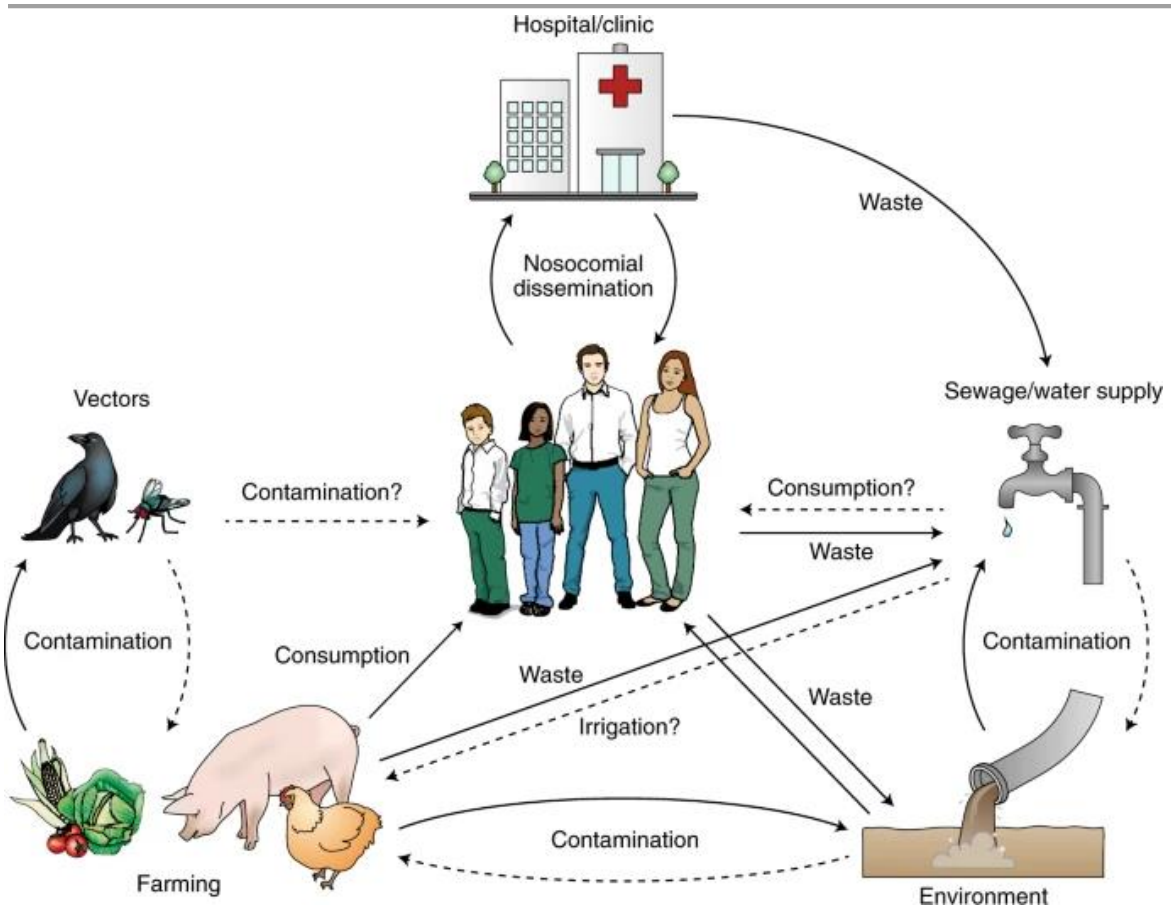


Figure 2-3 Potential transmission pathways between human, environmental and agricultural reservoirs of AMR, illustrating that AMR is a One Health issue (from Walsh 2018).

2.3 Monitoring microbial communities and AMR in the environment

Routine and targeted environmental monitoring is required to tackle AMR through a One Health approach (Bengtsson-Palme et al. 2023). Despite this, whilst AMR is monitored in clinical settings for humans and livestock, there is no standardised approach for AMR monitoring for the environment (Anjum et al. 2021). There have been numerous approaches to assess AMR in the environment, where the application depends on the surveillance area of interest.

One can gain insight into microbial communities from microbial culturing, where such methods can characterise viable and culturable target bacteria (Acharya et al. 2019). However, culturing alone cannot provide a complete picture of the microbial species in the sample, as most of the microbial species present cannot be cultured, either due to our incomplete knowledge of the conditions in which to culture the bacteria and/or culturing requiring more complex laboratory techniques (Stewart 2012). Understanding microbial ecology and diversity through culture independent approaches is therefore necessary to obtain a full representation of the communities

present in the sample (Michan et al. 2021) and the potential host communities of ARGs and MGEs. Since the Human Genome Project, our capability to characterise microbial communities through nucleotide sequencing has advanced significantly and is cheaper, more accurate and can provide results rapidly (Mardis 2011). DNA sequencing is frequently used in environmental AMR studies, offering a more efficient approach to capturing all genotypes and taxa in microbial communities (Slatko et al. 2018).

Techniques for sequencing include next-generation sequencing (NGS) (e.g., Illumina sequencing), shotgun sequencing (Chen and Pachter 2005) and MinION sequencing (Oxford Nanopore Technologies Ltd.). The 16S rRNA gene is usually used to distinguish taxonomic species, as it is highly conserved amongst organisms and the hypervariable region is distinguishable at individual taxonomic levels (Janda and Abbott 2007). Typically, the hypervariable region is amplified through polymerase chain reaction (PCR) and the amplified product is sequenced, where taxonomy is determined through comparison to a reference database (Mohsen et al. 2019).

Our understanding of natural microbiomes and their interactions with human health has been greatly improved through projects such as the ‘Earth microbiome project’ (EMP), a project which applied mass spectrometry and DNA sequencing of crowd-sourced samples to understand microbiomes across the globe in various environments (Thompson et al. 2017).

Microbial communities can be influential on AMR in the environment (Forsberg et al. 2014; Yu et al. 2020). For instance, identifying potential host species of key ARGs can be important to limit their dissemination in the environment. Currently, the most effective method for identifying host species is through microbial culturing and sequencing of isolates or through statistical analysis of co-occurrence patterns amongst ARGs and microbial taxa, as similar abundances indicate microbial taxa are carrying these ARGs (Forsberg et al. 2014; Li et al. 2015). However, identification of host-species through co-occurrence analysis requires a correct understanding and interpretation of sequencing data, which has been the source of much debate.

Datasets from sequencing are compositional, therefore the extent of information from a sample is dependent on the read depth (i.e., the extent to which a sample is sequenced; (Gloor et al. 2017)). Statistical measurements such as ANOVA or Kruskal Wallis on compositional datasets can lead to false discovery rates (Mandal et al.

2015; Weiss et al. 2017). In environmental studies, rarefaction has been frequently used to correct this problem, where samples are corrected to a common read depth (often the minimum read depth) (McMurdie and Holmes 2014; Gloor et al. 2017). However, this approach means data from samples can be omitted, and samples with different read depths can no longer be compared (McMurdie and Holmes 2014).

Quantitative Microbial Profiling (QMP) was introduced as an effective approach to overcome the weaknesses of rarefaction, whilst also allowing to estimate quantitative abundance of microbial communities (Vandeputte et al. 2017; Ott et al. 2021b) (Figure 2-4). First introduced by Vandeputte et al. (2017), this approach rarefies to the lowest sampling depth (i.e., sequencing depth divided by cell counts). The QMP approach was recently applied to environmental microbiomes and was found to be more effective than traditional rarefaction approaches (Ott et al. 2021b). Ott et al. (2021b) compared the traditional rarefaction approach (relative microbial profiling (RMP)) to QMP analysis on river water samples collected in Malaysia. They found the correction to QMP made it possible to observe the differences in diversity of the microbiomes in the samples, as well as compare the absolute taxa abundance more accurately.

One limitation introduced through the QMP approach is through estimating cell counts, which is frequently completed through qPCR of the 16S rRNA gene and converted using an estimate of 4.1 16S rRNA per bacterial cell (Klappenbach 2001). This method for cell estimation can lead to variable results between studies, where the qPCR result may deviate slightly depending on the method of DNA extraction, as well as the choice of primers (Morton et al. 2019). An alternative method to determine cell counts is through flow cytometry, which is more expensive and has lower throughput compared to qPCR (Morton et al. 2019). Despite the limitations, overall, the QMP approach proved to be more accurate to microbial community assessment in comparison to traditional rarefaction approaches (Ott et al. 2021b).

Like microbial community assessment, there are a wide range of culture dependent and culture independent approaches for monitoring AMR. Culture-based approaches include culturing indicator organisms on plates or filters, which identifies phenotypic resistance. For example, the World Health Organisation (WHO) Tricycle Protocol for monitoring AMR in the environment recommend culturing extended spectrum beta-lactamase producing *Escherichia coli* (ESBL *E. coli*) which are resistant to third and

fourth generation beta-lactam antibiotics (WHO 2021b). Resistance and susceptibility to antimicrobials can also be determined through antimicrobial susceptibility testing, which indicates the strength of resistance, and is well standardised (Reller et al. 2009). Culturing has the benefit of being cost-effective, easy to implement and can provide replicable results with low-error levels (McLain et al. 2016). However, culture dependent techniques only describe the phenotypic resistance of organisms, and information on the full genetic diversity of resistomes (i.e., all ARGs and MGEs of the microbial communities) are not captured.

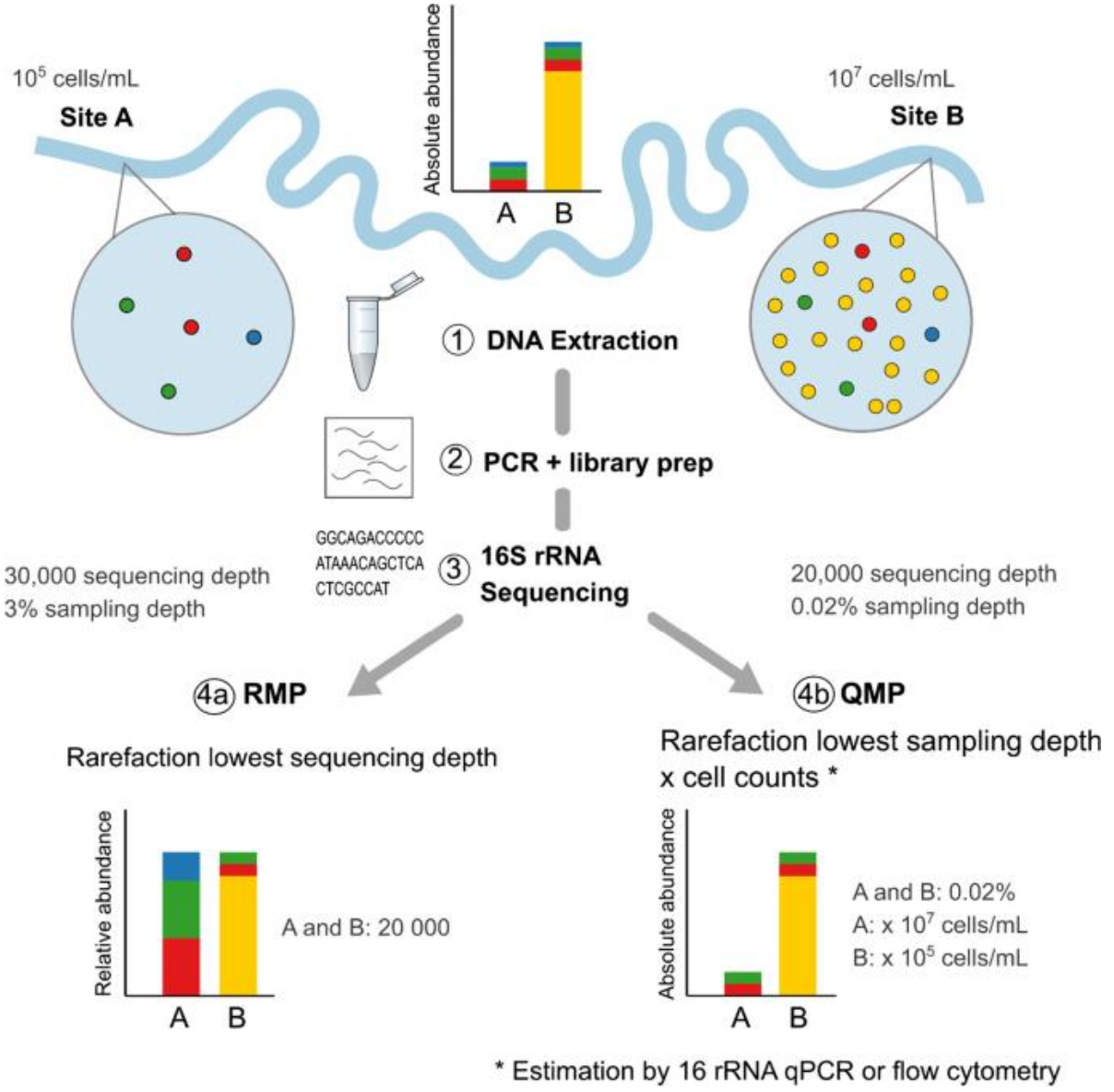


Figure 2-4 Schematic explaining the difference between relative and quantitative microbial profiling (RMP and QMP respectively). The RMP approach rarefies to the lowest sequencing depth per sample, whilst the QMP approach corrects for sampling intensity by rarefying to the minimum sampling depth and obtains the absolute abundances (per mL of river water) by multiplying by the estimated cell count. Figure from Ott et al, 2021b.

Genotypic methods have been applied either on their own, or to complement phenotypic data. This includes qPCR of target resistance genes, such as those which are indicative of anthropogenic resistance (e.g. *intI1* (Gillings et al. 2015), “anthropogenic” class 1 integrons (Quintela-Baluja et al. 2021) or *sul1* (Pruden et al. 2012a). These approaches can be used to indicate “hotspots” of AMR, where further phenotypic and genotype analysis could be done (Pruden et al. 2012a).

Comprehensive assessment of resistomes necessitates high-throughput approaches, including, metagenomic sequencing, shotgun sequencing and high-throughput qPCR (HT-qPCR). Sequencing based approaches can identify metagenomes, which can be compared against databases such as the Comprehensive Antibiotic Resistance Database (CARD) (Alcock et al. 2020). From a sequencing approach, novel ARGs/MGEs can be identified, whilst HT-qPCR uses known sequences and therefore cannot identify novel genes. (Pruden et al. 2012; Waseem et al. 2019). HT-qPCR however has lower detection limits compared to metagenomic sequencing and can identify very low relative abundances of target genes (Waseem et al. 2019; Pruden et al. 2021). HT-qPCR also only requires DNA at the nanolitre scale to perform a reaction, whereas metagenomic analysis requires a higher volume of sample (Waseem et al. 2019). Commercial services such as Resistomap Oy (Helsïnki, Finland), offer a customisable service, allowing the customer to select the genes to analyse, from a minimum of 12 to maximum of 384 genes. The disadvantages of HT-qPCR is that all assays within a run experience the sample qPCR cycle, which runs cycles and annealing temperatures which are not optimal for most genes (Sipos et al. 2007). In addition, both metagenomics and HT-qPCR can be costly, which is a barrier for implementation in AMR monitoring and research, particularly for developing countries (Waseem et al. 2019).

A recent review relating to AMR monitoring highlighted the inconsistency between methodologies and the need to establish a clear purpose, methodology, target environments, end-users, and integration across the One Health lens (Bengtsson-Palme et al. 2023). Further research utilising phenotypic and genotypic measurements for microbial communities and resistomes are required to better understand typical AMR levels across various environments and ensure more targeted and cost-effective routine AMR monitoring can take place (Bengtsson-Palme et al. 2023).

2.4 Uncovering the drivers for AMR in the environment

2.4.1 Anthropogenic drivers

The water cycle consists of rainfall, surface runoff, soil infiltration, retention in soil and plants, evaporation, and infiltration across the environment. Disruption to the natural water cycle and allochthonous inputs of bacteria and ARGs primarily occurs in urban, industrial, and agricultural settings (Almakki et al. 2019). The aquatic environment plays an important role as a pathway for the transmission and spread of ARGs and bacteria (Hooban et al. 2020). The majority of AMR research in the aquatic environment focuses on rivers, likely due to the essential role they play in aquatic ecosystems. Rivers can act as conduits in which AMR may be spread and stored (Figure 2-5) (Vaz-Moreira et al. 2014), link habitats and promote nutrient cycling, aeration, and the assemblage of microbial communities (Doretto et al. 2020). As well

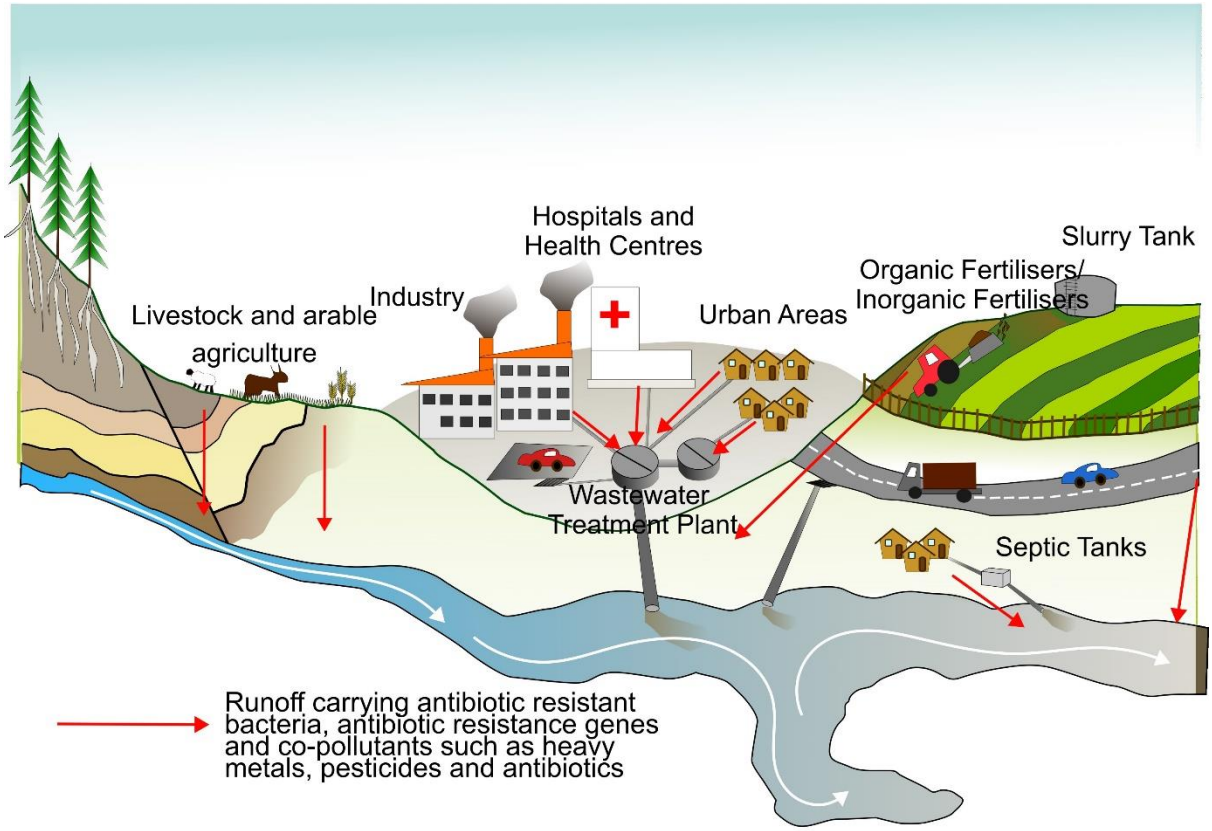


Figure 2-5 Sources of environmental antimicrobial resistance (AMR) and pathways into the environment (modified from Almakki et al. 2019)

as impacting surface water, high abundances of ARGs and resistant bacteria found in groundwater has been previously associated with high faecal contamination from humans and animals on the surface (Szekeres et al. 2018), which is particularly

problematic as groundwater is commonly used as a source of drinking water (Zainab et al. 2020).

Anthropogenic activity has the potential to expand the resistance present in environmental reservoirs (i.e., the resistome). Ingested antibiotics are incompletely metabolised (Serwecińska 2020), meaning after excretion, ARGs and resistant bacteria are released into the sewer network and then into the aquatic environment if inadequate wastewater treatment is provided. Due to the potentially high contamination from wastewater effluent, especially when not adequately treated, usually AMR in river catchments has been studied in “contaminated” environments (Chen et al. 2013; Amos et al. 2015; Koczura et al. 2016; Karkman et al. 2019; Roberto et al. 2019). It was even suggested that the level of anthropogenic activity and urbanisation in a catchment is a more important driver for environmental AMR than bacterial community interactions (Peng et al. 2020).

In terms of antibiotic selection driving ARG expansion and dissemination, it was found rivers in Asian countries had the highest antibiotic load in rivers (Singh et al. 2019). This was found to be due to higher antimicrobial use in Asian countries, less effective wastewater treatment and contamination from the pharmaceutical industry (Singh et al. 2019). While European and Australian rivers were found to have relatively low antibiotic loads compared to other continents (Singh et al. 2019), clinically important antibiotic resistant bacteria (ARBs) have been reported in highly urbanised areas in Europe, including fluoroquinolone resistant *E. coli* of O25b:H4-St131 lineage in the Thames River, UK (Dhanji et al. 2011), clinically relevant ARGs in a wastewater effluent contaminated lake in Switzerland (Devarajan et al. 2015) and CTX-M producing bacteria isolated from a river polluted by multiple wastewater treatment plants in Portugal (Tacão et al. 2022). In the Netherlands, which is a country with low antibiotic use, contamination with wastewater has been shown to significantly enhance the level of antibiotics and ARGs in the receiving rivers, sometimes persisting as far as 20 km downstream from the point of release (Sabri et al. 2020).

With antibiotic load in environments linked to antibiotic use, reducing the problem of AMR in the environment requires effective wastewater treatment as well as antimicrobial stewardship. In clinical settings, the UK government are aiming to improve surveillance efforts to improve antimicrobial stewardship and use. This

includes the formation of the English Surveillance of Antimicrobial Utilisation and Resistance (ESPAUR) group (Ashiru-Oredope et al. 2015). The ninth annual report of ESPAUR, demonstrated that the number of severe antibiotic resistant infections, including bloodstream infections, surgical site infections, urinary tract infections and skin/soft tissue infections had increased in England from 2014-2021, with only a temporary decrease in 2020 due to the impact of the COVID-19 pandemic (Ashiru-Oredope et al. 2023).

Reducing antibiotic and ARG load in rivers necessitates ensuring effective antibiotic stewardship in livestock as well as humans. Antibiotics are used in livestock-based agriculture for therapeutic use, metaphylatic use (i.e., the presence of clinical illness in one animal necessitates treatment for the whole herd/flock), and in some cases prophylactically for disease prevention or growth promotion (Woolhouse et al. 2015). Due to the food chain and interaction with humans, the use of antimicrobials in agriculture can have direct human health impacts. For example, the ban of the antibiotic avoparcin (an antibiotic in the same family as vancomycin) in animal feed in European countries in the 1990s triggered the reduction of vancomycin-resistant *Enterococci* in food and humans (Klare et al. 1999). In addition, the carriage of the ARG *mcr-1* in humans, which confers resistance to the antibiotic colistin (which is typically referred to as a ‘last resort antibiotic’ (Watkins et al. 2016)), has been shown to have probably originated from colistin use in agriculture (Liu et al. 2016).

Specifically in the UK, the use of antibiotics in agriculture has decreased by 55% between 2014 and 2021 (VMD 2022). In January 2022, regular use of antibiotics was banned in the European Union (EU), where use is now only permitted in the case of individual animal treatment (Coe et al. 2023). In addition, antimicrobials can no longer be used in agriculture to compensate for low quality animal husbandry practices and hygiene (Coe et al. 2023). Whilst this ban does not apply to the UK, the UK has previously stated it would align with EU regulations in the UK government’s 2019 5-year national action plan for AMR (GovUK 2019), but this was updated to state they would be following ‘similar’ provisions (GovUK 2022). Despite these reductions, an investigation of AMR in cows and sheep found that whilst antibiotic use was decreasing, resistance remained high against commonly used antibiotics (Hennessey et al. 2020).

AMR from livestock can spread through faecal contamination on fields, where faeces from livestock have been shown to contain an abundance of ARGs and MGEs (Xiao et al. 2016; Wepking et al. 2017; He et al. 2019; Wang et al. 2023). In addition, farm slurry is spread onto agricultural land, with the intention of maintaining soil quality and nutrient levels; however, this has been shown to result in an increase of ARGs to the local environment through soil runoff (Xiang et al. 2018; Swift et al. 2019). Other on-farm practices, such as rearing in pens versus open range animal management, also impact environmental AMR (Wang et al. 2023).

2.4.2 Environmental drivers

The previous section demonstrated that environmental AMR is often studied in the context of viewing ARGs as emerging environmental contaminants from wastewater and agriculture (Pruden et al. 2006). Such assessments must be considered relative to what the ‘background’ resistance looks like, but pristine rivers (i.e., those with little to no anthropogenic impact), are very rare. Even environments which have been previously considered pristine, such as Antarctica, have since been contaminated through anthropogenic activity, due to local wastewater releases (Hwengwere et al. 2022). This is also the case for the High Arctic, which showed signals of anthropogenic resistance, likely carried by migratory birds and humans visiting the region (McCann et al. 2019).

For truly understanding AMR in the environment, there is a need to characterise intrinsic resistance, such that ‘background’ resistance can be differentiated from resistance due to specific anthropogenic pressures, which extend resistance levels above background. This has been attempted through monitoring of single/isolated river catchments with various land-uses, including capturing a site devoid of human activities to inform on background AMR (usually the site further upstream of the catchment) (Pruden et al. 2012; Ott et al. 2021a). Anthropogenic AMR can also be distinguished through contrasting environments with different levels of anthropogenic exposure, such as comparing an Indian lake subjected to industrial pollution and fluoroquinolone antibiotics, with a less contaminated lake in Sweden (Bengtsson-Palme et al. 2014), or quantifying changes in AMR at remote pilgrimage sites along the Ganges where outside visitors are only seasonally present (Ahammad et al. 2014).

Studies have attempted to assess the environmental and climatic drivers for environmental AMR. Seasonality for example has been frequently assessed for the effect on environmental AMR (Keen et al. 2018; Son et al. 2018; Harnisz et al. 2020; Yang et al. 2022), with mixed results. Higher summer temperatures and more sunlight have been frequently associated with the proliferation of total bacteria and ARGs (Luo et al. 2010). In addition, increased organic matter has the potential to enhance the presence of ARGs in rivers (Zhang et al. 2021). However, seasons with high rainfall have also been associated with higher levels of resistance, due to the impact of surface runoff (Di Cesare et al. 2017) and storm drain outfalls or combined sewer overflows (CSOs) (Baral et al. 2018b; Honda et al. 2020). In addition, increased antibiotic consumption in winter months and resulting wastewater effluent contribution, has resulted in a greater environmental contamination in winter (Yan et al. 2013).

Seasonality in a catchment indirectly impacts the level of organic matter in the aquatic environment, where enhanced nutrients from organic matter has been previously associated with increased resistance in rivers (Zhang et al. 2021; Wu et al. 2023) and soils (McCann et al. 2019). The nutrients phosphorous (P), carbon (C) and nitrogen (N), coupled with light conditions, are essential in driving microbial ecology and interactions (Sterner and Elser 2017; Jarvie et al. 2018, 2020). The availability and scarcity of these nutrients is typically described through a stoichiometric ratio (i.e. C:N:P ratios) (Sterner and Elser 2017; Jarvie et al. 2018).

Previous studies have demonstrated the role anthropogenic disturbance plays in C:N:P ratios in rivers in the United Kingdom, where high nutrient contributions were associated with increased agriculture and land-use, whilst nutrient limitation is associated with less disturbed sites (Jarvie et al. 2018). In bacterial population studies, it has been demonstrated that the demand and rate of consumption of a nutrient will influence the predominant taxa in an environment, and the limitation of this nutrient will enhance competition (Tilman 1977; Ghoul and Mitri 2016). Whilst it has been recognised the essential role rivers play in transporting and cycling nutrients in the terrestrial and aquatic environment, the role nutrient availability plays in influencing the microbial community and persistence of resistance genes has not been previously explored, particularly at less impacted rural sites.

Whilst there is lack of research in rural sites in the United Kingdom, there is research into the environmental drivers of AMR in the Thames River, which partly goes through the city of London. This assessed class 1 integron and beta-lactam resistance and generated a predictive model attributing the likely sources of resistance (Amos et al. 2015). The model revealed the impact of wastewater treatment accounted for 49.5% of the variance in resistance levels, whilst heather and grassland land-uses had a strong negative effect on resistance levels, except in winter months. Precipitation was found to be a strongly influencing factor on AMR, depending on the land-cover.

Additional studies into AMR in aquatic environments in the United Kingdom, focus on bathing waters, due to the likely exposure pathways to humans. For example, an epidemiological survey of people who frequently used bathing water recreationally in the UK, indicated people who take part in water sports sessions are likely to be exposed to *E. coli* harbouring ESBL genes such as $bla_{CTX\text{-}M}$ (Leonard et al. 2018). An assessment of *E. coli* isolates of coastal bathing waters in Ireland similarly found isolates harbouring carbapenemase genes, including $bla_{NDM\text{-}1}$ and $bla_{OXA\text{-}48}$ (Hooban et al. 2021). Freshwater samples assessed through this study also found a high percentage of tetracycline resistant isolates. There has however been limited work assessing environmental drivers of AMR in riverine environments in the United Kingdom, especially those primarily impacted by rural runoff and waste exposures.

2.5 Identifying sources of AMR through microbial source tracking

2.5.1 Indicator organisms

Understanding the movement of faecal organisms in the environment is of utmost importance, where poorly treated or untreated sewage in the environment puts humans at great risk of exposure to pathogenic organisms. In addition, faecal pollution has been frequently associated with increased levels of AMR (Derx et al., 2023; Yu et al., 2020a).

Microbial indicators are habitually used for the identification of faecal pollution in water quality studies. Measurement of faecal indicator bacteria (FIB), can characterise the extent of faecal pollution, but doesn’t provide information on the potential origin (Hagedorn et al, 2011). Faecal pollution from human origin has been associated with higher health risks than that from animal or agricultural operations (Hagedorn et al., 2011). In addition, preventing further contamination necessitates identifying the source of pollution, so remediation action can be taken.

Microbial Source Tracking (MST) is an emerging discipline of biology which has several use cases, including but not limited to discriminating the sources of pollution in beaches (Jardé et al. 2018; González-Fernández et al. 2021), public health monitoring (Stewart et al. 2006) and for river catchment studies (Stapleton et al. 2007; Ballesté et al. 2020). MST first began through the identification of potential indicators which could be used to discriminate sources. For example, studies have applied faecal markers which can distinguish between two or more sources of faecal pollution. Organisms which have been applied for MST include *Bacteroidetes* spp. and *E. coli* (Bernhard and Field 2000b, a). *Bacteroidetes* has been a particularly effective MST marker due to their longer survival time in the environment and ease of detection (Bernhard and Field 2000b).

As previously discussed, development of sequencing technologies means it is possible to gain an insight into composition of whole microbial communities. Therefore, using whole microbial communities from sources, as well as sinks has been previously found to be an effective method for source apportionment. Previous work into this uses naïve Bayes, which estimates the likelihood that one community came from one source as opposed to another (Greenberg et al. 2010). This was however found to be inaccurate when there were unknown sources in the sample, inflating the apportionment of the specific sampled sources (Knights et al. 2011). Estimations based on Random Forest provided a more reasonable estimate (Knights et al. 2011).

2.5.2 SourceTracker for microbial source tracking

A commonly used method for whole community source tracking is the programme SourceTracker, a computational tool which estimates the relative contributions of mixed sources into a mixed sink environment (Knights et al. 2011). SourceTracker uses Markov chain Monte Carlo (MCMC) methods to explore the assignments of species to source environments within a given sink sample. This tool uses Gibb's sampling (MCMC algorithm) to estimate the source proportions and allocates the unexplained OTUs into sinks as from an 'unknown source'. In some studies, this method has been used in conjunction with a Bray Curtis dissimilarity index to determine the sources/sink similarity (Quintela-Baluja et al. 2019). SourceTracker is capable of sensitivity adjustment through parameter tuning and Gibbs sampling. However, these features add to running times and can be difficult to operate with large datasets (Carter et al. 2019).

Within the context of AMR, SourceTracker has been used to determine the sources of ARGs in river catchments. The application of SourceTracker in an urban river catchment in Lincoln (Nebraska, US) was demonstrated, where it was found that storm drain outfalls were found to contribute highly to microbiomes and ARGs in the environment, at around 54-57% (Baral et al. 2018a). This was later expanded to investigate wet and dry flows in the same catchment, revealing wet weather resulted in higher *E. coli* levels, and an increased contribution of stormwater, which was primarily derived of street sweepings (Baral et al. 2018b). SourceTracker was also applied to rural environments, where it was used to apportion ARGs in river sediments from sources such as chicken manures, pig manures and wastewater effluent (Chen et al. 2019b). The results were then compared with the human faecal marker crAssphage and indicated human faecal pollution contributed highly to microbes and resistomes, whilst animal manure had negligible contribution (Chen et al. 2019b).

2.5.3 Fast Expectation-Maximisation (FEAST) for microbial source tracking

In 2019, the Fast Expectation Maximization for Microbial Source Tracking (FEAST) was proposed as a method for whole community source tracking (Shenhav et al. 2019). This uses the expectation-maximisation algorithm which is a method for estimation in models where observed data is incomplete, which can hinder the accuracy of probabilistic methods such as MCMC (i.e., the approach used by SourceTracker) (Moon 1996; Do and Batzoglou 2008). Similarly to SourceTracker, it takes the sink microbial community and a separate group of source environments and estimates the fraction of the sink community that can be attributed to the source environments (Shenhav et al. 2019). It also can apportion sink communities into their source components 30-300 fold faster than SourceTracker and it has been found to be more accurate than SourceTracker and Random Forest approaches (Shenhav et al. 2019).

In some cases, SourceTracker has been shown to underestimate the proportion of unknown sources, when it is difficult to distinguish between source and sink communities (Shenhav et al. 2019; Yu et al. 2020). FEAST on the other hand, regards an unknown relative abundance for each source and adjusts its estimates to increase the contribution of the unknown sources to reduce the variability (Shenhav et al. 2019). FEAST also considers the uncertainty of sources, as opposed to directly

transforming them to relative abundance, and models species and genes together instead of treating them as separate entities (Shenhav et al. 2019)

There have been numerous environmental studies which have used FEAST analysis. Within the context of AMR, metagenomic analysis and determination of a wide depth and range of ARGs in environmental settings has permitted the application of FEAST to determine the origin of ARGs. For example, FEAST was used to apportion sources of ARGs and MGEs in the sediments of Lake Baiyang from the Fuhre River (Northern China) (Yu et al. 2020). Another study utilised FEAST to determine the sources of microbes, resistomes and virulence factors on microplastics in the environment (Li et al. 2022a). Rural land-uses were found to contribute highly, with the influence of urban land-uses increasing further downstream the river (Li et al. 2022a). FEAST has also been used to show the impact of environmental processes other than faecal contamination on ARG distribution in a river system. For example, a study investigating wastewater effluent on the receiving river found most of the downstream river was sourced from the upstream river (88% contribution), not wastewater effluent (Zhang et al. 2022). The low contribution of wastewater effluent to the river resistome was interpreted to be a result of rapid mixing and rapid advection of the wastewater effluent, making HGT of ARGs across taxa impossible (Zhang et al. 2022).

A limitation to SourceTracker and FEAST is the input requirement for sources sampled in a specific study. The use of microbiome data repositories such as EMP for source allocation has previously been limited by the wide geographic variability of sequencing data for source samples (Scott et al. 2002). Where specific study source samples are not available, a recently published scalable microbial source tracking tool STENSL (Microbial Source Tracking with Environment Selection) may be applied (An et al. 2022). STENSL was introduced to extend MST analysis through an unsupervised exploration of multiple sources available on publicly available microbiome data repositories (An et al. 2022). This approach has however yet to be utilised to the same extent as FEAST analysis, and has not yet been used in river studies.

Whilst both SourceTracker and FEAST have been regularly used in environmental studies, there are a lack of studies using these approaches in environments which are not frequently exposed to anthropogenic contamination (Hooban et al. 2020). Such studies would provide insight into the drivers of AMR in more natural

environments where the sources of AMR are more nebulous and less distinct, as opposed to urban environments with abundant point sources.

2.6 Modelling AMR

2.6.1 Modelling AMR for health protection

The use of mathematical models for predicting AMR patterns and spread is a rapidly developing field and has become highly useful for supporting decisions to inform public health (Opatowski et al. 2011; Spicknall et al. 2013). Models have been used to identify where further data and research is required, as well as to predict and manage risk (van Leeuwen and Vermeire 2007).

A previous systematic review found most AMR modelling studies were in relation to human health, modelling the development and dissemination of AMR (Birkegård et al. 2018). Modelling AMR development has several useful health protection applications, such as exploring the development of secondary infections of tuberculosis bacteria with immunocompromised individuals and modelling the likely outcome of antibiotic treatment (Alavez-Ramírez et al. 2007). Models which can predict the spread of AMR have applications in nosocomial settings, where AMR spread is likely. Research relating to application of models in clinical care have modelled the outcome of persistent antibiotic treatment of antibiotic wild-type and resistant strains, which can be used to assess whether additional treatment options are required (Wang et al. 2017). Additional clinical AMR modelling studies have used hospital size and length of stay as confounding variables to predict the emergence of AMR, where larger hospitals were found to strongly correlate with sensitive and resistant infections (Kouyos et al. 2011). This model also demonstrated the importance of environmental reservoirs, because the result showed that environmental AMR is particularly important in communities with small hospitals (Kouyos et al. 2011).

2.6.2 Modelling AMR in river systems

Despite the importance of environmental AMR, there have been relatively few models developed that model AMR pathogen fates in the environment, i.e., a systematic review found over 89% of AMR modelling research was related to human hosts (Niewiadomska et al. 2019). However, modelling AMR in river systems has recently received a lot of research interest (Hellweger et al. 2011; Hellweger 2013; Gothwal and Thatikonda 2018, 2020; Van Heijnsbergen et al. 2022; Jampani et al. 2023;

Niebaum et al. 2023). Modelling AMR in river systems is complex and requires incorporation of several factors (Figure 2-6). One of the first AMR modelling studies in river environments aimed to explore tetracycline resistance in a river using a simple mechanistic model (Hellweger et al. 2011). The model predicted river concentrations of tetracycline antibiotics in the Poudre River (Colorado), and considered particulate and dissolved organic matter, as well as the water and stream bed compartments. The model was later expanded to incorporate the influence of metal co-selection, where it was found that copper (Cu) and zinc (Zn) could explain the presence of resistant bacteria (Hellweger 2013).

A mathematical AMR model to predict the concentration of fluroquinolone resistant bacteria in the Musi River, a tributary to the Krisna River in India has been developed (Gothwal and Thatikonda 2018). The development of this model was motivated by the heavy contamination of the tributary and considered variables for organic matter, fluoroquinolones, heavy metals, and susceptible and resistant bacteria both in the water column and in the sediments. The model was used to simulate different pollution scenarios and found that antibiotics and organic matter were strong drivers for the presence of resistant bacteria. Both this model and the model from Hellweger (2013) incorporated natural stream transport processes, such as advection, dispersion, adsorption, degradation, settling, resuspension, and diffusion, as well as biological processes such as bacterial growth and plasmid HGT.

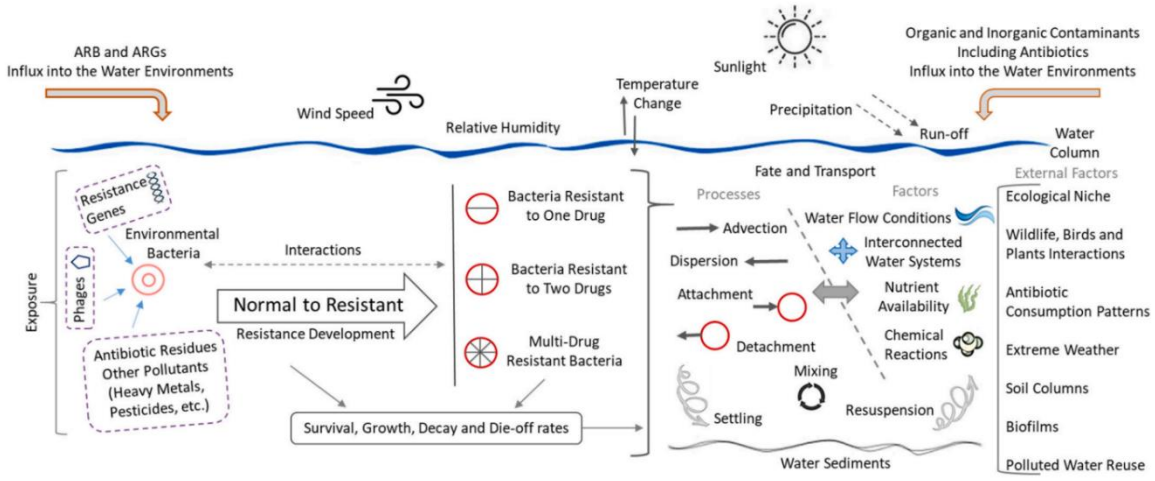


Figure 2-6 Complexity of modelling antimicrobial resistance in environmental settings (from Jampani et al, 2023)

In a later study Gothwal and Thatikonda (2020) incorporated stochastic differential equations into the model to account for random variability of anthropogenic and

environmental factors. The authors found that stochasticity in the environment had no impact on the presence of resistant bacteria in river sediment (Gothwal and Thatikonda 2020).

Models that combine Geographic Information Systems (GIS) can be useful for investigating the spatial and temporal variability of any natural system (Feijtel et al. 1997). For example, the GREAT-ER (Geo-Referenced Regional Exposure Assessment Tool for European Rivers) model, a GIS based chemical exposure prediction tool, was built as an Add-In to ArcGIS and has been used as a prediction tool at the catchment scale. This model has been widely applied in European studies of pollutants in river systems, such as cleaning agents (Schulze and Matthies 2001; Sabaliunas et al. 2003), heavy metals (Fox et al. 2000; Hüffmeyer et al. 2009), pharmaceuticals (Schowanek and Webb 2002; Burns et al. 2018) and antibiotics (Archundia et al. 2018; Zhang et al. 2020).

Adaptions of GREAT-ER have been used to predict the fate of antibiotic resistant *E. coli* (Van Heijnsbergen et al. 2022; Niebaum et al. 2023). This study applied the GREAT-ER model as a case study in the Dutch-German catchment of the Vecht River, representing dry summer simulations and typical flow conditions (Van Heijnsbergen et al. 2022). A later improvement of the model incorporated uncertainty and variability using Monte-Carlo simulations (Niebaum et al. 2023). This latter model considered emissions from WWTPs as well as diffuse emissions from runoff, soil, wildlife and resuspension and mobilisation of *E. coli* from the bottom sediments (Niebaum et al. 2023). Overall, it was found that the GREAT-ER model predicted *E. coli* concentrations with a slightly higher range of concentrations than observed values (Niebaum et al. 2023). Environmental parameters which were highly influential on *E. coli* concentration in the catchment included flow rate and *E. coli* removal efficiency in WWTPs.

2.6.3 The Soil and Water Assessment Tool

Despite the successful application of GREAT-ER to model antibiotic resistant *E. coli*, this adaptation of the model is not yet open source, therefore it is not readily available to end-users, such as policy makers or river catchment scientists. Open-source GIS-based models, such as the Soil and Water Assessment Tool (SWAT), developed by the United States Department of Agriculture (USDA), have an integrated bacteria sub-module. SWAT was initially developed for the purpose of

predicting the temporal impact of land-management practices on water (Neitsch et al. 2011) and is a useful tool that can be applied to numerous disciplines, with its successful and diverse applications evident from the specialised SWAT international conferences (Abbaspour et al. 2017).

The SWAT model was initially developed as a response to the Clean Water Act in 1972, which aimed to protect natural waterbodies from chemical and biological pollution and urbanisation (Saleh and Du 2004). The Hydrological Simulation Programme Fortran (HSPF) model developed by the Environmental Protection Agency (EPA) has similar capabilities, but research has found its calibration process to be difficult and time consuming, requiring numerous input parameters (Saleh and Du 2004; Xie and Lian 2013). In addition, the SWAT tool has been suggested as a more appropriate method (compared with the HSPF model) where observed information required for calibration purposes is scarce (Xie and Lian 2013).

SWAT models have been applied in numerous settings, including to assess the climatic and human induced changes on water resources and streamflow (Milewski et al. 2019; Oo et al. 2020), agricultural and human pollution (Taylor et al. 2016) and to inform management practices. Oo et al. (2020) used different climate change scenarios to predict the impact on flow rate using a SWAT tool in a catchment in Myanmar. The model demonstrated the potential for flooding and low stream flow with increasing seasonal extremes (i.e., wetter winters and dryer summers). Application of the SWAT model to investigate nitrogen and phosphorous diffuse pollution has been completed in studies in the UK where Taylor et al. (2016) modelled the impact of mitigation measures on agricultural pollution, such as introducing red clover to crop rotations. This study demonstrated introducing red clover reduced phosphorous loss by 19.6%, whilst buffer strips resulted in a reduction of 12.2-16.9% (Taylor et al. 2016).

The SWAT model has been implemented for the purposes of simulating microbial concentrations in river systems, using the bacteria sub-module (Sadeghi and Arnold 2002). Using a SWAT model can be an inexpensive approach to estimating microbial loads and hotpots, and can be used to support in situ monitoring, and due to the integration of non-point sources such as agriculture and wildlife, it’s use is well suited to agricultural catchments (Jeong et al. 2019). An advantage of SWAT over HSPF for this use case is the simulation of persistent and less persistent pools of bacteria in

the same model run. For example, a SWAT model was developed for the Upper Salem River Watershed in New Jersey which simulated persistent bacteria (*E. coli*) and less persistent bacteria (faecal coliforms) in one model run (Niazi et al. 2015).

SWAT has been used to model bacteria in remote locations such as the Athabasca River Basin in Alberta, Canada, where there are only a few in situ monitoring stations (Meshesha et al. 2020). The model was used to identify the influence of pH on *E. coli* concentrations, which was found to be substantial (Meshesha et al. 2020). Management or remediation strategies on microbial water quality can also be simulated using the SWAT model, for example, the influence of best management practices on *E. coli* levels in an agricultural catchment (Hernandez-Suarez et al. 2020). In this case, management practices such as septic tank restoration was found to be successful in achieving target reductions in *E. coli* concentrations (Hernandez-Suarez et al. 2020). Therefore, identifying the most effective remediation strategies through a SWAT modelling approach, can save cost and be used to provide estimates on the potential environmental impact.

Whilst the SWAT model was developed and designed to be used in river catchments in the United States, the model has been applied to river catchments internationally. For example, the SWAT model was applied to an agricultural catchment in West Ireland, where *E. coli* was modelled through simulating inputs from cattle and sheep manure and wastewater effluent (Coffey et al. 2010a, 2013). Whilst this model yielded acceptable estimates of *E. coli* concentrations in the river, the authors suggested more reliable input information could be used to improve the model accuracy (Coffey et al. 2010a), particularly for predicting daily loads (Coffey et al. 2013). An additional study in the same catchment applied the SWAT model to predict the parasite *Cryptosporidium*, demonstrating the adaptability of this model to different organisms of interest (Coffey et al. 2010b).

Whilst there has been substantive research in the application of SWAT models to predict bacteria concentrations, there is scant work on the modelling of AMR bacteria or genes in river catchments using SWAT models. A SWAT-ARB (Antibiotic Resistance Bacteria) model was developed in Virginia Tech University as part of a doctoral thesis, which was based on ARG data (Thilakarathne 2020). Whilst SWAT was able to model resistance, it's accuracy could not be determined due to the lack of observed data. This was later expanded on for a Masters thesis, also in Virginia

Tech University, which explored the application of the SWAT-ARB model to catchments in the US (House 2020). However more sampling is needed to validate its predictions (House 2020). The SWAT-ARB model is also not yet open source, which means it is not readily available for end-users such as policy makers, researchers or other scientists. The lack of availability of SWAT-ARB also limits the capability to improve the model.

2.7 Conclusions and key knowledge gaps

The discovery of antibiotics transformed healthcare. However, the emergence of resistance is a significant problem, which would require addressing clinical, animal and environmental exposures. AMR should therefore be monitored and researched through a One Health Lens, especially when studying AMR within systems with overlapping sources and sinks, such as rural river catchments without dominant sources.

In the United Kingdom, specifically there have been few studies investigating AMR transmission and spread in such rural systems. Therefore, there is need for work on AMR in rural catchments with different types of land-use to gain an understanding of what drives AMR in typical UK rivers without acute pollutant sources. In addition, analysis of rural catchments may reveal the local and regional drivers of AMR, such as climate or hydrology.

Microbial Source Tracking technologies have developed into useful tools for estimating point source contribution of microbial communities and AMR in river catchments. Whole community microbial source tracking technologies such as SourceTracker and FEAST have been useful to estimate source contribution using whole community data from sequencing. Understanding the primary sources of AMR in rural river catchments could provide some insight into where mitigation is best targeted. In addition, the factors that influence the persistence of non-native microbes and associated resistance genes, such as nutrient availability are not well understood.

The use of mathematical models for AMR is developing rapidly. For environmental studies, models which combine Geographical Information Systems (GIS) can be useful for considering spatial variability. The Soil and Water Assessment tool is widely used for simulating non-point sources, but there is little research on its application for

predicting AMR. Using a SWAT model to model AMR bacteria can be used to help guide monitoring and identify AMR hotspots.

Chapters 3-5 will attempt to address the research gaps presented:

Chapter 3 will aim to compare two rural river catchments in the UK (the Coquet (Northumberland) and the Eden (Cumbria) with similar land-uses and different regional characteristics to interpret the impact of hydrometeorological characteristics on driving microbial communities and AMR in a river system.

Chapter 4 focuses on the Coquet catchment to assess the contribution of point source samples at different sites along the river using the FEAST approach. This also looks at nutrient limitation, through nitrogen phosphorous (N:P) ratios, as an environmental factor that has potential to influence the persistence of AMR along the course of the river.

Chapter 5 describes the development of a SWAT hydrological model for the Coquet catchment, which simultaneously models the concentration of *E. coli* and ESBL *E. coli* in the river.

Chapter 3 Antimicrobial Resistance in Rural Rivers: Comparative Study of the Coquet (Northumberland) and Eden (Cumbria) river catchments

3.1 Introduction

Antimicrobial resistance (AMR) is recognised as a global health and societal issue. In 2019, it was estimated that 4.95 million deaths could be associated or directly attributed to AMR (Murray et al. 2022) and it has been conservatively projected that up to 10 million additional deaths per year might be expected by 2050 (O'Neil, 2014). As a response to growing concerns, the United Nations, led by the World Health Organisation, is developing an integrated surveillance programme for AMR and antimicrobial use (AMU), which was recently highlighted in the G20 Summit Leaders' Declaration aimed at "strengthening global health and implementing a One health approach" (UN, 2023). However, AMR prevails across human, animal, and environmental sectors, and surveillance must be inclusive, even in places where the consequences of AMU are less directly evident (UNEP, 2023).

AMR is intrinsic in the environment with antibiotic resistance genes (ARGs) that encode antibiotic resistance being found in ~30,000 yr-old glacial sediments (Dcosta et al. 2011). However, with intensive antibiotic use in clinical and agricultural settings and pollution, resistance has expanded across the biosphere via mutations and horizontal gene transfer (HGT), often on mobile genetic elements (MGEs), and microbial selection (Bengtsson-Palme et al. 2018). Anthropogenic activity, such as agricultural AMU (Xiang et al. 2018; Neher et al. 2020; Burch et al. 2022) and wastewater discharges (Zhang et al. 2022), has altered resistance across the natural environment, including the intrinsic resistome (all ARGs and MGEs present in nature prior to anthropogenic impacts) expanding across nature. The problem, however, is how to characterise such expansion. Most environmental studies focus on resistance "hot spots", which says very little about how we are changing nature.

One way of assessing intrinsic AMR is to study locations without extensive AMU or waste sources, such as rural river catchments that include different types of land use. River catchment studies can contrast potentially subtle effects of different inputs and land uses, with the river itself acting as a "biomarker" for natural AMR transmission and spread (Vaz-Moreira et al. 2014). Within this context, the United

Kingdom (UK) five-year National AMR Action Plan includes environmental AMR surveillance of rivers at the catchment scale, partly for better understanding of what intrinsic resistance looks like in the UK.

As noted, while ARGs and their microbial hosts in river settings have been well studied, focus is almost always on contaminated environments, including those impacted by wastewater releases (Dhanji et al. 2011; Amos et al. 2014; Devarajan et al. 2015; Tacão et al. 2022), agricultural activity (Seiler and Berendonk 2012), pharmaceuticals (Šimatović and Udiković-Kolić 2020; Wilkinson et al. 2022) and heavy metals that can drive co-selection for ARGs (Gupta et al. 2022; Zhang et al. 2023a). Also, most studies tend to focus on urban catchments and there is less data on rural and-or less impacted landscapes. Data in less impacted landscapes is critical to providing context for more impacted catchments, helping baseline the extent of anthropogenic impacts.

Finally, there is a lack of work that compares catchments with similar land-use to reveal underlying drivers of environmental AMR. Such comparisons are difficult, often due to inconsistency in sampling methods and analytical techniques between different environmental AMR research studies (Hassoun-Kheir et al. 2021). Investigations have been conducted in large catchments through comparing AMR in sub-catchments with land-use varying from pristine to highly cultivated (Mukherjee et al. 2020; Neher et al. 2020). These studies can provide a valuable perspective on the impact of anthropogenic AMR contamination with the context of more pristine land-use. However, comparisons of multiple river catchments with predominantly rural land-use, but different hydrometeorological characteristics, can be used to interpret the potential impact of local hydrology and climate as AMR drivers. Differences in rainfall and run-off will likely impact the microbiome and resistome in the receiving rivers, due to the influence of differential runoff (Almakki et al. 2019). The temporal distribution and intensity of precipitation within catchments and linking to hydrologic characteristics have not been reported in most studies (Hamilton et al. 2020).

In addition to understanding the distribution and abundance of ARGs in a catchment, analysis of the microbial taxa in resistome studies is important to reveal potential host species. Recent analysis of next-generation sequencing (NGS) data using

quantitative microbial profiling (QMP) showed that this was an effective approach to overcome the weaknesses of traditional normalisation techniques for NGS data, such as rarefaction (McMurdie and Holmes 2014), whilst also allowing quantitative abundance estimates of microbial communities to be made (Ott et al. 2021b). This approach, first introduced by Vandeputte et al. (2017) rarefies NGS reads to the lowest sampling depth (i.e. sequencing depth divided by cell counts), instead of the traditional minimum read depth (Gloor et al. 2017), which can lead to data from samples being omitted (McMurdie and Holmes 2014) and false discovery rates from subsequent statistical testing (Mandal et al. 2015; Weiss et al. 2017).

The objectives of this study were to compare the hydrology, microbiomes (using QMP approaches), and resistomes in two rural river catchments in the UK, using the Coquet and Eden Rivers as case studies (Figure 3-1). The differences between catchments were then assessed to identify drivers of rural AMR within UK catchments. Overall, this study demonstrates the benefit of using multiple case-studies in the assessment of environmental AMR and shows how this could inform AMR surveillance.

3.2 Materials and methods

3.2.1 Site Description and catchment sampling

The Coquet River catchment is in NE England (Northumberland) and spans from the Cheviot Hills to the seaside town of Amble (Figure 3-1B). It has numerous small towns, including Shillmoor, Sharperton, Thropton, Rothbury, Warkworth and Amble. The total Coquet catchment area is 606 km^2 and the length of the river is 60 km. In contrast, the Eden River catchment is in NW England (Cumbria) and it is larger than Coquet catchment, being 2324 km^2 in area (Figure 3-1B). The Eden is split into six sub-catchments, of which the Upper Eden (670 km^2) and Lower Eden (461 km^2) were selected for work here due to their similar size as the Coquet. In the Eden sub-catchments, there are several small towns, such as Kirkby Stephen, Appleby in Westmorland, and Temple Sowerby as well as the city of Carlisle near the Irish Sea. For practical purposes, these catchments were chosen due to their close geographic proximity and to assess catchments spanning the west-east precipitation gradient in the Northern UK (due to eastern tracking weather systems and orographic effects), which significantly impacts western vs eastern hydrology.

The Coquet and Eden catchment areas were extracted using ArcMap (Attal 2017; ESRI 2018), and land-use was classified using Land Cover Map (LCM) 2015 (Rowland et al. 2017) (Figure 3-1). Land-use types were grouped in ‘urban’, ‘rural’ and ‘pristine’ as shown in Appendix Table A-1. The percentage of land-use within a 2 km buffer around each sampling site was calculated as per Amos et al. (2015) (Figure 3-1C). Sample sites were selected to capture a variety of land-uses along the river, whilst also allowing sampling to be safe and logistically suitable. For example, most samples were collected at mid-stream from bridges.

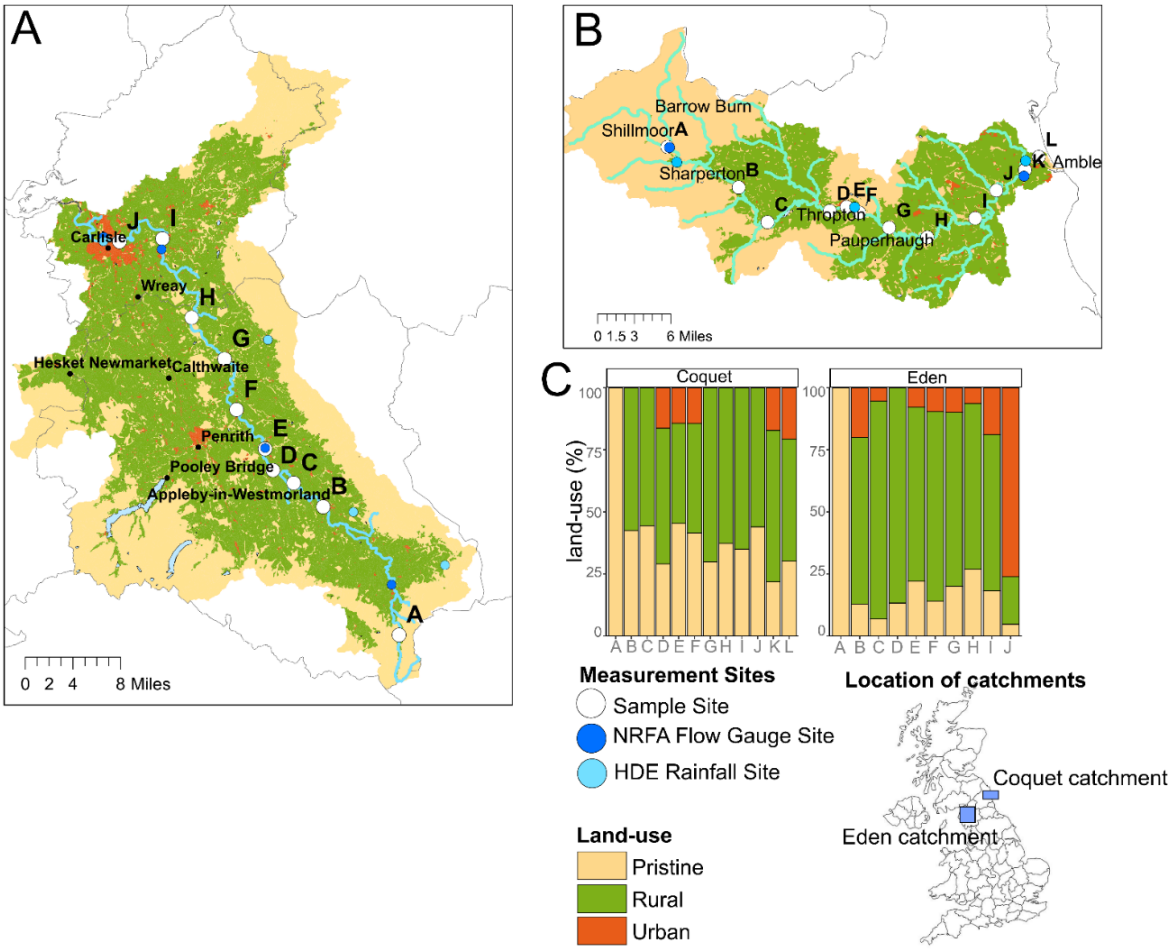


Figure 3-1 A) the Eden catchment map with sample locations and land-use B) The Coquet catchment map with sample locations and land-use C) Percentage of Pristine, Rural and Urban land-use within a 2 km buffer in sampling locations. Measurement sites include sample sites for this study, National River Flow Archive (NRFA) Gauge Site (see Section 2.7) and DEFRA Hydrology Database Explorer (HDE) Rainfall measurement sites (see Section 2.7). Land-use classifications according to Land Cover Map 2015 (Rowland et al, 2017) were grouped into Pristine, Rural and Urban according to Appendix Table A-1

Exact locations of sample sites are shown in Figure 3-1. Twelve and ten sampling sites were chosen for the Coquet and Eden, respectively, where Site A was sampled

farthest upstream, and Sites L and J were sampled farthest downstream. Sampling always was performed from up to downstream, over three separate days in each catchment on the specified dates in 2020 and 2021 (Appendix Table A-3). Sampling had been planned to take place over a shorter time, but sampling trips needed to be rearranged due to numerous disruptive lockdowns during the COVID-19 pandemic. The extended time between sampling campaigns may have introduced differences in microbial community and resistomes arising from seasonality.

When onsite, river water quality was assessed for temperature, dissolved oxygen (DO) and conductivity using an HQ40 portable multimeter (HACH) and pH using a 500 series portable pH meter (Jenway). River water samples were collected in triplicate, using a bucket cleaned with 70% ethanol solution between uses. Five litres were collected in total, with three litres being used for DNA extraction and two litres used for all other analyses. River volumetric flow rate was estimated in situ during two of the three sampling trips for each catchment (September and October/November), where river velocity was estimated using the float method (Jowett 1997; Michaud and Wierenga 2005).

The cross-sectional area of the river at each site was calculated using the measured river width and depth, which was multiplied by the surface velocity. A correction factor of 0.85 was applied to surface velocity data (Michaud and Wierenga 2005; Ott et al. 2021a). Flow rate in the March sampling trip was not performed due to including microbial plate colony culturing (Section 2.6), which required extra field- and lab-processing time.

3.2.2 DNA extraction and quantification of 16SrRNA, ARGs, MGEs and Microbial Source Tracking (MST) probes

For each site visit, three litres of river water (3 x 1 litre composite samples) were filtered through 0.22 µm cellulose filter paper (Merck Millipore), before subsequent processing. DNA was extracted from microbial cells trapped on the filter paper using the FastDNA Spin kit for soil (MP Biomedicals, UK). Following extraction, samples were assessed for purity using a NanoDrop 1000 Spectrometer (Thermoscientific, UK) and DNA concentration was measured using the Qubit® dsDNA High Sensitivity (HS) Assay Kits (Invitrogen, UK). The extracted DNA was diluted to 5 ng/µL for quantitative polymerase chain reaction (qPCR) analysis to minimise inhibition.

Human and ruminant *Bacteroidetes* MST quantification (HuBac and RuBac respectively) was performed using primers and probes supplied by the Environment Agency (EA) (Environment Agency 2008). For the MST probes, and 16S rRNA (total bacteria), Taqman qPCR reactions were conducted using SSoAdvanced™ Universal Probes Supermix (BioRad) (Appendix Table A-4). Faecal coliforms (Appendix Table A-4) were quantified using the qPCR SYBR green-based method assay. SYBR-green reactions were conducted using SSoAdvanced Universal SYBR® Green Supermix (Bio-Rad). Assays were completed in duplicate using the Bio-Rad CFX c1000 System (Bio-Rad), with a negative control.

3.2.3 Next-generation sequencing

Amplification of the 16S rRNA gene was confirmed through PCR and qPCR. The Illumina MiSeq platform at NU-OMICS, Northumbria University (UK) was used to sequence the hypervariable V4 region 515F-806R of the 16S rRNA gene with V2 500 cycle chemistry. Sample preparation and sequencing was conducted using the Schloss MiSeq Wet Lab SOP (Kozich et al. 2013), including a positive control (mock community, ZymoBIOMICS Microbial Community DNA Standard, Zymo Research) and negative control (H_2O).

Raw sequences were available as FASTQ files and were processed using QIIME2 (v. 2021. 4) (Estaki et al. 2020). Reads were denoised into Amplicon Sequence Variants (ASVs) with DADA2 (Callahan et al. 2016, 2017). Then Naïve Bayes classifiers pre-trained on SILVA 138 99% OTUs full-length sequences were used for taxonomic assigned to genus level. ASVs of <0.1% of the mean sample depth were removed to account for MiSeq bleed (Comeau et al. 2017). The taxonomy and ASV table biom file were produced for downstream analysis in R Studio (v.4.2.2), with the *phyloseq* (v.1.42.0) and *vegan* (v.2.6-4) packages. ASVs not classified to phylum level, mean value <1 and maximum value <10 were removed. This resulted in a total of 2991 taxa for 66 samples (compared to 4197 taxa pre-quality filtering), with a minimum of 4280, median of 26212 and maximum of 119457 reads.

3.2.4 Quantitative Microbial Profiling and Hill diversity analysis

Quantitative Microbial Profiling (QMP) was used to rarefy ASVs as described previously (Vandeputte et al. 2017; Ott et al. 2021b), using R programming from http://www.raeslab.org/software/QMP (Vandeputte et al. 2017). Samples were

rarefied to an equal sampling depth (i.e., sequencing depth divided by cell counts), using the R function *rarefy_even_sampling_depth* (seed=711). Following rarefaction, abundances were multiplied with an estimated cell concentration/mL of river water, which was calculated by dividing the 16S rRNA concentration/mL by 4.1, which is the estimated 16S rRNA gene copy numbers per bacterial cell (Klappenbach 2001).

Hill numbers were used to measure the species diversity. These have been used more frequently in macroecology as an improved method of determining species diversity and address the problem of rare taxa in diversity influences (Boeken and Shachak 2006; Chao et al. 2014; Alberdi and Gilbert 2019a). The influence of rare and abundant taxa are assessed through changing the order of diversity (q) where at q=1 relative abundances of ASVs are assessed according to their original values, at q<1 rare ASVs are overweighed and at q>1 highly abundant ASVs are overweighed (Alberdi and Gilbert 2019b). Hill numbers for QMP abundances were calculated using the *hilldiv* R package (v.1.5.1) (Alberdi and Gilbert 2019b) and diversity profiles were visualised using ggplot2 (v.3.4.2). As described in Ott et al. (2021b), the Sørenson-type over-lap dissimilarity measure at q=1 was plotted on a Non-metric Multi-dimensional Scaling (NMDS) plot, to visualise the proportion of nonshared ASVs between sample sites and between different sampling months.

3.2.5 Relative abundance of ARGs and MGEs using high-throughput qPCR

Resistomes were characterised through high-throughput qPCR (HT-qPCR), using the Resistomap Oy (Helsinki, Finland) SmartChip Real-time PCR system. DNA from 58 unique samples were analysed. Thirty-two representative samples were analysed for the full array of 384 ARG and MGEs offered by Resistomap (Appendix Table A-6). Based on these analyses, 96 genes were selected for analysis of the remaining samples (Appendix Table A-7). DNA samples were first diluted to 10 ng/µL based on concentrations measured with the NanoDrop 1000 Spectrometer (Thermoscientific, UK). DNA samples, qPCR reagents and primer sets were mixed in 100 nL reaction SmartChip™ wells, using the SmartChip™ Multisample Nanodispenser (TakaraBio). It should be noted that there were originally 66 collected samples, but eight had insufficient DNA for HT-qPCR analysis (Appendix Table A-5), including all the samples from Coquet Site A.

Data processing was performed through R programming. The threshold cycle (CT) of

28 was used as previously suggested (Stedtfeld et al. 2018). DNA were analysed in three qPCR reactions. Genes which were only present in one technical replicate (i.e. one out of three qPCR reactions) were excluded as false positives. For notable detections, quality control was assessed by checking for issues flagged by Resistomap, such as multiple melt curves. The abundance of each gene was calculated as the relative abundance in proportion to the 16S rRNA gene as previously described (Muurinen et al. 2017). ARGs and MGEs were transformed into absolute copy numbers by multiplying with 16S rRNA concentration for each sample.

3.2.6 Physiochemical analysis and microbial colony counts

Water samples were filtered through 0.22 µm polyethersulfone (PES) syringe filters within 24 hours of collection and then analysed for soluble chemistry within 48 hours. Assessment of water chemistry was conducted using HACH LANGE kits of Chemical Oxygen Demand (COD) (LCK 314: 15-150 mg/L), ammonium-nitrogen (NH_4-N) (LCK 304: 0.015-2 mg/L), orthophosphate as phosphorous (PO_4-P) (LCK 349: 0.05-1.5 mg/L) and Total Nitrogen (TN) (LCK 138: 1-16 mg/L) analysis.

Microbial plating and colony counts were conducted for the third sampling trip for both catchments (March 2021), focusing on non-resistant and resistant *Escherichia coli* (*E. coli)* isolation. River water was filtered onto 0.45 µm cellulose membrane filters (Sartorius™), which were placed on ChromoSelect agar for *E. coli* (Sigma-Aldrich), and incubated for 38 hours at 37 °C. This process was repeated, except filters were placed on Chromoselect Agar with extended beta-lactamase (ESBL) supplements (Sigma-Aldrich) (i.e., 1.5 mg of Ceftazidime and Cefotaxime, 1 mg of Ceftriazone and Aztreoname, and 5 mg of Fluconazole) to quantify *E. coli* resistance to ESBL antibiotics. The amount of water filtered was 2-20 mL for non-ESBL plates and 200-500 mL for ESBL plates. Each site was plated in triplicate.

3.2.7 Catchment hydrology

Three river gauging sites within the Coquet catchment and four sites within the Eden were identified (Appendix Table A-8) from the National River Flow Archive (NRFA; https://nrfa.ceh.ac.uk). For each site, the flowrate on each sampling day, various river flow indices, and catchment descriptors were obtained from the NRFA (Table 3-1). The flow indices were normalised by upstream contributing area to allow comparisons between sites. Daily rainfall data was downloaded for the seven

gauging sites from the DEFRA Hydrology Data Explorer (https://environment.data.gov.uk/hydrology/explore) (Appendix Table A-8). The Antecedent Precipitation Index over the previous five days (API_5) was calculated according to (CEH 2023). This provides a measure of catchment wetness.

Table 3-1 Definitions of Hydrological Parameters (from NRFA, 2023)

Parameter	**Definition**
Catchment area (km^2)	Area of the catchment upstream contributing to the site
PROPWET	The fraction of the year that soils can be expected to be quite wet, where saturated soils are more likely to contribute to flooding
BFI	Baseflow Index: Proportion of the total flow that comes from groundwater
SAAR (mm)	Average annual rainfall in the standard period (1961-1990) in millimetres
Mean Flow (m^3/s)	Record mean-gauged flow at gauging stations
Mean annual runoff (mm/year)	Long-term mean annual flow as measured at the gauging station normalised by catchment area
95% Exceedance (Q95) (mm/day)	Low flow parameter: flow which was equaled or exceeded 95% of the flow record
5% Exceedance (Q5) (mm/day)	High flow parameter: flow which equaled or exceeded for 5% of the flow record
Flow rate on the day of sampling (m^3/s)	Gauged flow rate on the day of sampling (From NRFA)

3.2.8 Statistical analysis

Data processing and statistical analysis was performed in the R environment (R Core Team 2018). Graphics were developed with *ggplot2* and finalised with Inkscape

(v.09.4). The Kruskal-Wallis test was performed to analyse differences between the Coquet and Eden catchments and sampling months for MST probe, colony count, and ARG abundance data. The Wilcoxon test was used to find significantly different abundances of taxa at Phylum level between the Coquet and Eden catchments (p<0.01). Otherwise, significance was defined as p<0.05.

The log_2fold change between ARG and MGE concentrations between the rivers Coquet and Eden was computed using the *DESeq2* package, which utilised the Wald test, and p value were adjusted according to the Benjamini Hocherg method (Benjamini and Hochberg 1995). The log_2fold change was then plotted against statistical significance with a volcano plot using *ggplot2*.

Quantile-Quantile plots (Q-Q plots) were used for the microbiome and resistome data to identify outlier samples. Two datapoints were excluded from the analysis which is discussed further in Section 3.3.2 and 3.3.3. Co-occurrence analysis was also conducted with the ARG and MGE genes, and abundance of ASVs by order level, to determine possible hosts for ARGs. This was conducted through an initial Spearman correlation, where significance values were adjusted according to Benjamini Hochberg (Benjamini and Hochberg 1995). Strongly positive correlations (r_s>0.8, p<0.01) were further visualised using network analysis based on the *igraph* R package and *Gephi* software (Bastian et al. 2009). Spearman correlations, with significance values adjusted according to the Benjamini Hochberg method (Benjamini and Hochberg 1995), were used to assess relationships between water quality, microbial, AMR, and hydrological indicators which were visualised using the *corrplot* (v. 0.92) package.

3.3 Results

3.3.1 Differences in hydrological conditions in the Coquet and Eden

To provide some context of the precipitation and hydrology of both catchments, flow data and FEH catchment descriptors were extracted from the UK Centre for Ecology & Hydrology Flow Archive (Table 3-2). As background, the Standard Average Annual Rainfall (SAAR) reflects altitude and the influence of the Atlantic Ocean, with the highest values in the Eden catchment, particularly in the Eamount at Udford and Kirkby Stephen sites.

Table 3-2 Catchment descriptors in the Eden and Coquet at relevant sites from the National Flow Archive (NRFA). Definitions of Hydrological Parameters are in Table 3-1

		Eden (Upstream-Downstream)				**Coquet (Upstream-Downstream)**		
		Eden at Kirkby Stephen	Eden at Temple Sowerby	Eamount at Udford	Eden at Sheep mount	Usway Burn at Shillmoor	Coquet at Rothbury	Coquet at Morwick
Catchment Area (km^2)		69.4	616.4	396.2	2286.5	21.4	346	569.8
SAAR (mm)		1492	1142	1768	1182	1056	905	850
PROPWET		0.65	0.66	0.66	0.64	0.45	0.45	0.44
BFI		0.25	0.37	0.51	0.48	0.38	0.47	0.44
Mean Annual Runoff (mm/year)		1195	786	1258	752	824	534	486
Mean Flow (mm/day)		3.27	2.15	3.45	2.06	2.26	1.46	1.33
95% Exceedance (Q95) (mm/day)		0.21	0.28	0.53	0.38	0.32	0.22	0.20
5% (Exceedance (Q5) (mmm/day)		13.07	7.68	10.60	6.45	7.91	4.45	4.52
Flashiness Q5:95 (-)		62.24	27.43	20.00	16.97	24.72	20.23	22.60
NRFA Flow Rate (m^3/s) on day of sampling	Sep	1.497	10.79	NA	55.7	0.621	2.47	2.72
	Oct/ Nov	5.168	40.89	NA	96.94	3.145	4.457	6.28
	Mar	1.221	13.31	NA	65.86	1.486	6.066	9.32

High SAAR is associated with long periods of soil saturation (PROPWET), enhancing the generation of saturation excess surface runoff and higher annual runoff. The ratio of Q5 and Q95 is a measure of flashiness, with high values indicating a flashier flow regime. In each catchment, SAAR, mean annual runoff, and flashiness decrease with decreasing elevation towards the catchment outfall. BFI values, the ratio of baseflow to total streamflow, are similar in the two catchments, except for Kirkby Stephen. This may be due to the hydrological influence of the

extensive peat soils, which enhance the flashiness of that sub-catchment. The Eamont flow regime is affected by the Haweswater and Wet Sleddale Reservoirs and Ullswater, which may impact on low and high flow values (characterised by Q5 and Q95) and BFI.

The calculated API_5 for both catchments at different rainfall gauges on the day of sampling is summarised in Appendix Table A-9, i.e., 1.52 ± 0.321 mm in the Coquet and 1.94 ± 1.26 mm in the Eden. Although the API_5 was similar for both catchments, there is higher variability in the Eden, primarily due to high API_5 during the October sampling trip (Appendix Table A-9).

3.3.2 Coquet and Eden catchment microbial communities - quantification and diversity

Estimated bacterial cell concentrations in the Coquet (based on 16S rRNA data) varied from $6.37 \times 10^3 \pm 1.37 \times 10^3$ to $1.87 \times 10^5 \pm 7.97 \times 10^4$ cells/mL and in the Eden from $1.6 \times 10^3 \pm 7.48 \times 10^3$ to $8.79 \times 10^4 \pm 3.17 \times 10^4$ cells/mL (Appendix Table A-10). The lowest cell abundances were recorded in the upstream sites (Site A) in both the Coquet and the Eden, whilst highest abundances were recorded at Site J in the Coquet downstream of Felton and Site J in the Eden within Carlisle city (Figure 3-1). Overall, when broadly comparing the sampled sites in the catchments, there were no significant differences in cell count numbers between the Coquet and Eden based on 16S rRNA data (Kruskal Wallis test p=0.55). There were also no significant differences in cell counts when comparing sampling months (Kruskal Wallis Coquet: p=0.85, Eden: p=0.95) (Appendix Table A-11), however in general lower cell counts were observed in colder months (March/November).

Microbiomes were assessed by 16S rRNA sequencing with Illumina MiSeq. Following QMP normalisation, there were 2975 taxa amongst samples. Sampling depth was highest in the Eden samples, especially in colder months such as November and March (Appendix Figure A-1). Conversely, lower sampling depth is seen in the October and September Eden samples, as well as March/November samples at downstream sampling points in the catchment. Lower sampling depth in general was observed in the Coquet samples.

The microbial abundance was plotted for each site in both catchments (Figure 3-2), showing the top 25 ASVs in each. Both catchments showed low levels of abundance

in upstream sites and an increase in the midstream. The Coquet shows a particularly high abundance at site J, such as a high abundance of *Flavobacteriaceae* in September. Outlier analysis revealed the levels were substantially higher than the typical ASV abundance across all sites and these data were removed from further analysis (Appendix Figure A-2).

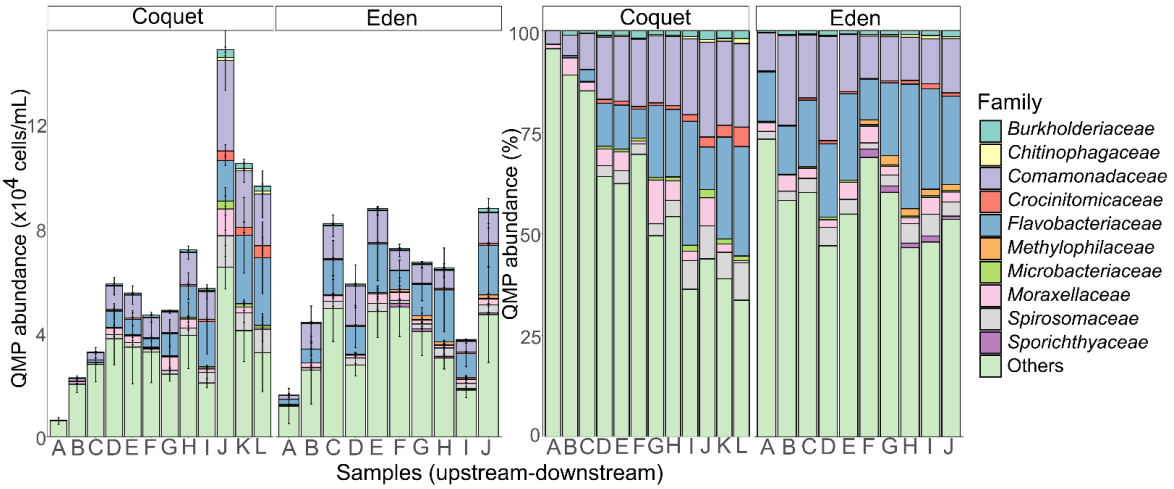


Figure 3-2 Bar plots showing Quantitative Microbial Profiling (QMP) abundance of the 25 most abundant amplicon sequencing variants (ASVs) grouped into Families, with the remaining grouped into 'Others'. The error bars show the deviation between sampling months.

Analysis of microbial abundances using the QMP approach permits the analysis of significantly more abundant taxa between catchments to enable comparisons. Taxa that had significantly different abundance across all samples in both catchments at Phylum level are presented in Figure 3-3. Overall, eight taxa were significantly different between catchments (Wilcoxon test, p<0.01).

In the Coquet, *Bdellovirbionota* and *Pastescibacteria* are significantly more abundant compared to the Eden (Wilcoxon test, p<0.01). In the Eden, *Gemmatimonadota* and *Synergistota* are more significantly abundant compared to the Coquet (Wilcoxon test p<0.01).

The order of diversity (q) was plotted against microbial diversity represented through Hill numbers (Appendix Figure A-3) Decreasing richness of species was evident from up to downstream in the Coquet and decreasing evenness of species. The same pattern was seen in the Eden, except in September when communities were less rich and more uneven along the river.

Beta diversities in each catchment were visualised using a Sørenson-type overlap

dissimilarity measure at q=1 (Appendix Figure A-4). There were differences in microbial community structure in the upstream and downstream sites in both catchments, except for some upstream sites in the Eden, which sometimes clustered with downstream sites. Microbial community diversities in both rivers varied across sampling months, which was particularly apparent in the Eden catchment.

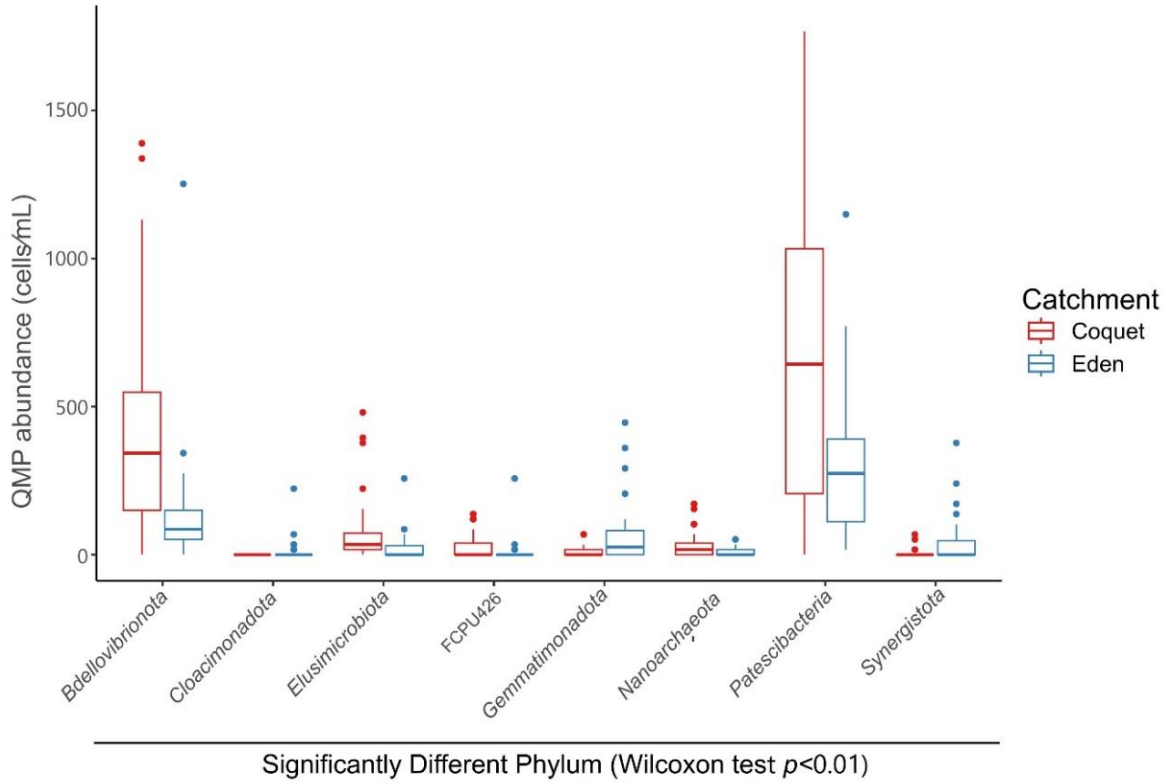


Figure 3-3 Significantly different abundant Phylum in the Coquet and Eden catchments. Abundance was assessed through quantitative microbial profiling (QMP) and significant differences were determined through a Wilcoxon test, where the significance threshold was set at p<0.01.

3.3.3 Comparing the resistomes and mobilomes in the Coquet and Eden catchments.

Relative and absolute abundance of ARGs sorted by antibiotic group are shown in Figure 3-4 and Appendix Table A-15. Outlier analysis indicated that Coquet Site D in September had particularly high relative and absolute abundance of ARGs and MGEs compared to sites in the same catchment, and in the Eden, Site A had particularly high relative abundance of ARGs and MGEs (Appendix Figure A-5). As Site D in the Coquet was based on a single sample and it was a statistical outlier, it was removed from further analysis, although data from the site are included in

Appendix Table A-15. For Eden Site A in March, outlier analysis indicated the sample data were in the normal range for the catchment once data were converted to absolute abundances (i.e., ARG and MGE copies/mL). Therefore, Site A in the Eden in March was included in further analysis.

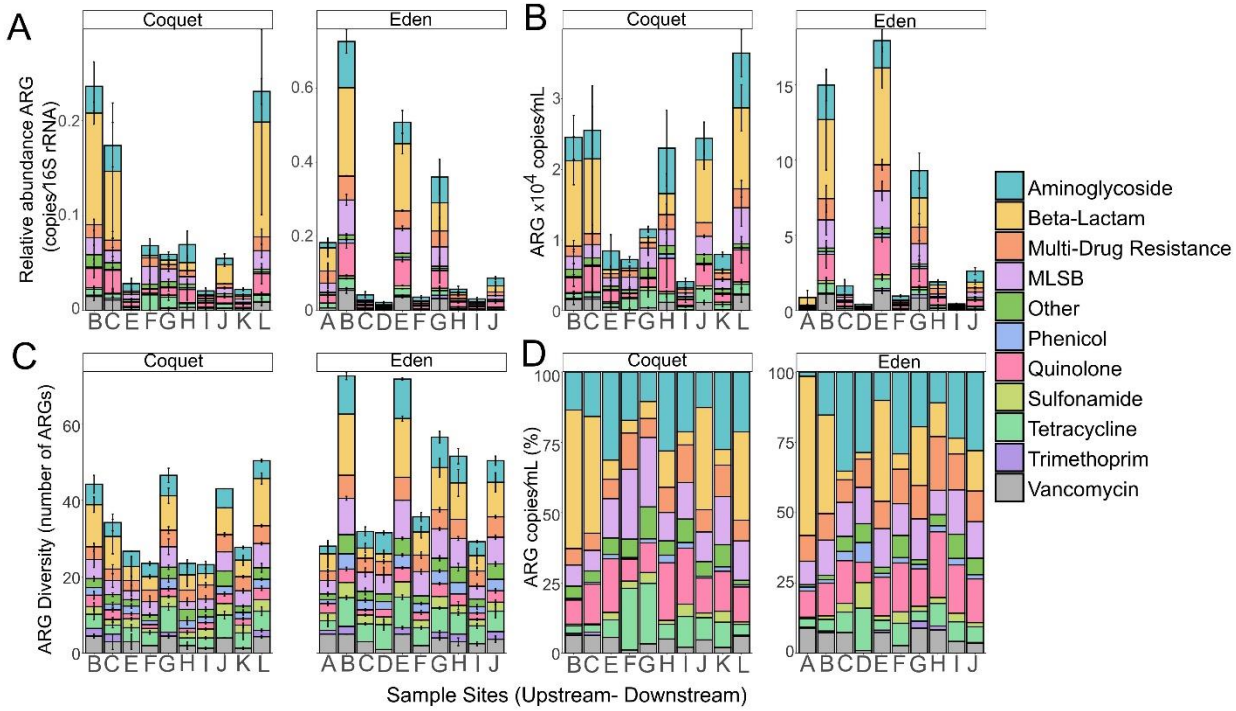


Figure 3-4 A) Relative abundance of ARGs/16S rRNA grouped by antibiotic class in the Coquet and Eden B) ARG copies/mL for the Coquet and Eden, c) Diversity of ARGs, D) percentage of absolute abundance (ARG copies/mL) in each catchment.

The Coquet had an average total relative abundance of 0.09 ± 0.13 ARG copies/16S rRNA (mean ± standard deviation) and absolute abundance of 1.6 $x10^4$ ± 1.7 $x10^4$ ARG copies/mL. The average relative abundance in the Eden sites was 0.2 ± 0.31 copies/16S rRNA and the absolute abundance was 5.2 $x10^4$ ± 7.83 $x10^4$ copies/mL. The Eden had on average higher relative and absolute abundance of ARGs compared to the Coquet, but this was not significant (Kruskal Wallis, p>0.05). However, the Eden catchment had significantly higher ARG diversity, based on the 96 gene assay (Kruskal Wallis, p=0.0207), with 45.6 ± 20 genes detected out of 85 potential ARGs, compared to the Coquet's average of 32.5 ± 16.3 genes (Appendix Table A-16).

In both catchments, MGEs (including integrons; see Appendix Table A-16) were detected at all sites (Figure 3-5), with an average relative abundance of 0.21 ± 0.22

MGE copies/16S rRNA in the Coquet and 0.34 ± 0.69 MGE copies/16S rRNA in the Eden catchment. In the Coquet, the average absolute abundance of MGEs was 3.98 $x10^4$ ± 3.6 $x10^4$ MGE copies/mL and 5.8 $x10^4$ ± 6.4 $x10^4$ MGE copies/mL in the Eden. The MGE abundance was not significantly different between the catchments (Kruskal Wallis, p=0.46) (Appendix Table A-16).

The extent of shared ARGs and MGEs in the Coquet and Eden catchments is provided in Appendix Figure A-6. Overall, within the 96 ARG/MGE gene assay, the catchments had similar resistomes, where the Coquet had no unique genes and the Eden had four unique genes, two conferring resistances to Beta-Lactams (*bla*$_{KPC}$ and *bla*$_{GES}$), one tetracycline gene (*tetPB_1*) and one MLSB gene (*lsaC*). In terms of abundance, the *bla*$_{KPC}$ gene and *tetM* gene were found to be significantly more abundant in the Eden catchment (Appendix Figure A-7).

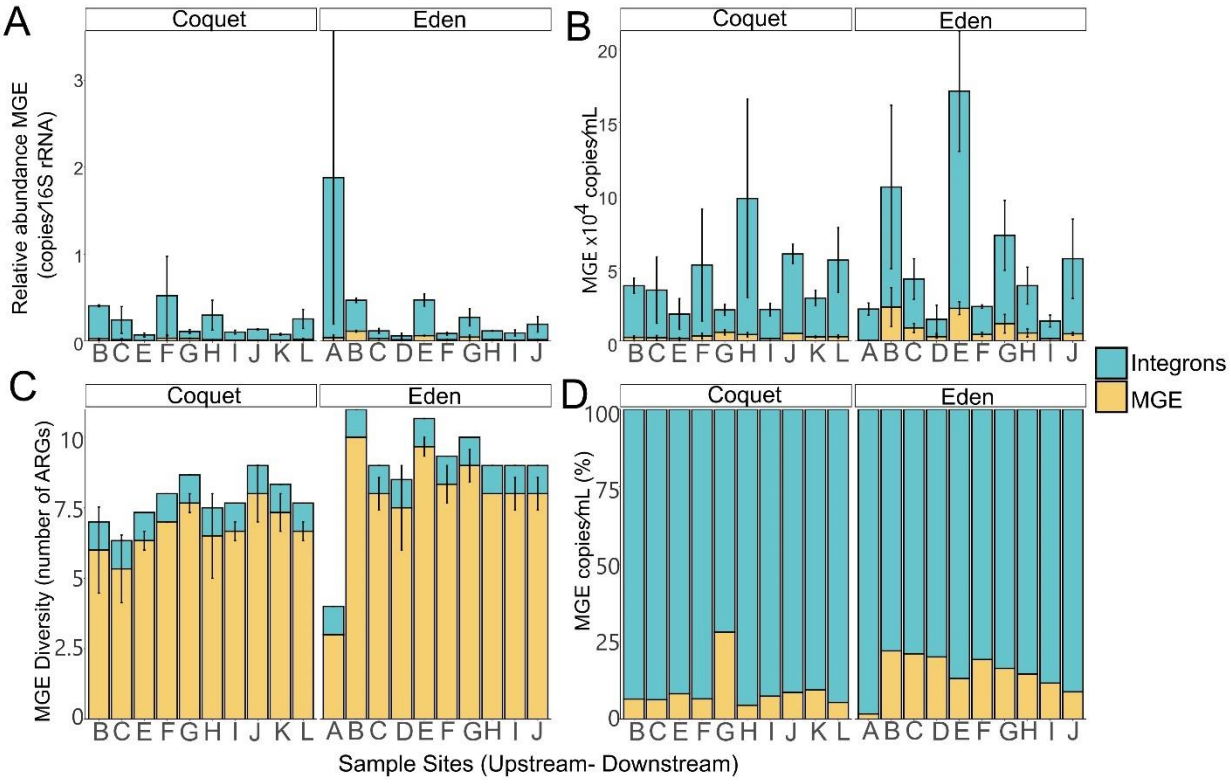


Figure 3-5 A) Relative abundance of MGEs/16S rRNA B) MGE/copies/mL C) Diversity of MGEs, D) Percentage of Absolute abundance (MGE copies/mL) in each catchment.

3.3.4 Possible drivers for microbial community and resistomes in catchments

Microbial source tracking probes (MST probes) for human and ruminant faecal derived *Bacteroidetes* (HuBac and RuBac, respectively) were used to determine

potential contributing faecal sources in both catchments. Concentrations of HuBac and RuBac derived *Bacteriodetes* were significantly higher in the Eden compared to the Coquet (HuBac: Kruskal Wallis, $p=6.83_{x10^{-5}}$, RuBac: Kruskal Wallis, $p=0.00257$) (Appendix Table A-11). HuBac and RuBac concentrations in both catchments increased as one moved downstream, maximising at mid-catchment, but then decreased downstream (Appendix Figure A-8). This pattern was especially evident in the Eden, where upstream RuBac concentrations steadily increase from upstream sites to the Site F, before decreasing towards the bottom of the catchment. On average, RuBac was about 10-fold more abundant than HuBac concentrations, likely due to the high levels of agricultural land-use in both catchments (Figure 3-1).

Microbiomes and resistomes in both catchments were assessed using network analysis (Figure 3-6; Appendix Table A-17). When both catchments were analysed together, there were 22 nodes, where the highest correlating taxa was *Bacteroidales* (3 degrees). The highest correlating ARGs were bla_{TEM} (three degrees) and *cfXA*, bla_{VIM} and *tetO_2* (two degrees). The MST marker RuBac, also correlated significantly with bla_{VIM} and *sul4* (Spearman, r>0.8, $p<0.01$).

Analysed separately, the Eden and Coquet both had a similar number of nodes in their networks, with 64 and 67 nodes respectively. In the Eden catchment, the highest correlated ARG included *strB* (12 degrees), *tetQ* (10 degrees), *tetW* (8 degrees), and the highest correlating MGEs were *tnpA_5_* (5 degrees), *tnpA_1* (4 degrees) and *ISAba3* (3 degrees). The highest correlating taxa were *Bacteroidales* (7 degrees), *Acideaminococcales, Campylobacterales* and *Peptostreptococcales-Tissierellales* (6 degrees). Both HuBac and RuBac also significantly correlated with noteworthy clinically relevant ARGs in the Eden catchment (e.g., *mcr1* and *carB*, respectively) (Spearman, r>0.8, $p<0.01$).

In the Coquet, the highest correlating ARGs were *tetM* (16 degrees), *aacC4* (5 degrees), *eerm36* and *mcr1* (4 degrees). The highest correlating taxa were *Aeromonadales, Bacteriodales, Candidatus Azambacteria and Pirellulales* (all 3 degrees). Interestingly, no MGEs were strongly correlated with taxa in this catchment, although this is probably because the number of MGEs quantified were small (only 10 MGEs).

The data from the two catchments were pooled to identify additional factors driving

AMR. A correlation matrix was developed comparing parameters measured in this study, grouped into categories called water quality indicators, microbial indicators, and AMR indicators (see Figure 3-7 for specific indicators). AMR indicators, whilst not significantly correlated in most cases, are positively correlated with water quality (e.g., conductivity and pH) and the microbial indicators. For example, MGE abundance is significantly positively correlated with 16S rRNA (Spearman, p<0.05).

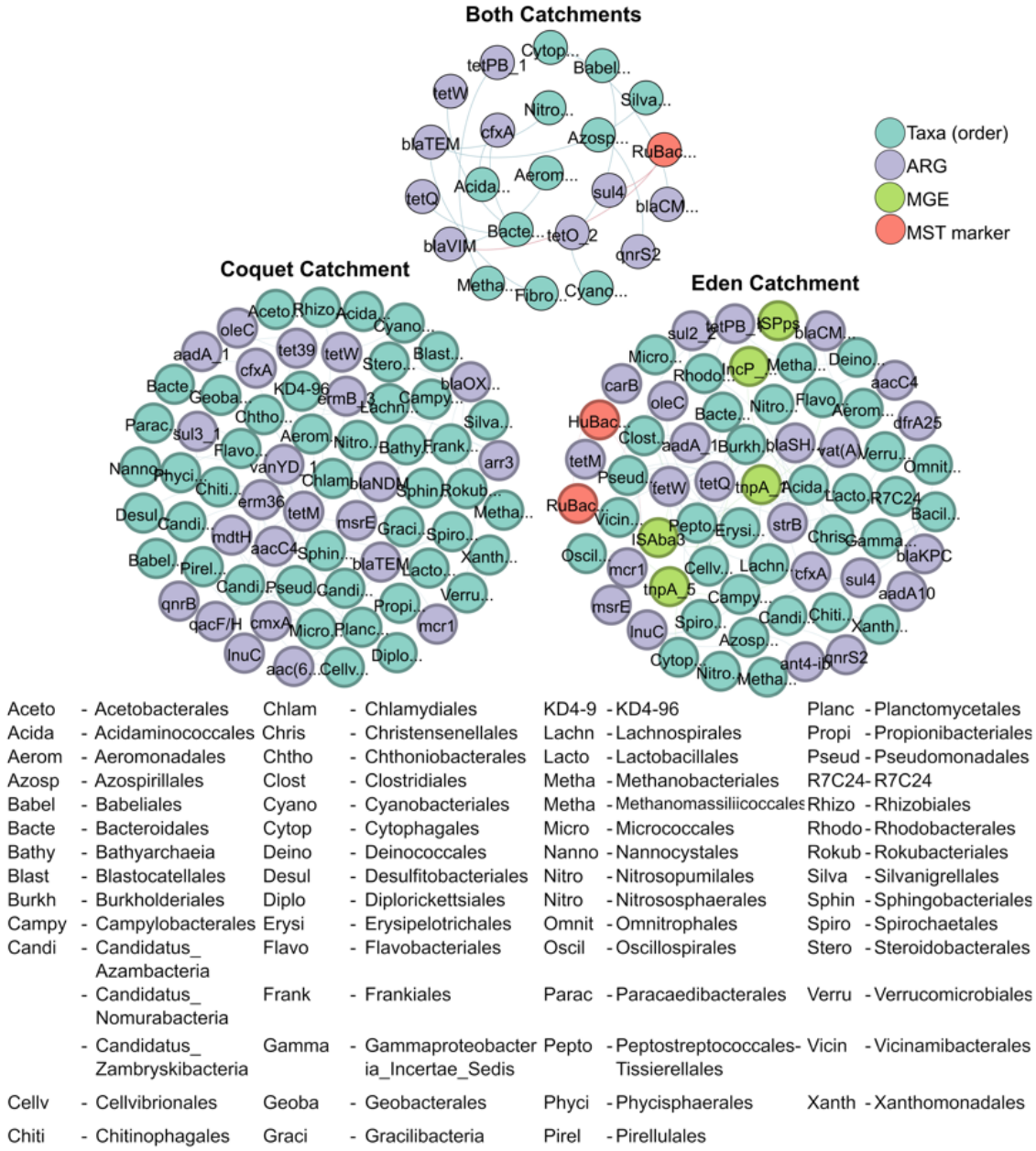


Figure 3-6 Network Analysis for both catchments, and individual catchments revealing co-occurrence patterns amongst taxa at order level (assessed through quantitative microbial profiling (QMP), ARGs/mL, MGEs/mL and MST markers (ruminant and human *Bacteroidetes* (RuBac and HuBac). A connection indicates a strong spearman correlation (rs>0.8) and significant (p<0.01), which is adjusted with a Benjamini Hochberg correction.

Interestingly, whilst microbial indicators are negatively correlated with DO (i.e., as DO declines, microbial indicators increase), AMR indicators have a weak, but positive correlation with DO. Although the hydrology indicators are not significantly correlated with water quality, microbial or AMR indicators, they do positively correlate with HuBac and RuBac, and negatively correlate with COD. The Antecedent Precipitation Index over the previous 5 days (API_5) negatively correlated with temperature, pH, DO, conductivity, COD and NH_4. There is also a negative correlation with the flow rate and ESBL coliforms.

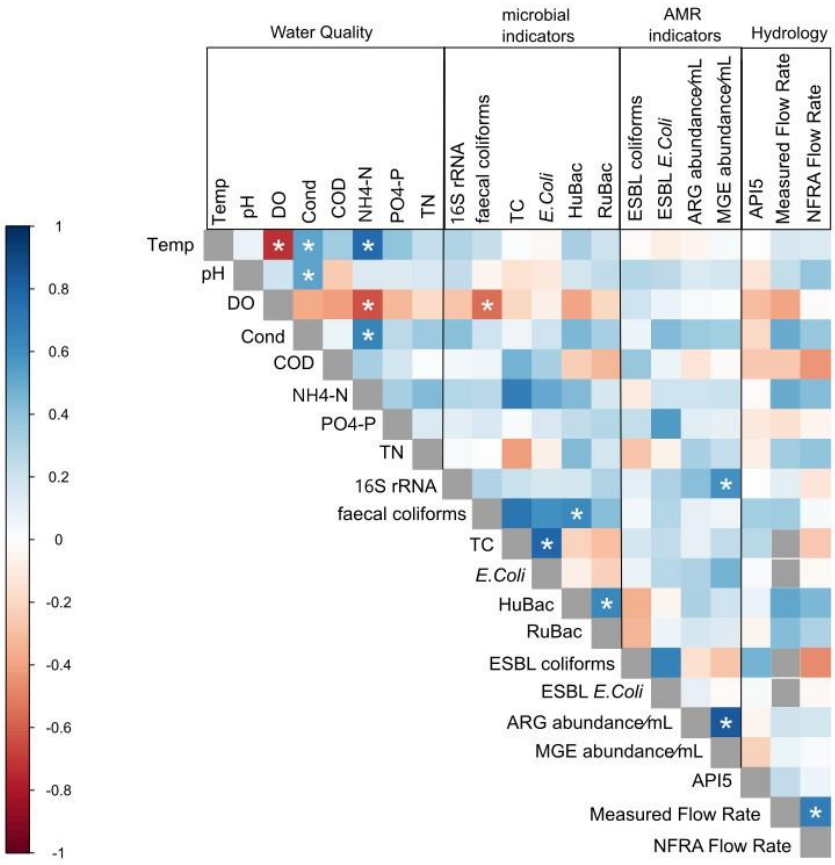


Figure 3-7 Correlation matrix for water quality parameters, microbial indicators, AMR indicators and Hydrology parameters. The Hydrology parameters include the measured flow rate (i.e., the flow rate measured for this study) and the National River Flow Archive (NRFA) Flow rate (i.e., flow rate on the day of sampling measured through river gauges) (Table 3-2). Correlations are Spearman correlations, where white stars indicate a significant correlation ($p<0.05$) with a Benjamini Hochberg correction.

3.4 Discussion

3.4.1 Impact of hydrological factors on microbial communities in the Coquet and Eden catchments

The hydrology of both catchments was characterised, and the Eden generally had more extreme differences in flow rates, rainfall and runoff compared to the Coquet catchment. Due to the connections of the Eden River with reservoirs such as Haweswater, Wet Sleddale and Ullswater, more extreme precipitation will impact river flows, especially in the middle of the catchment (NRFA, 2023). More extreme flows will clearly influence flow rates but will also 'flush' upstream and on-land contaminants downstream, impacting river water quality lower in the catchment as seen previously (Chung et al. 2008; Zhang et al. 2014). This effect is most evident in the Eden through the human and ruminant *Bacteroidetes* levels, which increase towards the middle of the catchment before decreasing in concentrations farther downstream. Decreases farther downstream are probably due to greater urbanisation and land-management in those reaches, where there are interventions such as fencing or riparian buffers that can protect stream water quality from cattle-related runoff and pollution (Grudzinski et al. 2020).

In general, the Coquet has higher numbers of the *Bdellovibrionota* and *Patescibacteria* phylum, which are often prevalent in less polluted surface water or groundwater environments (Brown et al. 2015; Herrmann et al. 2019; Im et al. 2019; Chaudhari et al. 2021; Li et al. 2021). In particular, the *Patescibacteria* phylum is often present in nutrient limited conditions (Tian et al. 2020). *Bdellovibrionota*, is a phylum that preys on other bacteria (Sockett and Lambert 2004). This has been previously associated with low abundance of microalgae (Yang et al. 2023), and maintenance of a healthy and diverse ecosystem through removing dominant bacterial groups (Zhang et al. 2023b). Higher numbers of such strains are consistent with nutritional conditions and evidence of low pollutant inputs in the Coquet, where the Coquet River microbial communities reflect less impacted conditions than the Eden, which is consistent in all the genetic and microbial data for the catchment.

Conversely, the Eden catchment had broadly greater abundances of microbes associated with soils, limnic environments and sediments, such as the phylum *Gemmatimonadiota* (Mujakić et al. 2022) and *Synergistota*, which is often present under conditions impacted by animal faeces, surface soils, and wastewater releases (Bhandari and Gupta 2012). Longer periods of high soil wetness (PROPWET) and greater soil saturation in the Eden catchment appears to result in greater runoff that contains phylum commonly present in the soils entering the river. This speculation is

supported by significantly higher ruminant *Bacteroidetes* abundances in the Eden compared to the Coquet catchment.

3.4.2 Drivers of AMR in the Eden and Coquet catchments

The river water resistomes in the two catchments were similar, although the diversity, and absolute and relative abundances of ARGs and MGEs were slightly higher in the Eden compared to the Coquet catchment, with the diversity of ARGs in the Eden being significantly higher.

Comparison of the catchments revealed that from the 96 gene assay used in the study, the Eden catchment had four unique genes compared to the Coquet. This included the beta-lactam genes, *bla*$_{KPC}$ and *bla*$_{GES}$ that encode resistance to carbapenem antimicrobials, which were significantly more abundant in the Eden catchment compared to the Coquet. Both these are plasmid-mediated genes that are genetically mobile and can be shared through horizontal gene transfer (HGT) (Queenan and Bush 2007; Bennett 2008; Mengistu et al. 2022). The high prevalence of *tetM* was previously found to be a consequence of environmental pollution caused by livestock (Munck et al. 2015).

Carbapenemase genes have been more frequently studied in the context of agricultural and wastewater related contamination of natural waterbodies (Mills and Lee 2019). In particular, *bla*$_{KPC}$ represents carbapenem resistance in *Enterobacterales*, which is a problem in hospital settings, especially in the NW England (Stoesser et al. 2020). In this study, the *bla*$_{KPC}$ gene was detected once in Site G, and twice in two sites, Site B and E. All sites have a high percentage of agricultural land-use, but also have small settlements nearby. Therefore, the presence of *bla*$_{KPC}$ could be a result of community wastewater, septic tanks, or agricultural contamination. Interestingly, *bla*$_{KPC}$ was not found in the Coquet, which is in NE England. More sampling is needed to determine the source of this gene, especially what it might suggest relative to the spread of hospital-associated AMR in the environment through waste releases. This should also utilise more targeted qPCR approaches to confirm detections.

Our network analysis revealed that the ARGs in the Coquet and Eden have multiple potential hosts. In both catchments *Bacteroidales* was highly correlated with ARGs, consistent with previous studies finding that *Bacteroidales* often carry abundant

ARGs including tetracycline and beta-lactam genes (Li et al. 2022c). *Aeromonadales*, known to harbour clinically relevant drug resistance (Kneis et al. 2022), also significantly correlated with ARGs. The network analysis further highlighted differences in microbial, resistome and mobileome interactions between the catchments. There were multiple MGEs strongly correlating with ARGs in the Eden, whereas no strong correlations were seen between ARGs and MGEs in the Coquet. This further indicates greater anthropogenic impact in the Eden, where the presence of MGEs is indicative of acquired resistance (Datta and Hughes 1983) and greater human and animal waste inputs to the river.

In addition, human and ruminant *Bacteroidetes* had strong and significant correlations with ARGs in the Eden, whilst there were no correlations with these markers in the Coquet, indicating the increased likelihood of human and/or agricultural related resistance in the Eden catchment compared to the Coquet. However due to the complexity of the microbial interactions in the environment, this is speculation at best, although it is broadly consistent with qualitative differences between the two catchments and might be useful for considering the effects of differences between river catchments in general terms (Carr et al. 2019).

To understand ecological drivers, correlation matrices were developed to investigate the links between water quality, microbial indicators, resistance indicators and hydrology. Interestingly, unlike previous studies in more contaminated catchments (Ho et al. 2021; Ott et al. 2021a), there were few strong statistically significant positive correlations. There is a positive correlation between AMR indicators, such as ESBL *E. coli* abundance and conductivity (a good indicator for dissolved solids in a river; Abdulsattar et al. 2020). Conductivity may thus be linked to inorganic and organic pollution from fertilisers, or runoff from roads (particularly as most sites were sampled close to roads), which may contain heavy metals that can increase resistance through co-selection (Knapp et al. 2017; Robins et al. 2022). However, this cannot be verified with available data.

Unlike previous studies in heavily contaminated catchments, DO was only a weak indicator for AMR, whereas previous studies have found a strong, significant negative correlation (Ho et al. 2021; Ott et al. 2021a). This could be due to the difference in AMR sources, where resistance may be primarily derived from diffuse

agricultural sources (here) as opposed to point source contamination from wastewater (Ho et al. 2021; Ott et al. 2021a). This observation highlights the importance of studying less contaminated sites to understand the drivers for AMR in “background” environments. In a heavily contaminated site, DO may be an effective indicator for resistance and be used as a marker for AMR (Ott et al. 2021a). However, in less contaminated sites, which are primarily influenced by diffuse rural activity, alternative markers for AMR need to be considered for monitoring.

3.4.3 The need for further environmental surveillance

The comparison of the Coquet and the Eden catchments illustrates the importance of increased integrated AMR surveillance and insights it can bring relative to different catchment dynamics and AMR. The quantitative approach used here lends further support to recent recommendations by the UN Environment Programme (UNEP, 2023). Moreover, this study has indicated that the local geography of the Eden catchment, elevated river flows, rainfall and runoff are associated with greater agricultural contamination and increased resistance, a connection that would not have been clear without parallel data from the Coquet catchment.

This study exemplifies the importance of integrated sampling and analysis and method standardisation, which allowed us to better compare and understand the drivers of resistance in “typical” UK rivers. With higher flow rates apparently increasing rural in situ resistance, the expected more dynamic rainfall events due to climate change (Watts et al. 2015) may increase resistance in rivers due to surface runoff. Whilst rivers in contaminated environments in the UK have been well studied, understanding of regional catchments like the work here will provide insights into the temporal and spatial variation of AMR.

Increased spatial surveillance could further inform environmental risk assessments through understanding of how catchments differ in terms of AMR relative to each other (Burch et al. 2022). Furthermore, this study demonstrates the need for monitoring in different environments, where ultimately the data can be used to inform large scale routine monitoring for AMR in environments (Bengtsson-Palme et al. 2023; Hart et al. 2023). However, there is a need for metadata as well as AMR data, such as water quality and nutrient conditions, to explain drivers of AMR within environmental studies.

3.5 Conclusions

This study aimed to compare the microbiome and resistome of two rural catchments in the United Kingdom. Through comparison of the Coquet and Eden catchments, this study demonstrated that the geography and hydrology of the Eden catchment led to slightly elevated resistance across sites, which may not have been clear through isolated catchment studies. Standardised methodologies for microbiome and resistome measurement are necessary to enable comparison between studies and catchments. In addition, increased surveillance for AMR in non-contaminated areas, as well as contaminated areas are important to understand the different drivers associated with environmental AMR. However, the study overall shows that even rural catchments can have noteworthy levels of resistance potential, but in such locations, the sources are more nebulous and may be dominated by non-point source runoff. Further work is needed on similar rural catchments to corroborate and more generalise the results herein.

Chapter 4 Defining the sources of AMR in a rural catchment: A case study on the Coquet River (Northumberland)

4.1 Introduction

The natural and developed environment have been increasingly recognised as a source, driver and sink for antimicrobial resistance (AMR) (Graham et al. 2019) and the development and spread of AMR in the environment is now established as a public health concern (Ahmad et al. 2021). However, identifying the dominant pathways of spread that lead to exposures is often difficult, which is particularly true in settings that have a wide range of sources and sinks, where none of which are blatantly obvious.

The sources of environmental AMR are complex due to a variety of point and diffuse sources in environmental settings, and the coalescence of intrinsic and acquired antibiotic resistance genes (ARGs) in a continually changing environment (Li et al. 2018). Knowledge of the sources of AMR in the environment may therefore improve the capacity to initiate appropriate mitigation strategies (Chen et al. 2023).

Existing methods of detecting AMR in river environments and identifying AMR hotspots include monitoring the occurrence of antibiotic resistant bacterial phenotypes, such as extended spectrum beta-lactamase (ESBL) producing *Escherichia* coli (*E. coli*) (Anjum et al. 2021). Whilst this is an effective, easy to implement and inexpensive method of monitoring, in mixed environments, such as rivers, it is difficult to distinguish sources of resistance. ESBL *E. coli* phenotypes “look” functionally the same regardless of their source. Additional methods for source tracking are needed, including culture independent, genetic microbial source tracking (MST) methods. These include the UK Environment Agency ruminant and human *Bacteroidetes* markers previously discussed in Chapter 3, which are markers that are broadly associated with the gut of original hosts (Environment Agency 2008). MST markers including *Cattellicoccus marimammalium* and human associated *Bacteriodetes* and *Lachnospiraceae* not only identified signals faster than culture dependent methodologies, but also detected sources of faecal pollution after sewer overflow events, which was not possible through *E. coli*-based detection (Cloutier and McLellan 2017). The disadvantages of using MST markers are they are not

geographically consistent (Mayer et al. 2018), and that there is also a risk of cross-reactivity (Boehm et al. 2013).

Community source tracking meets these limitations by using information derived from whole microbial communities detected in source samples, as well as the sinks, and approximating the percentage contribution of the source samples (McGhee et al. 2019). Methods such as the SourceTracker approach (Knights et al. 2011) have been used to detect contamination from multiple sources and account for potentially unknown sources (Mathai et al. 2020). This method uses a Bayesian algorithm based on the Markov chain Monte Carlo (MCMC) method to calculate the probability that taxa belong to specific source categories (Mathai et al. 2020). Overall, SourceTracker specificity has been found to be greater than existing qPCR MST assays for animals (Harwood et al. 2014). An assessment by Staley et al. (2018) showed that the method was 91% accurate when using libraries from local sources, highlighting the importance of using geographically specific source samples.

Despite its benefits, the approach used by SourceTracker is computationally expensive, and only fits a limited number of sources (Shenhav et al. 2019). An alternative method, called Fast Expectation Maximization for Microbial Source Tracking - termed ‘FEAST’- was recently developed based on a multinomial model distribution, reducing computational running time from days to hours (Shenhav et al. 2019). In addition, FEAST appears to be more accurate than earlier approaches. For example, SourceTracker sometimes had difficulty in distinguishing between sources with similar bacterial communities (Brown et al. 2018). Comparing SourceTracker and FEAST directly, FEAST was identified as a more suitable approach for environmental water bodies where there is low faecal input (Xu et al. 2022). Sourcetracker further has been found to underestimate the impact of unknown sources, leading to false positives (Chen et al. 2023; Wen et al. 2023).

SourceTracker and FEAST have both been frequently employed for environmental studies in aquatic environments (Zhang et al. 2022), but also in soil studies (Yang et al. 2021) and even forensic investigations (Carter et al. 2019). The FEAST model has been applied to microbiomes (Shenhav et al. 2019), but also resistomes (usually determined through metagenomic sequencing) (Chen et al. 2020) and even microplastic distributions (Li et al. 2022a). However, previous river studies have

focused on settings which have received extensive urban contamination, which are more likely to have “impacted" allochthonous microbiomes and resistomes due to dominant originating point sources (Chen et al. 2019a, 2023; Wang et al. 2020; Li et al. 2022a; Zhang et al. 2022). There is a recognised lack of studies in environments which are less exposed to anthropogenic contamination (Hooban et al. 2020). One such environment may be the rural UK, where rivers receive inputs from small wastewater treatment works and agricultural runoff, which both have the potential to enhance the natural resistome but are often relatively small or are diffuse sources in themselves.

Following inputs of diffuse and local point sources into natural river systems, the factors which dictate the survival of non-native bacteria and ARGs are not well understood (Mahaney and Franklin 2022). Previous studies in microcosms have shown that following the release of untreated waste into river systems, populations rapidly die off (Mahaney and Franklin 2022). Factors such as nutrient availability may be important in influencing the microbial composition, where the microbial community in river systems are more capable of surviving the nutrient limited conditions compared to non-native bacteria (Hibbing et al. 2010). The impact of nutrient availability can be determined through the stoichiometric nitrogen- phosphorous (N:P) ratio, where an unbalanced ratio (i.e., a phosphorous or nitrogen limited ratio relative to each other or relative to carbon supply) may impact ecological functions (Ibekwe et al. 2016).

The aim of this Chapter is to (1) identify the primary contributing sources of the microbial community and AMR in the rural Coquet catchment (Northumberland, UK) using the FEAST approach and (2) determine the environmental factors influencing the persistence of non-native ARGs in this river system, including the role of nutrient limitation as a defining factor in community composition (Tilman 1982). Here different rural sites were sampled from upstream to downstream of the river catchment to test the hypothesis that as a river receives sources, such as a wastewater effluent point sources, the local limiting-nutrient conditions change, which alters the microbial community and AMR that is further carried down the river. Specifically, N:P ratio

conditions were used to investigate whether nutrient availability was a driving factor for the survival of non-native bacteria.

4.2 Materials and Methods

4.2.1 Site description

The sampling campaign was designed to capture the influence of four types of “land-use:” (i.e., input sources to the river), pseudo-pristine land, a rural area with a small wastewater treatment plant, and two small towns with medium size wastewater treatment plants. To capture the different land-uses, the samples were collected in ‘Clusters’, where a cluster comprised of a distinct land-use and captured upstream and downstream river and source sites. Clusters and associated sample sites are described in Table 4-1 and shown in Figure 4-1. Each cluster (except Site A, which was used in additional analysis to account for ‘background’ resistance) included samples from a specific wastewater effluent source, riverbank soil samples, upstream river samples as ‘source’ sites, and a downstream river water site as ‘the sink’.

Sample collection was between March 2022 and August 2022 (Appendix Table B-2). Clusters A and B were sampled on separate days to Clusters C and D to allow time for lab processing. Clusters were all sampled in the same week, except for the second sampling campaign, which had a short delay due to a positive COVID-19 test. An extra sample was collected directly downstream of the wastewater effluent in Cluster B (i.e., B3); however, it was not possible to collect from this site in other sampling campaigns due to problems with site access. Site access issues also prohibited collection of summer soils samples at Site D (i.e., summer D2_SOIL) as well as associated wastewater samples (i.e., DS2) in the final sampling campaign.

4.2.2 Catchment sampling

River water from the midstream was collected as grab samples by suspending a sampler by a rope from a bridge or using a telescopic sampler, if sampling had to be performed from the riverbank. Wastewater effluent samples were collected directly from the wastewater treatment plant (WWTP) site. A riverbank soil sample was obtained 100 m and 200 m upstream of the river samples where possible, using a 50 cm centrifuge tube to obtain a core of ~12 cm below the root (Baral et al. 2018a).

Table 4-1: Sampling Clusters, the sample sites within the clusters and the description of the cluster. All river water samples are sampled along the main Coquet River unless otherwise specified. Sample site location coordinates can be found in Appendix Table B-1

Cluster	Sampled sites	Description
A	**A** (river water) **A_SOIL** (100m & 200m) (soil samples)	*Purpose for sampling* • Control site at the top of the Coquet catchment. *Land use/point source input* • None *Samples collected* • Samples were collected in the river and the riverbank soil.
B	**B1** (river water) **B2** (river water: tributary) **B3** (river water: tributary) **B4** (river water) **BS1** (wastewater effluent) **B1_SOIL** (100m & 200m) (soil samples) **B4_SOIL** (100m & 200m) (soil samples)	*Purpose for sampling* • Capture a predominantly rural location including sheep and cow farms, with a wastewater treatment plant in the upstream tributary. *Land use/point source input* • Wastewater treatment effluent in an upstream tributary (contributing population ~882). *Samples collected* • Samples were collected from wastewater effluent in the tributary, downstream of the tributary, and along the main Coquet River (upstream and downstream of the tributary). Soil samples were also taken upstream of the main river sites.
C	**C1** (river water) **C2** (river water) **CS1** (wastewater effluent) **C2_SOIL** (100m & 200m) (soil samples)	*Purpose for sampling* • Capture downstream river site from a small town *Land use/point source input* • Wastewater effluent (contributing population ~2107). *Samples collected* • Samples were collected from the wastewater effluent, upstream of the wastewater effluent discharge, and immediately downstream. • Soil samples were also taken upstream of the downstream site.
D	**D1** (river water) **D2** (river water) **DS2** (wastewater effluent) **D2_SOIL** (100m & 200m) (soil samples)	*Purpose of sampling* • Capture downstream river site from a small town further downstream the Coquet *Land use/point source input* • Wastewater effluent (contributing population ~1090). *Samples collected* • Samples were collected from the wastewater effluent, upstream of the wastewater effluent discharge, and immediately downstream. • Soil samples were also taken upstream of the downstream site.

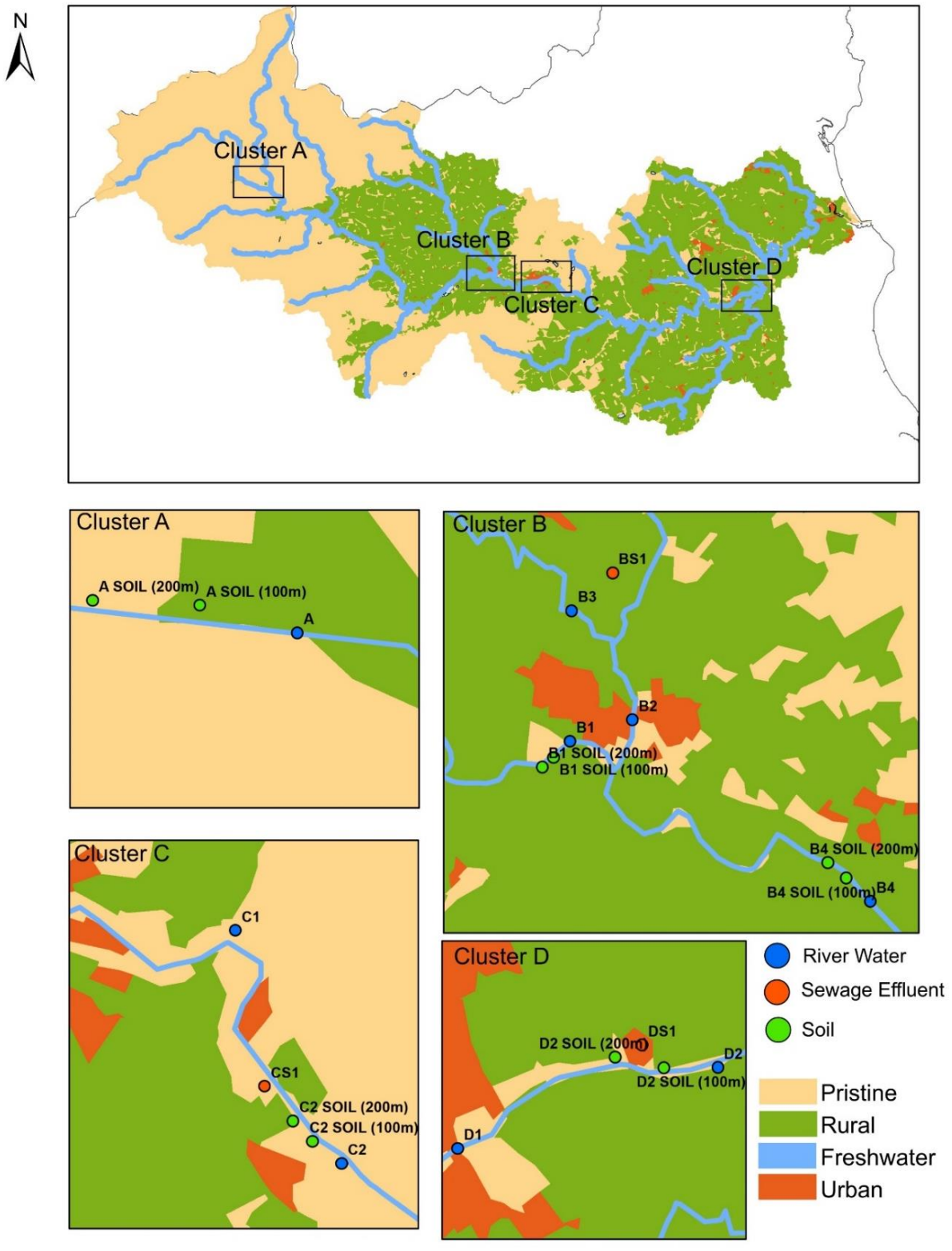


Figure 4-1 Map of Sampling sites and location of different sampling sites and land-use

For river water and wastewater effluent samples, temperature, dissolved oxygen (DO), pH and conductivity were measured on-site as previously described (Chapter

3.2.1). River volumetric flowrate was approximated using the float method as described previously (Michaud and Wierenga 2005). The cross-sectional area of the river was estimated using measured river widths and depths which were multiplied by the surface velocity. A correction factor of 0.85 was applied to the surface velocity data (Michaud and Wierenga 2005; Ott et al. 2021a). In cases where the sample was collected at the riverbank, the width could sometimes not be measured, the velocity was recorded.

Microbial plating and colony counting of *E. coli* and ESBL *E. coli* for wastewater effluent and river water samples was conducted within 24 hours of sampling as previously described in Chapter 3.2.6. The amount of river water filtered for plating ranged from 10 to 20 mL for non-ESBL plates and 100 to 300 mL for ESBL plates, whilst for wastewater effluent, the filtered volume was 50 µL for non-ESBL plates and 2 mL for ESBL plates. Physiochemical analysis for ammonium nitrogen (NH_4-N), orthophosphate as phosphorous (PO_4-P), chemical oxygen demand (COD) and Total Nitrogen (TN) was conducted within 48 hours of sampling, as described in Chapter 3.2.6.

4.2.3 DNA extraction and quantification of 16SrRNA, ARGs, MGEs and MST probes

River water and wastewater effluent were filtered as described in Chapter 3.2.2. DNA was extracted as described in Chapter 3.2.2 using FASTDNA Spin Kit for soil (MP Biomedicals, UK), with the exception that 1 litre of composite wastewater effluent sample was used for DNA extraction, compared to 3 litres composite river water sample. Riverbank soil samples were pooled into a composite sample, where a total of 1.5 g of soil (3 x 0.5 g) was used for DNA extraction. DNA concentration and purity were assessed as described in Chapter 3.2.2 and diluted to 5 ng/µL for downstream analysis to prevent inhibition.

Human and ruminant *Bacteroidetes* MST probes (HuBac and RuBac), faecal coliforms and 16S rRNA (total bacteria) were quantified using quantitative polymerase chain reaction (qPCR) as previously described in Chapter 3.2.2.

4.2.4 Next-generation sequencing

The Illumina MiSeq platform at NU-OMICs, Northumbria University, UK was used to sequence the hypervariable V4 region 515F-806R of the 16S rRNA gene with V2 500 cycle chemistry using the Schloss MiSeq Wet Lab SOP (Kozich et al. 2013), as previously described in Chapter 3.2.3.

Raw sequencing data were available as FASTQ files and were processed and denoised into Amplicon Sequence Variants (ASVs) with DADA2, using R programming (Callahan et al. 2016). Naïve Bayes classifiers were used for taxonomic assignment to genus level. ASVs of <0.1% of the mean sample depth were removed to account for MiSeq bleed (Comeau et al. 2017). The taxonomy and ASV table *biom file* were produced for downstream analysis in R Studio (v. 4.2.2), where phylum level data with mean value <1 and maximum value <10 were removed. This resulted in a total of 4431 taxa for 77 samples, with minimum reads of 4280, a median of 2612 and a maximum of 119457 reads.

Quantitative Microbial Profiling (QMP) analysis was conducted to rarefy ASVs to sampling depth and provide quantitative data of the present taxa as previously described (Chapter 3.2.4) (Vandeputte et al. 2017; Ott et al. 2021b). This was completed independently for river water, wastewater effluent and soil samples.

4.2.5 Relative abundance of ARGs and MGEs using High-throughput qPCR

The relative abundance of 96 ARGs and mobile genetic elements (MGEs) was characterised using high-throughput qPCR (HT-qPCR) performed by Resistomap Oy (Helsinki, Finland) as described in Chapter 3.2.5. Relative ARG and MGE data (i.e., gene copies normalised to 16S rRNA copy numbers) provided by Resistomap were transformed into absolute copy numbers by multiplying with the 16S rRNA abundance/mL for each sample that were independently determined using qPCR.

4.2.6 Microbial Source Tracking with FEAST

Community-based microbial source tracking analysis was performed using FEAST in R Studio using the package *FEAST*, with the maximum number of iterations set to 1,000 (Shenhav et al. 2019). The FEAST source/sink model for this study is described in Figure 4-2. FEAST analysis was only completed for Clusters B-D in the Coquet catchment because Cluster A had too few point source samples. The FEAST

model was run for each sampling trip and repeated five times to reduce false positives as previously described (Chen et al. 2023). Microbial abundance measured through QMP analysis (i.e., microbiome data) was used to assess the percentage contribution of upstream wastewater effluent, soil samples and unknown contributions in each cluster. Separately, using the FEAST model, absolute ARG and MGE abundances (copies/mL) (See Section 2.5), were employed to determine the relative percentage contributions of source sites to resistomes in each sink.

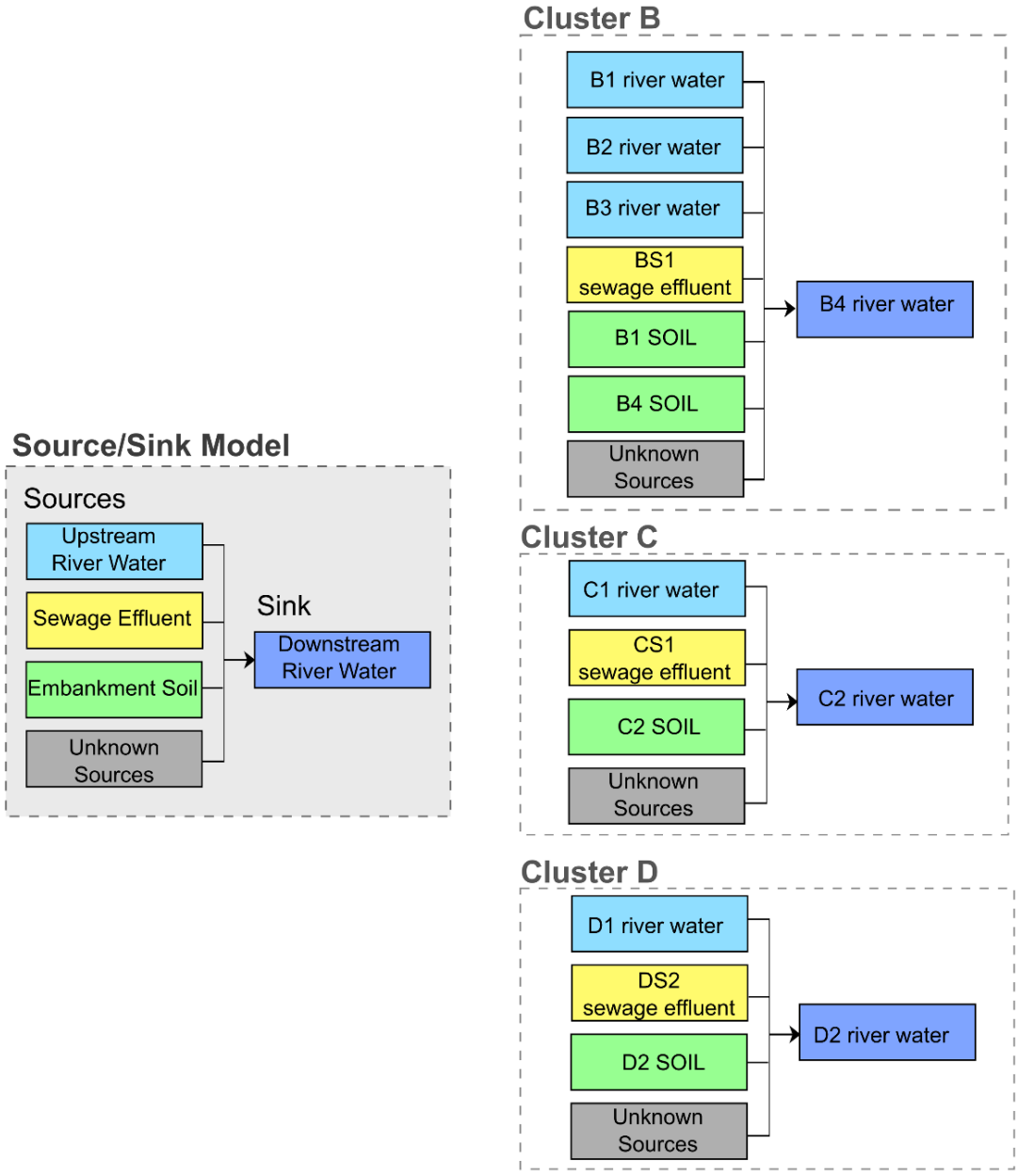


Figure 4-2 FEAST model for SourceTracker analysis (left) and specific sources and sinks used for the FEAST model in this study.

4.2.7 Statistical analysis

Statistical analysis and data processing were performed using R Studio. Graphics were developed with *ggplot2* and finalised with Inkscape (v.0.92.4). A Kruskal Wallis test was used to assess confirm the five repeated FEAST model runs were not significantly different. A paired t-test was used to assess the significant differences between the estimated FEAST percentage contribution in microbiome/resistome estimates and in summer/winter samples. The significance threshold was set as $p<0.05$ unless otherwise stated.

The impact of nutrient limitation through the catchment was explored using the stoichiometric total nitrogen (TN) and orthophosphate (PO_4^3) ratio (N:P ratio). The TN was directly measured in this study for each site, whereas PO_4^3 was converted from measured values of orthophosphate as phosphorus P (PO_4-P) (HACH 2022). The Mantel test was carried out using the package *vegan* (v2.6.4) using Spearman correlations ($p<0.05$). The $\log_2$fold change of ARG and MGE river water concentrations between the control site (i.e., background resistance), and the clusters (A vs B, C, D) were computed using the *DESeq2* package, which utilised the Wald test, where *p* values were adjusted according to the Benjamini Hochberg method (Benjamini and Hochberg 1995).

The percentage contribution of wastewater effluent flow to the river water flow in sites C2 and D2 was approximated using calculated wastewater daily flow data from Chapter 5.2.4 (Table 5-2) and the simulated average daily river flow rate (measured over 2019-2022) from the SWAT model developed in Chapter 5. The average daily river flow rate was estimated to be 5.62 $_{x10^5}$ m^3/day and 7.41 $_{x10^5}$ m^3/day for site C2 and D2 respectively.

4.3 Results

4.3.1 Microbial taxa and resistome

Microbiomes were characterised using 16S rRNA sequencing data with Illumina MiSeq. Following QMP normalisation, there were 3,245 taxa amongst river water samples, 2,289 taxa in wastewater effluent samples, and 231 taxa within soil samples. In river water samples, the most abundant ASVs were Flavobacterium (Genus level), Acinetobacter (Genus level) and Simplicispira (Genus level) (Appendix Table B-3). In wastewater effluent samples, the most abundant ASVs were

also Flavobacterium (Genus level), Comamonadaceae (Family level) and Limnohabitans (Genus level) (Appendix Table B-4), and in soil samples, the most abundant ASVs were Candidatus Udaeobacter (Genus level), Pseudarthrobacter (Genus level) and Hyphomicrobium (Genus level) (Appendix Table B-5).

Figure 4-3 shows quantitative microbial abundances (QMP) for each site in each cluster, showing the top 25 ASVs and labelled point sources in the river. For the most part, except for Cluster B, there is an increase in absolute microbial abundances downstream of point sources, and QMP abundances were generally greater in summer compared to winter.

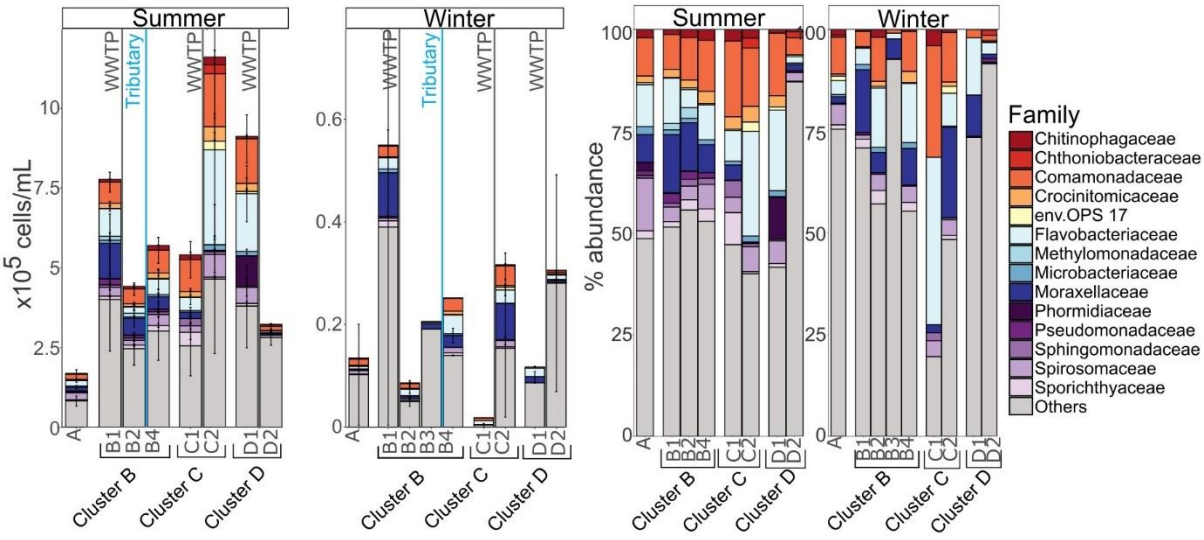


Figure 4-3 Bar plots showing the 25 most abundant ASVs grouped in families, with remaining pooled as 'Other'. Point sources including wastewater effluent (WWTP) and tributary sites are indicated in the bar plot, sites B2 and B3 are part of the tributary. The two left-hand side plots have different y-axis scales.

Despite having less unique taxa in the wastewater effluent and soil samples, the abundance of taxa per mL of wastewater and per gram of soil samples was greater than in river water samples (Appendix Figure B-1).

Figure 4-4 shows a bar plot summary of the absolute abundance of ARGs per site and Cluster, grouped by antibiotic type. Looking at Cluster C, C2, which is downstream of wastewater effluent (CS1), has the highest abundance of ARGs in summer and winter months. In winter, there are overall lower abundances of ARGs. There are higher abundances of ARGs in site B1 and site B2, whereas D1 and D2 have lower ARG abundances. Aminoglycoside, multidrug resistance (MDR, which

primarily codes for non-specific resistance mechanisms, such as efflux pumps) and quinolones are the dominant antibiotic classes of ARGs. Site C2 has a higher percentage of tetracycline ARGs than other sites.

In wastewater effluent samples, MLSB and tetracycline ARGs have a higher percentage of the total abundance (Appendix Figure B-1) and in soil samples aminoglycoside and quinolone ARGs predominate (Appendix Figure B-2).

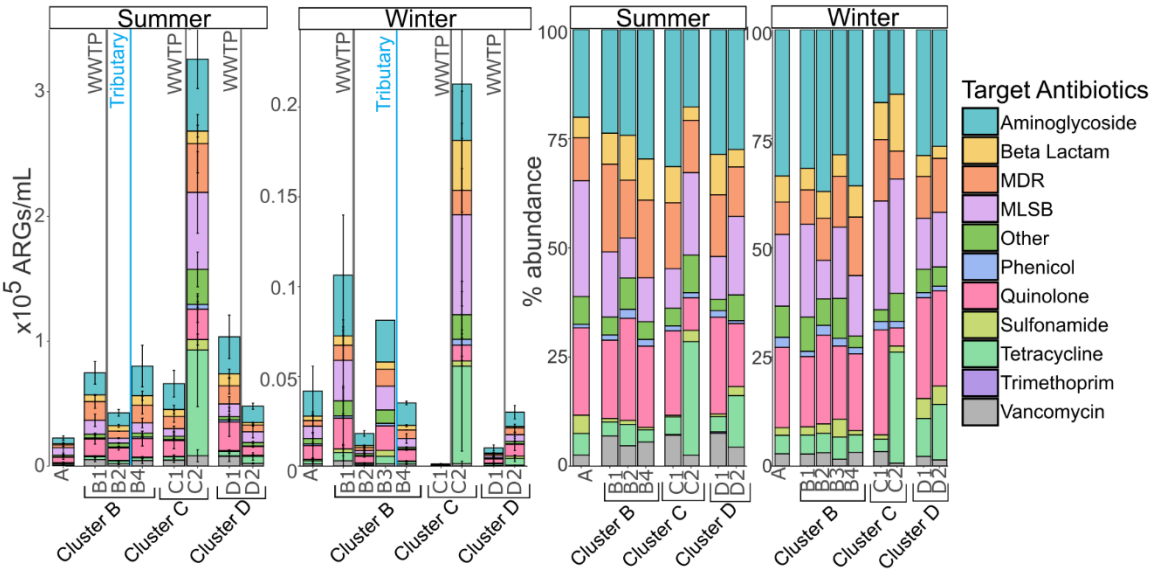


Figure 4-4 Bar plots showing the absolute ARG abundance for river water. Point sources including wastewater effluent (WWTP) and tributary sites are indicated in the bar plot, sites B2 and B3 are part of the tributary.

4.3.2 Source Contributions of AMR along the Coquet catchment

The FEAST model was used for source attribution with Clusters B to D in the Coquet catchment. Cluster A had limited point source samples and was not subject to FEAST analysis. Five independent runs of the model were completed to reduce false predictions. The results for the five model runs were functionally the same (Figure 4-5) and the predicted percentage contribution across runs for both microbiome and resistome based predictions were not significantly different (Kruskal Wallis, $p<0.05$). Therefore, the percentage contributions were averaged for model runs.

The average and standard error for contribution of sources across the sampling campaigns are shown in Figure 4-6 and summarised in Appendix Table B-6 for both estimations based on the microbial community (estimated through QMP), and the

resistome (based on HT-qPCR) data, as well as t-tests for seasonality and microbiome/resistome. Overall, there were no significant differences between source contributions in summer and winter (t-test, p<0.05). The following section will outline the estimated contributions amongst sampling Clusters, B, C and D using FEAST analysis. This section has been divided to discuss each sampling cluster separately with Cluster B being examined first.

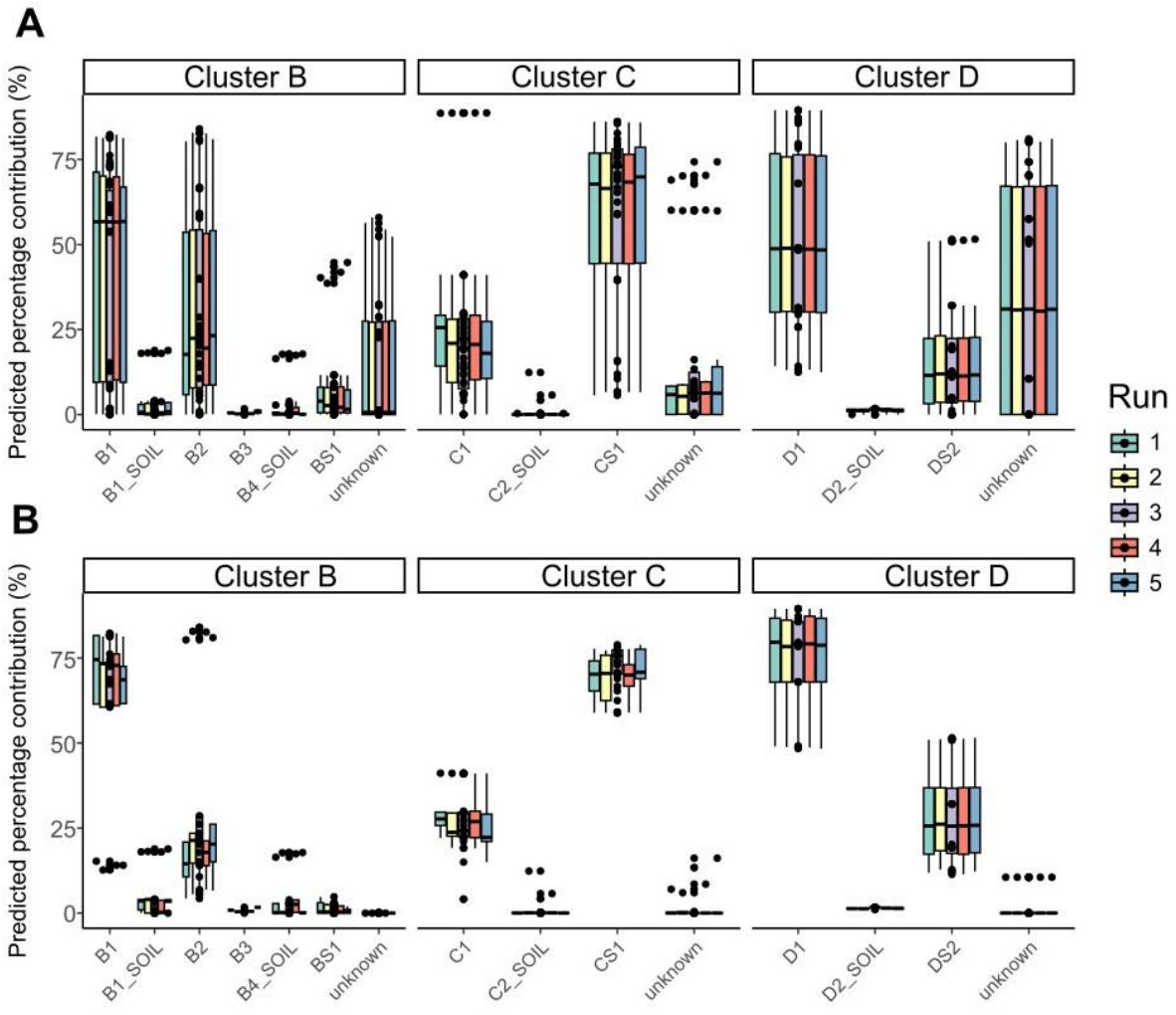


Figure 4-5 Results of FEAST analysis using (A) microbial community and (B) the resistome, showing the relative contribution of various sources in Clusters B, C and D. The results from five runs of FEAST showed there was no significant difference between percentage contribution in different runs (microbiome: Kruskal Wallis p=0.9999, resistome, Kruskal Wallis: p=0.9997)

Cluster B

Based on the microbiome data, the FEAST analysis revealed a high percentage contribution of upstream river water microbiomes, particularly in summer, where B1

and B2 contributed an estimate of 38 ± 19% and 36 ± 19% (mean ± standard error), compared to winter, where the site B1 contributed an estimate of 4.9 ± 3.3% and B2 had an estimate contribution of 30 ± 29%. There was also a higher amount of unknown source contribution in winter months compared to summer (39 ± 16% and 21 ± 10% respectively). The wastewater effluent (BS1), which was discharging into a tributary upstream of site B4, was found to have a higher percentage contributing to the downstream microbiome in the winter compared to the summer at 25 ± 16% and 6 ± 3% respectively. Soil samples had very low contributions throughout.

Based on the FEAST source contributions using the resistome data, there was a notably higher percentage of the source ARGs attributed to the upstream river when compared with microbiome-based estimates. This was particularly the case for site B1, which was estimated to contribute 72 ± 6% in summer and 44 ± 30% in winter. There was also a lower estimated contribution of the wastewater effluent BS1, which was estimated to contribute around 1% in both summer and winter. The percentage of unknown sources for resistance genes were also predicted to be almost zero, which was found to be significantly different from microbiome-based estimates (t-test, $p<0.05$)

Cluster C

Based on the microbiome data, the wastewater effluent source (CS1) was found to contribute highly to site C2 at 44 ± 23% in summer and 47 ± 35% in winter. The sampled upstream river site (C1) was also found to contribute highly to the downstream microbiome in summer, although with large variability (32 ± 29%), whereas in winter, the upstream river was contributing less to downstream microbiomes (14 ± 4%). Unknown sample contributions were higher in winter compared to summer (Winter: 40 ± 31%, Summer: 24 ± 18%) and soil contributions were found to be negligible in summer and winter.

Overall contributions of sources were not found to be significantly different based on microbiome and resistance gene data. However, based on resistance gene data, CS1 was estimated to contribute more highly to ARGs at the downstream site, at 69 ± 5% in summer and 72 ± 3.6% in winter. Like Cluster B, the unknown contribution for resistance gene data was estimated to be very low, i.e., <1% in the summer and 5 ± 5% in the winter.

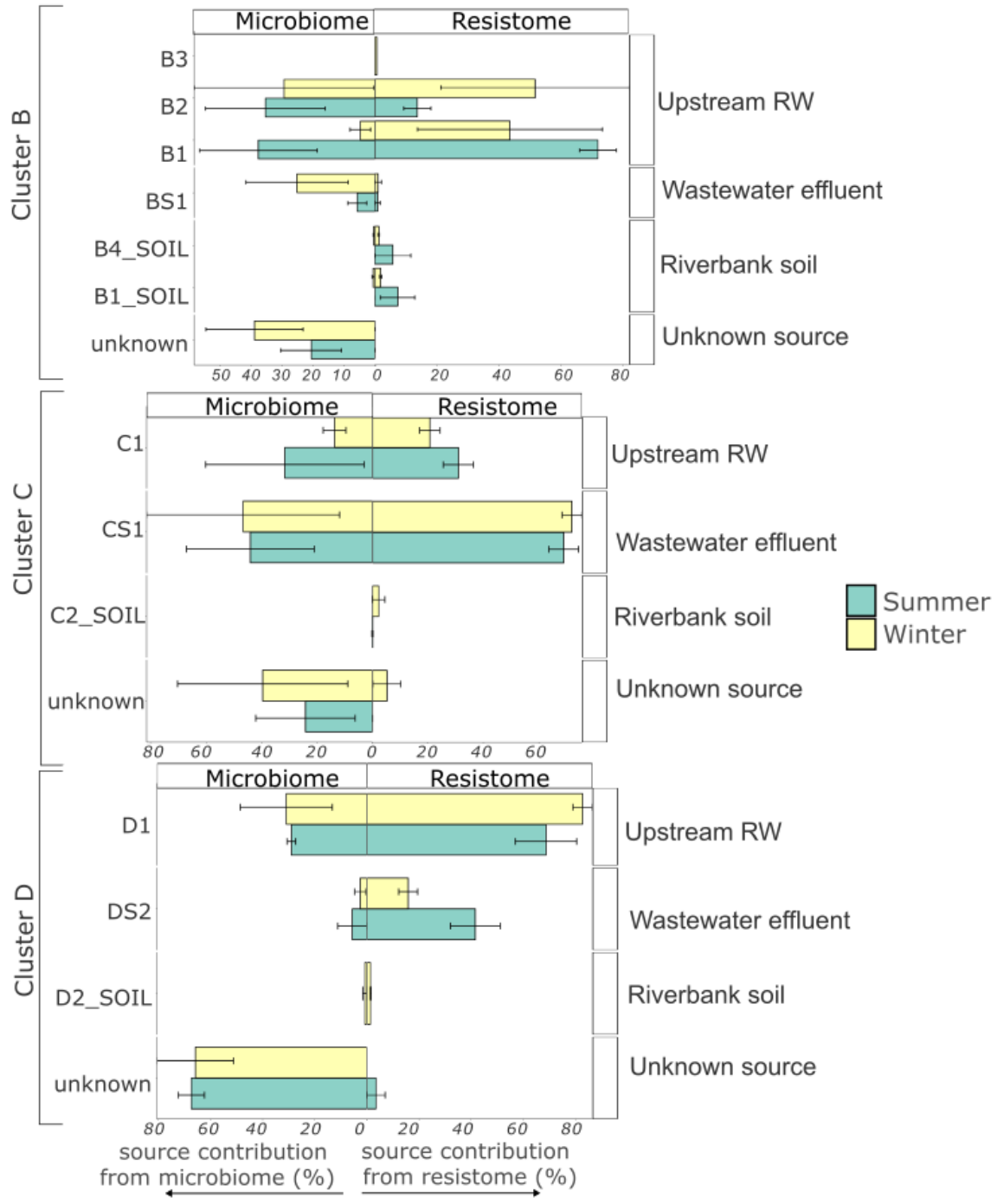


Figure 4-6 Estimated source contributions from the FEAST analysis for sink B4 (top), sink C2 (middle) and sink D2 (bottom)

Cluster D

Unlike Cluster C, based on microbiome data, the estimated source contribution to microbial community at the wastewater effluent site (DS2) was low (6 ± 6% in

summer and 3 ± 2% in winter) and most of the contributing microbial community was from unknown sources (67 ± 5% in summer and 66 ± 15% in winter).

When considering the resistance gene data, the predicted contribution of sources differs from that based on microbiome data. The contribution of ARGs in the upstream river was significantly higher compared to the microbiome estimations (t-test, p<0.01) (69 ± 9% in summer and 83 ± 4% in winter), and the wastewater effluent contribution to in-situ ARGs was higher (42 ± 10% in summer and 16 ± 4% in winter). In contrast, the unknown contribution was significantly lower based on resistance gene data compared to microbiome data (t-test, p<0.001) at 4 ± 4% in summer and <1% in winter.

4.3.3 Environmental drivers of microbial community and ARG persistence

Changes in the resistome in river samples in the Clusters were assessed through volcano plots (Figure 4-7), showing the log_2fold change of the absolute abundance of ARGs and MGEs (copies/mL) plotted against statistical significance (Wald test, p<0.05).

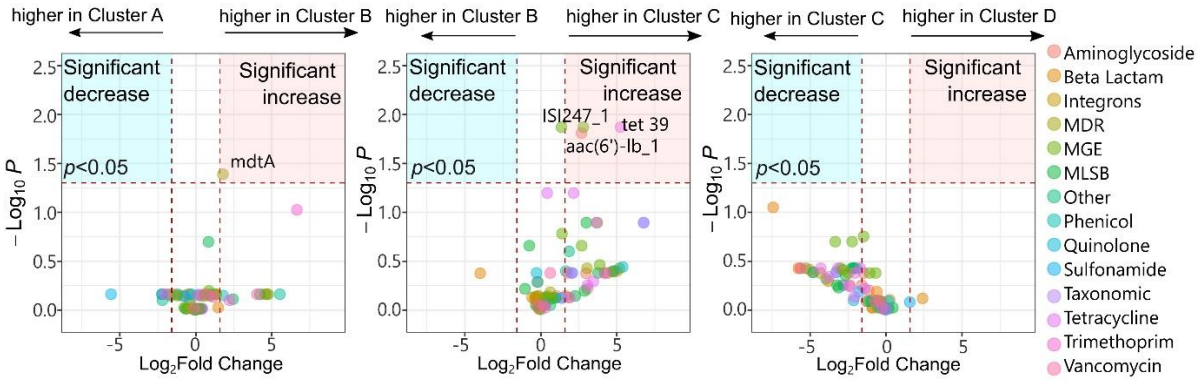


Figure 4-7 Volcano plots displaying Log2fold change against the statistical significance (Walds test, p<0.05), applying Benjamin Hochberg adjustment. Volcano plots represent the log2fold change of genes in river water sites in Cluster A, B, C and D

In Cluster B, only the MDR gene *mdtA* significantly increased relative to Cluster A (Wald test, p<0.05). Further downstream in Cluster C, there was an increase in resistance genes relative to Cluster B, with two ARGs and one MGE increasing significantly, the tetracycline ARG *tet39,* the aminoglycoside ARG *aac('6')-Ib_1* and the MGE *ISI247_1*) (Wald test, p<0.05). From Cluster C to Cluster D, there was a decrease of ARGs, with no significant increase of ARGs. There is, therefore, not a

consistent increase (or decrease) of ARGs and MGEs from upstream to downstream in the Coquet.

To understand the potential drivers for microbial community in different sampling clusters, Mantel tests were conducted to assess the correlation of the weighted Unifrac distance metric of the microbial community, with the Bray Curtis dissimilarity coefficient for water quality parameters, flow rate and microbial source tracking markers (human and ruminant *Bacteroidetes*: HuBac and RuBac) (Figure 4-8).

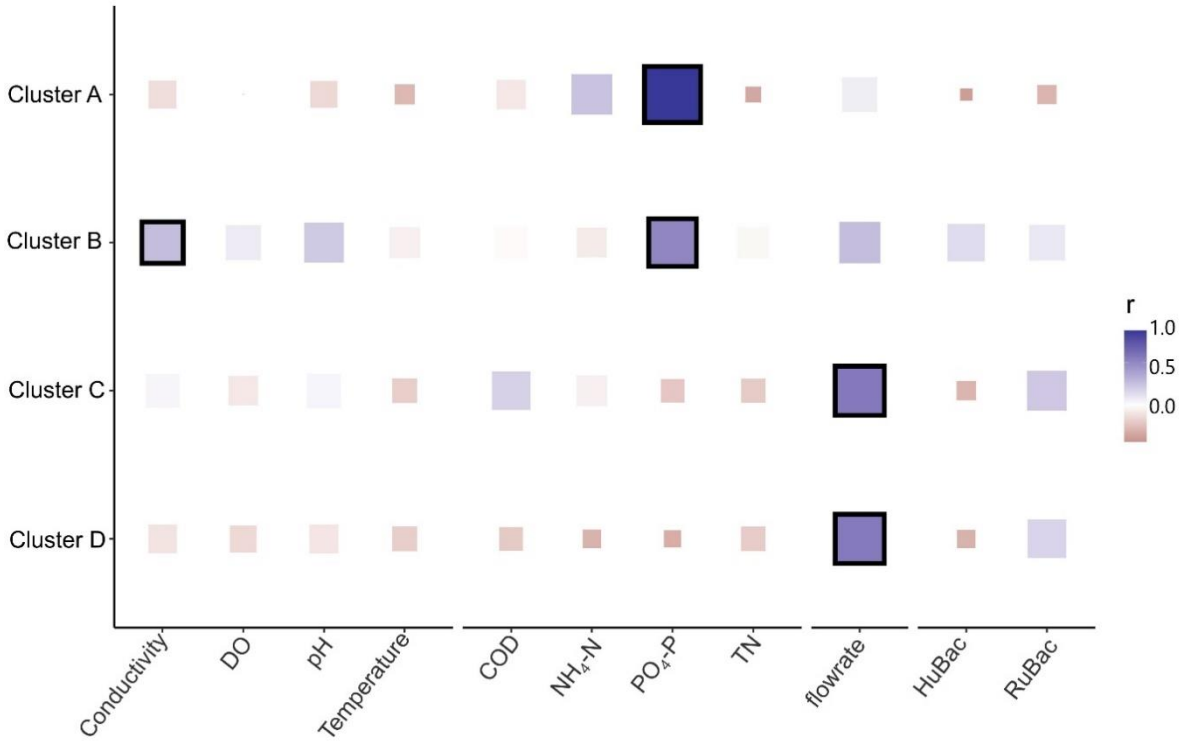


Figure 4-8 Mantel test results showing Spearman correlation between weighted UniFrac distance for microbial community (QMP) and Bray-Curtis dissimilarity coefficient for water quality parameters. The fill colour and size of the box indicates the correlation coefficient, and a black outline indicates the correlation is significant ($p<0.05$), which is adjusted with a Benjamini Hochberg correction.

In Cluster A, PO_4-P was strongly and significantly positively correlated with the microbial community ($r>0.8$, $p<0.05$). In Cluster B, conductivity was moderately significantly positively correlated with the microbial community ($r>0.25$, $p<0.05$) as well as PO_4-P.

In Cluster C and D, the flowrate was positively correlated with the microbial community ($r<0.5$, $p<0.05$). Both Cluster C and D had negative correlations with

water quality parameters, such as conductivity and DO, and positive correlations with RuBac.

The impact of nutrient limitation through the sites sampled in this study is summarised using N:P ratio data in Figure 4-9 and Appendix Table B-10. There was a wide range of observed N:P ratios, however most of the sites in the catchment are P-limited (i.e., higher N:P ratios); following previous studies finding an N:P value >16 is indicative of P limitation to the microbial community (Koerselman and Meuleman 1996). The exception is sites immediately downstream of wastewater effluent (B3, C2 and D2), which have N:P ratios below 16, indicating an increased supply of phosphorous and limitation of another nutrient, such as nitrogen or carbon. However, further downstream from the wastewater release point, the river appears to return to being P-limited again.

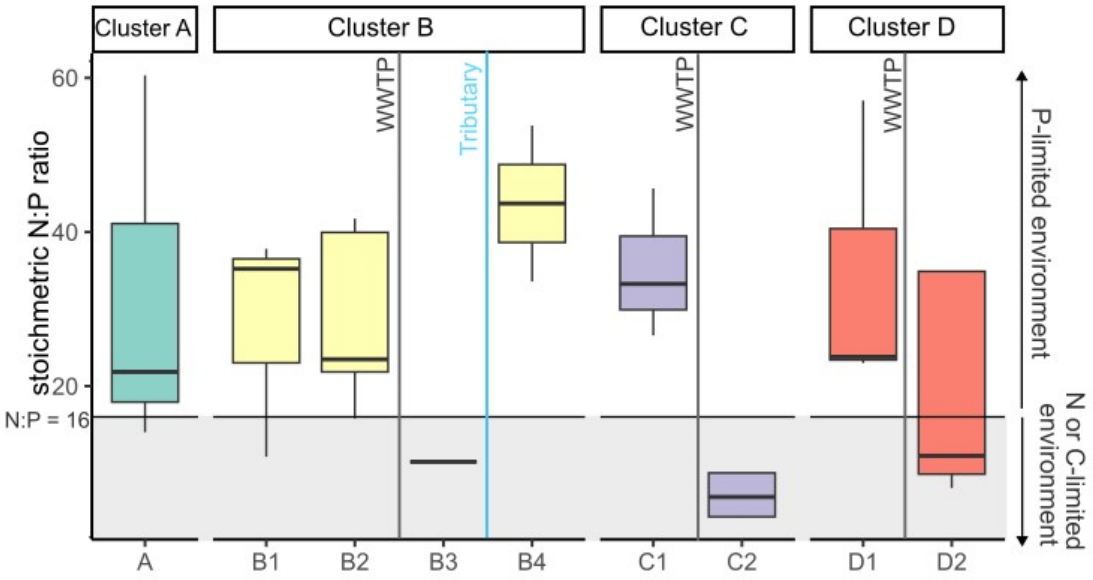


Figure 4-9 Box plots showing N:P ratio for river water sites down the Coquet catchment Point sources including wastewater effluent (WWTP) and tributary sites are indicated in the bar plot. The threshold for P-limitation at N:P=16 is indicated on the plot

4.4 Discussion

4.4.1 Importance of point sources in the Coquet catchment

This analysis of the Coquet catchment permitted the investigation of the sources and drivers influencing the river microbial communities and resistomes. FEAST analysis was applied and revealed that for the majority of sites along the catchment, point sources such as wastewater-like sources or specific soils, were less important for the

makeup of the microbial community. The high percentage of unknown source contribution at the sink site in Cluster B (B4), suggested large contributions of diffuse sources (e.g., non-point source runoff), that then become the driver of microbial community composition at downstream river sites. Therefore, in this rural environment, it is likely that the microbial community is not driven by defined sources, but by unknown diffuse sources that alter nutrient conditions that, in turn influence microbial selection in the river. Whilst this particular study investigated the sources of AMR and microbial communities in sites along the Coquet river, there is further opportunity for sub-analyses to focus on the impact of the wastewater effluent within the tributary site, through investigating site BS1 as a source for downstream site B2.

When the FEAST analysis was employed on the resistome data obtained through HT-qPCR (quantifying 96 genes), unknown sources had a substantially lower contribution and wastewater effluent was estimated to have a higher contribution to downstream resistomes, compared to microbiomes. Typically, in the context of AMR, the FEAST approach is used on metagenomics data, which has the capability of identifying a wide range, as well as novel ARGs (Waseem et al. 2019). The apparent low contribution of unknown sources could indicate that unknown or diffuse sources do not contribute to the resistome in the sink sites and contributing sources are mostly the upstream river and in the case for the sink site in Cluster C (C2), wastewater effluent. However, as the HT-qPCR in this study only identified 96 genes, interpretations such as this should be treated with caution. Shenhav et al. (2019) previously found the accuracy of FEAST decreased with sequencing depth, indicating that near-complete information is required for this model. In addition, a recent review which included 41 studies from 19 countries, demonstrated aquatic environments have abundant native ARGs and contaminant sources are not the sole contributor (Hooban et al. 2020). Due to the high contribution from unknown sources to the downstream microbial community, unknown sources will likely additionally shape the natural resistome. However the resistance genes from these unknown sources may extend beyond the 96 gene assay. Therefore, further work should use FEAST analysis with a wider array of resistance genes obtained through HT-qPCR to identify additional contributing sources that may be contributing to the natural resistome.

Whilst most of the sites along the Coquet had a low contribution from point sources, Cluster C was a notable exception. At the sink site C2, there was an increase of ARGs, MGEs and microbes associated with wastewater effluent input, overriding more diffuse sources at a local level. C2 contained a high abundance of ARGs, especially tetracycline ARGs, which have been typically identified in rivers heavily influenced by urbanisation and anthropogenic activity (Chen et al. 2013; Ling et al. 2013).

Whilst the sink sites in Cluster C and D (C2 and D2 respectively) had similar sampling conditions, where a wastewater effluent release was immediately upstream of the site, only C2 exhibited an influx of ARGs and MGEs. The difference between the downstream ARG abundance in C2 and D2 could be due to the differences between the ratio of river flowrate and wastewater effluent flowrate, which would impact the dilution of ARGs in the receiving river (Sabri et al. 2020). Whilst the flowrates were not measured at these sites, the percentage of wastewater flow contributing to the river flow was approximated using simulated data from Chapter 5, and was estimated to be 0.06% for site C2 and 0.02% for site D2. Therefore, there is likely less dilution of the receiving wastewater effluent at site C2, leading to a local increase in resistance. The impact of advection dispersion processes could also be impacting the level of dilution. Whilst downstream river sampling was taken at a sufficient distance from the point source input to allow the river to be sufficiently mixed, factors such as river flow rate and turbulence would have influenced how efficiently wastewater effluent is mixed with the river water. Further analysis should consider several sites at different distances downstream of the point source site, sampling at different seasons to capture the impact of seasonality and flow rate.

Another possibility for the difference between C2 and D2 could be the location of the discharge pipe in the rivers. At site C2, the upstream wastewater effluent discharge pipe was located in close proximity to the riverbank, whilst at site D2, the pipe was located in the centre of the river. Due to shear stress, the velocity of the river is faster in the middle than along the banks (Han et al. 2015). The downstream transport and advection of the wastewater effluent would likely be more efficient in the centre of the river, which could be what is observed in site D2 (Haberstroh et al. 2021). This would likely reduce the coalescence of native and non-native taxa and make the exchange of ARGs challenging (Zhang et al. 2022), where conjugation processes require direct

cell to cell contact (Michaelis and Grohmann 2023) and transfer of genetic material can take several minutes to hours (Raleigh and Low 2013). The impact of location of the discharge pipe is to the authors knowledge not an area previously explored, and it could be a contributing factor in the rapid dissemination of non-native microbes and ARGs, and an area for further research.

4.4.2 Persistence of non-native microbes and ARGs along the Coquet catchment

Environmental drivers that contribute to the survival of specific ARGs and microbes are not well understood (Mahaney and Franklin 2022). The volcano plot analysis revealed between Clusters along the river, there was no long-term accumulation of ARGs/MGEs from up to downstream. Previous studies have similarly observed a local increase in antibiotic resistant bacteria downstream combined sewer overflows (CSOs), which have rapidly declined further downstream (Mahaney and Franklin 2022). This work investigates the environmental drivers from an ecological perspective which influenced the persistence of non-native microbes and resistance genes in the Coquet catchment.

An assessment of environmental factors on the microbial community through the Mantel test revealed correlations with PO_4-P concentration in upstream locations, something which was not apparent in downstream sites. Physiochemical parameters driving microbial community and resistome structures have been previously observed (McCann et al. 2019; Zhou et al. 2020; Wu et al. 2023). The link with phosphorous builds on previous work, investigating ARG abundance in the pristine and naturally P-limited Arctic soils (McCann et al. 2019). The findings of the Arctic soil studies indicated ARG abundance was closely linked to increased P-availability, driven through the input from other sources, such as migrating wildlife (McCann et al. 2019). Differences in environmental conditions including nutrient availability and flowrate have previously been shown to have a substantial impact on microbial communities (Xu et al. 2020; Wu et al. 2023). Our results suggest that microbial communities at less human-impacted sites are controlled by nutrient supply, with P as the limiting nutrient in the upper Coquet. This means that ecological conditions select for resident bacteria. Therefore, ARGs found at such non-impacted sites are ARGs associated with strains selected by nutrient conditions (Wu et al. 2023).

Moreover, flow rate was positively correlated with microbial community composition in downstream (i.e., Cluster C and D) sites, which has previously been observed in assessments from other studies (Baral et al. 2018b). As the measured flow rate was consistently higher in winter compared to summer, this correlation could reflect the seasonality in the catchment influencing the microbial community (Staley et al. 2015). Flow rate as an environmental driver for microbial community composition may indirectly explain changes in the limiting nutrient condition, possibly due to increased source inputs under high flow conditions. Further, as flow rates increase, local resident times within parcels of water become shorter and provide selective conditions for faster growing strains in such reaches. This interpretation could be assessed through further studies.

The stoichiometric N:P ratio was measured along the Coquet River in order to investigate the influence of nutrient limitation on microbial communities and resistomes. The results indicated the Coquet is primarily a P-limited environment, except for immediately downstream of wastewater effluent discharges. Assessment of the stoichiometric ratios in UK rivers have found that most headwater streams and rivers are P-limited, and a high concentration of nutrients are often associated with heavily urbanised catchments, or those with intensive agriculture (Jarvie et al. 2018). Interestingly, the stoichiometric ratio mirrors the pattern observed with the resistance genes in the volcano plots, where in between clusters, the river ‘recovered’ to natural P-limited conditions. In this river system, the nutrient limited conditions may impact the survival of non-native bacteria, where the native bacteria have better adapted to the nutrient limited conditions than those introduced from the wastewater effluent (Hibbing et al. 2010). The findings of this study indicate that in the absence of selective pressure and nutrient availability, the temporary influx of faecal contamination and non-native ARGs are generally not maintained down the course of the river. Further research could investigate the impact of light exposure, which has previously been found to impact microbial communities (Ensz et al. 2003), and measure carbon (C) to investigate carbon availability/limitation on microbial communities and resistomes persistence in rural rivers.

Overall, maintaining healthy water quality, effective wastewater management and reducing runoff from agricultural sites is important to reduce the potential for selective pressure and maintenance of novel ARGs (Singh et al. 2019). Further

exploration of less impacted rural rivers are required to corroborate the findings of this study.

4.5 Conclusions

This study aimed to identify the contributing sources of the microbial community and AMR, as well as the factors influencing the persistence of these sources in the Coquet catchment. The FEAST approach was used to quantify the contributions of the microbial community and resistome. In general, most of the contributing taxa were either from diffuse sources or native taxa except for Cluster C, which had a high faecal input. It was theorised this could be due to either differences in wastewater/river flow rate ratio, or the placement of the discharge pipe within the river, which warrants further exploration, and possible selection for strains based on ambient nutrient conditions in the river. It was observed that novel ARG dissemination was not maintained in downstream sites, likely due to an absence of selective pressure. In addition, the nutrient limiting environment may impact the survival of non-native taxa, where they are not adapted to the nutrient limited conditions, which in turn, suggests that the presence of specific ARGs is resulting from host selection of other reasons. Therefore, mitigation strategies should focus on maintaining adequate water quality and wastewater management, which reduces the potential for selective pressure and maintenance of novel and more clinically relevant ARGs. Further evidence from clean river water sites is needed to support these findings, as well as additional research on the impact of nutrient availability and light conditions on the persistence of AMR.

Chapter 5 Application of the Soil and Water Assessment Tool (SWAT) to inform environmental AMR monitoring: A case study in the Coquet River catchment

5.1 Introduction

Performing antimicrobial resistance (AMR) monitoring in the environment is important because it provides an indication of how and where anthropogenic activity is impacting the wider resistome (UNEP 2023), including evolution of new AMR in environmental reservoirs (Bengtsson-Palme et al. 2023). Within a river catchment, monitoring can be used to assess the local environmental impacts of direct AMR and related releases from point versus diffuse sources, whilst determining differences between catchments related to hydrologic factors and land use (Bengtsson-Palme et al. 2023a).

Catchment scale studies such as the studies in this thesis, can provide a useful insight into environmental drivers of AMR. This is crucial for assessing AMR from a One Health perspective, where environmental AMR exposures can potentially impact human, animal, and crop health (UNEP 2023). However, AMR monitoring in the environment is currently not standardised, is expensive and resource- and time-consuming (Meshesha et al. 2020). Particularly when the seasonality and stochasticity of environmental settings are concerned, designing a monitoring approach through catchment sampling may be challenging in terms of knowing what, where and when to monitor (Bengtsson-Palme et al. 2023).

The main multinational environmental AMR monitoring programme is the World Health Organisation's (WHO) Tricycle Protocol, where extended-spectrum beta lactamase-producing *Escherichia coli* (ESBL *E. coli*) is plate-cultured to enumerate isolates as an indicator of AMR. ESBL *E. coli* are generally resistant to third and fourth generation beta-lactam antibiotics and enumeration through colony counting is an effective and low-cost method to quickly estimate AMR in the environment. ESBL *E. coli* abundances are not a perfect indicator, but their concentration is often correlated with other AMR determinants (Bengtsson-Palme et al. 2023), and can be used as an indicator to identify and triage areas of interest (e.g., AMR "hot spots"), and where deeper analysis might be conducted (e.g. genomic and metagenomic analysis). Monitoring *E. coli* is also part of human and livestock AMR surveillance

programmes (Anjum et al. 2021), such as the WHO Global Antimicrobial Resistance and Use Surveillance System (GLASS) programme (WHO 2020).

In river catchments, such as the ones studied in this thesis, waterborne contamination from agricultural areas can be important sources of AMR. However, diffuse inputs of AMR may not have been precisely captured and identified because sampling sites tended to focus on sites immediately following known point sources (i.e., wastewater treatment effluent release). In addition, diffuse contamination may vary and potentially become heightened during periods of higher rainfall and river flowrates, due to increased runoff (see Chapter 3) (Ahmed et al. 2018; Almakki et al. 2019). There is, therefore, a need for tools to understand and then approximate locations where there may be heighted environmental AMR and greater exposures to humans, animals, and crops along the river. This requires a numerical model.

The research into models that describe the spread of AMR in river systems has been increasing (Van Heijnsbergen et al. 2022) and whilst the field is still in its infancy, they have the potential to be used as important decision-making tools (Opatowski et al. 2011). Unfortunately, systematic reviews on AMR models demonstrate the vast majority have focused on human health (Birkegård et al. 2018; Niewiadomska et al. 2019), most not even considering extending models to environmental settings. However, models that use Geographic Information Systems (GIS) can use spatial data analysis to explain the diffuse and point sources in aquatic environments and have potential use for AMR. This includes models such as the GREAT-ER (Geo-referenced Regional environmental Exposure Assessment Tool for European Rivers), which has previously been adapted to assess the loadings of *E. coli* and ESBL *E. coli* in the Dutch-German Vecht River catchment (Van Heijnsbergen et al. 2022; Niebaum et al. 2023). While GREAT-ER was found to be successful for predicting *E. coli* and ESBL *E. coli* concentrations, adaptation of the GREAT-ER model to simulate bacteria concentrations is not yet open-source. An open-source AMR model for simulating ESBL *E. coli* concentrations in river catchments would be a large step forward, where its use could widely inform AMR monitoring.

The Soil and Water Assessment Tool (SWAT) is an open-source physical-based watershed-scale model developed by the United States Department of Agriculture, which had an initial purpose of predicting the temporal impact of land-management

practices on water (Saleh and Du 2004). SWAT has had a variety of applications, including simulating microbial concentrations, through editing bacterial inputs from both point and diffuse sources (Sadeghi and Arnold 2002). SWAT models have been used for predicting faecal coliforms (Niazi et al. 2015), *E. coli* (Coffey et al. 2010a) and even the parasite Cryptosporidium (Coffey et al. 2010b). However, there is currently little research into the application of such models on AMR transport and fate. A SWAT-based model was used on ARGs in the environment in a Doctoral and Masters thesis from Virginia Tech University (House 2020; Thilakarathne 2020). This work centred on the development of a SWAT-ARB model for various catchments in the US. However, determining the accuracy of predictions from the model has been challenging due to the lack of observed data, and this model adaptation is also not yet open-source.

The aims of this chapter are to build a SWAT model to simulate the spatiotemporal distribution and dynamics of *E. coli* and ESBL *E. coli* in the Coquet river catchment. This chapter uses data from the sampling campaign described in Chapter 4 to model *E. coli* and ESBL *E. coli* levels in the river, through inputting the wastewater effluent and agricultural inputs along the river and comparing output concentrations to the observed data at the sample site locations. In addition, the sources of the bacteria at different sites in the SWAT model are calculated and compared to the source estimations from the FEAST analysis from Chapter 4.

5.2 Materials and Methods

5.2.1 SWAT model

The SWAT model is a semi-distributed watershed based hydrological model, which has the purpose of simulating the impact of land management practices on flow and water quality. The key components include climate, hydrology, erosion, sedimentation, plant growth, nutrients and the impact of point sources (i.e, wastewater), as well as runoff of pesticides and manure from agricultural processes. The SWAT model operates through dividing a watershed into sub-basins, and then further into hydrological response units (HRUs). An individual HRU is a combination of homogenous land-use, soil type and slope in a sub-basin. The concept of HRUs are so hydrological models can simulate the response of different land surfaces to precipitation, runoff and other hydrological processes. For user accessibility, the SWAT model operates through the ArcGIS extension software ArcSWAT, which

automates entry of data, transforms raw data into input files that can be read by the SWAT model and allows Geographic Information System (GIS) data (e.g., location of point sources, land-use distribution, slope etc.) to be easily integrated into the hydrological model.

5.2.2 Sub-basin delineation

The study area is the Coquet River catchment in Northumberland (NE England), which spans from the Cheviot Hills to the town of Amble at the North Sea coast. The total catchment area is 606 km^2 and the elevation of the catchment ranges from 825m in the hills, to 48m towards the bottom of the catchment. The wastewater effluent discharge sites and monitoring sites for the model were the same as Chapter 4, where sampling was conducted in 'Clusters' as one moves downstream the river catchment (Figure 5-1; Chapter 4.2.2).

The catchment model was constructed using ArcSWAT (version 2012.10.25) on the ArcGIS interface (ArcMap v.10.8.2). Model set-up was also conducted according to Figure 5-2 with the aid of instructional videos (Frankenberger and Daneshvar 2019).

The data sources for the main model inputs are shown in Appendix Table C-1, including ALOS PALSAR, CORINE, and other sources. A digital elevation model (DEM) was used to calculate the topography, flow direction, slope and extract the boundaries of the catchment. The stream network was delineated, and the monitoring points were both estimated automatically by the watershed delineation process (based on the tributary streams) and monitoring sites from Chapter 4 were manually input to enable comparisons with observed data. The catchment was then divided into 41 sub-basins for the purpose of modelling.

5.2.3 Input hydrologic, land-use and climate data

Six soil types were identified from the European Soil Data Base (ESDB 2006). The SWAT model requires input of soil hydrological properties, which were determined through further investigation of various literature relating to the soil geological series (Avis and Harrop 1983). The final properties input into the model are shown in Appendix Table C-2. Land cover data was based on information from the CORINE Land Cover Data set 2018 (CLC, 2018), as this had the closest similarity with SWAT land-use categories (Appendix Table C-3). Hydrological Response Units (HRUs) (see Section 5.2.1 for description) were designed for the model, based on thresholds

of 5% for Land-use, 20% for soil class and 20% for slope class according to Frankenberger and Daneshvar (2019) (i.e., the land-use which occupies less than 5% of the sub-basin in eliminated). Each basin contained multiple HRUs, and there were 195 HRUs in total.

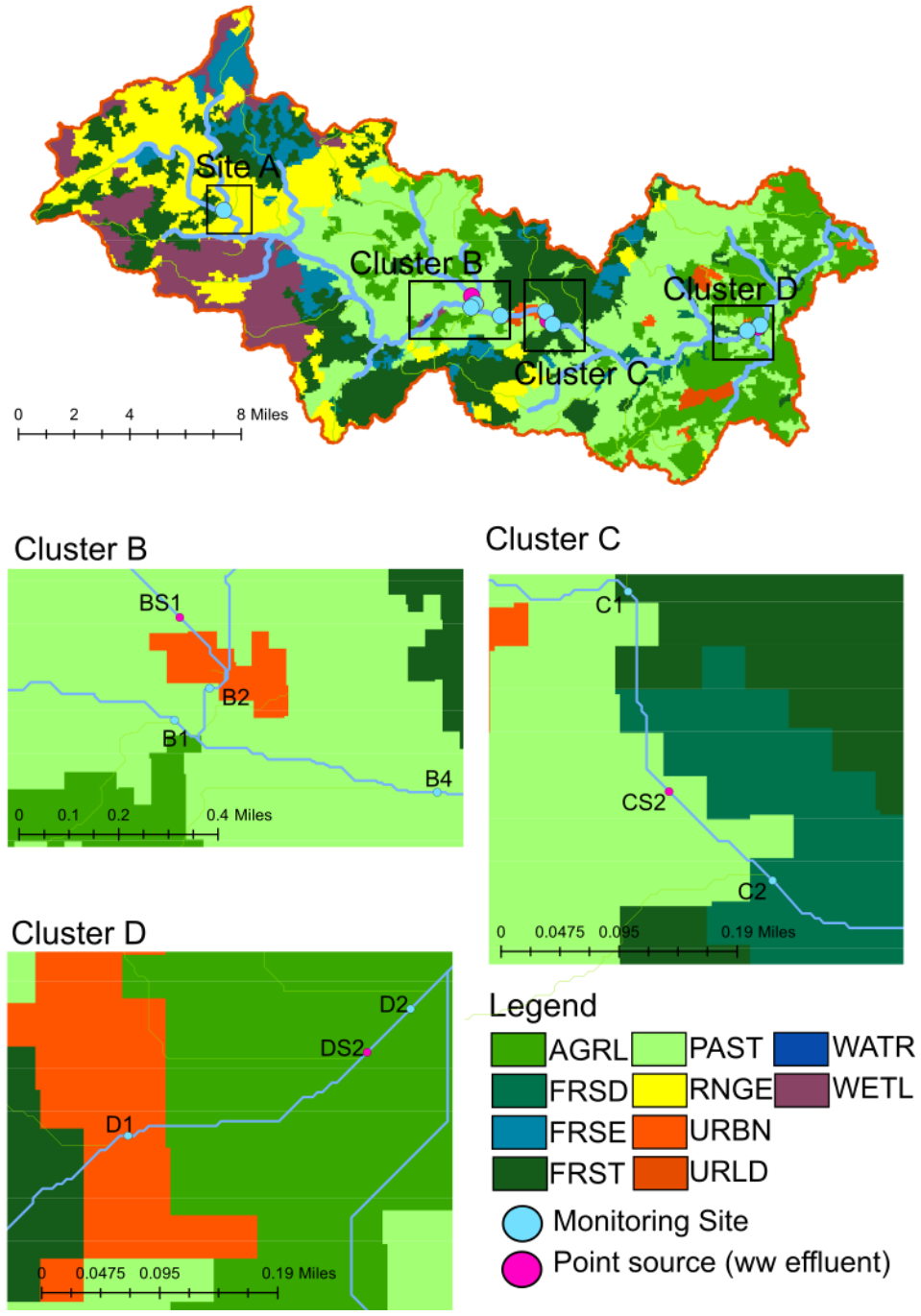


Figure 5-1 Coquet catchment with the monitoring sites selected for model comparison with observed data, including point source inputs from wastewater effluent (ww effluent) and land-use categories, AGRL (Non-irrigated agricultural land), FRSD (Broad-leaved forest), FRSE (Coniferous forest), FRST (Mixed forest), PAST (Pastures), RNGE (Natural Grasslands), URBN (Discontinous urban fabric), URLD (Sport and leisure facilities), WATR (Water bodies), WETL (Inland marshes/Peat bogs)

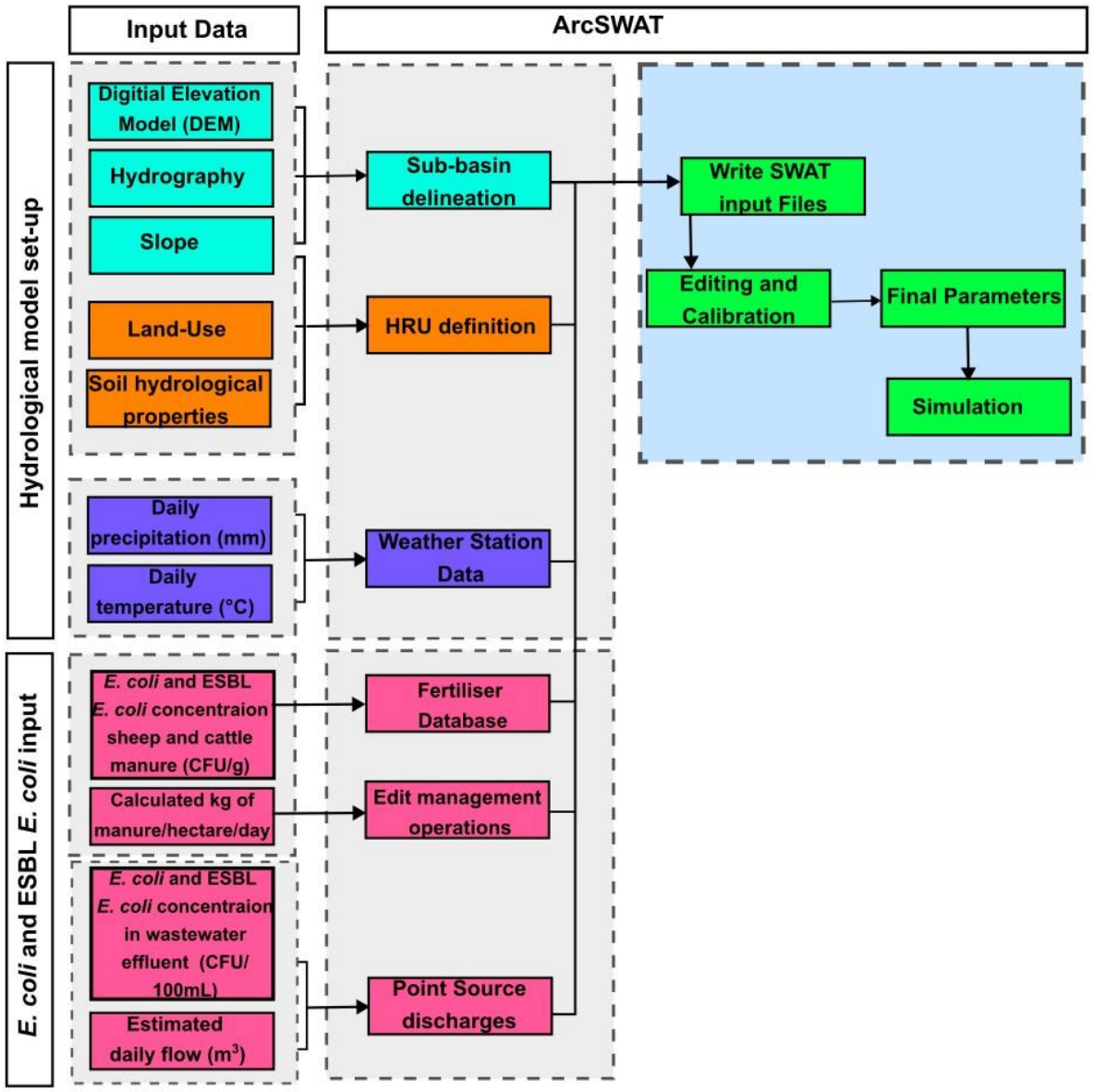


Figure 5-2 Schematic showing SWAT model process

Daily observed climatic data was used for preparing the weather input files for the SWAT simulation. Climatic data included daily precipitation and daily maximum and minimum temperature data over 10 years (01/01/2013-31/12/2022). Daily precipitation data (mm) was accessed from the DEFRA Hydrology Data Explorer from three gauging sites (Appendix Table C-4; (https://environment.data.gov.uk/hydrology/explore). Due to a lack of historical daily temperature data (°C) available, daily maximum and minimum temperatures were estimated using the function *rnorm* in the R environment, based on monthly maximum and minimum temperature accessed from the Met Office Historic Station Data (Met Office, 2023), where Durham was the closest station to the Coquet River

(Appendix Table C-4). As this weather station is located outside the Coquet catchment, there are likely to be minor differences between daily temperature between Durham and sites in the Coquet catchment, however the site was deemed geographically close enough that the temperatures would be similar and acceptable for hydrology simulations.

5.2.4 Bacteria characterisation for SWAT model

SWAT models bacteria transport through source loadings from a range of inputs, including wildlife, livestock, point sources and septic systems. Aiming to replicate the source inputs from Chapter 4, livestock (i.e., sheep and cattle) and wastewater effluent were considered as *E. coli* and ESBL *E. coli* sources for the SWAT model. SWAT models the transport of bacteria through the land through surface runoff or as part of the sediment transported in the river, with the assumption that only bacteria in the top 10mm of the soil can be transported through runoff (Sowah et al. 2020). SWAT has the advantage of simultaneously modelling two types of bacteria, referred to as ‘persistent’ and ‘less-persistent’ bacteria, which are purely descriptive terms and used to differentiate between the two types of bacteria. For the purposes of simulating *E. coli* and ESBL *E. coli* simultaneously, *E. coli* was input as a ‘persistent’ bacterium and ESBL *E. coli* was input as a ‘less-persistent’ bacterium and their persistence parameters were set to identical values. All bacteria concentrations were expressed as colony forming units/100mL (CFU/100mL) unless otherwise specified.

Livestock (diffuse discharges)

The influence of livestock on *E.* coli and ESBL *E. coli* in the Coquet River was approximated through inputting (1) the approximate *E. coli* and ESBL *E. coli* concentration in sheep and cattle manure and (2) approximating the daily mass of manure generated from cattle and sheep in each sub-basin.

Firstly, the *E. coli* concentration in sheep and cattle manure (in colony forming units/gram: CFU/g) was input into the model using values from literature (Table 5-1) through editing the SWAT fertiliser database. Whilst the ESBL *E. coli* concentration of cattle manure could be determined through literature values, the ESBL *E. coli* concentration in sheep manure could not be found in the literature. Therefore, the ESBL *E. coli* concentration in sheep manure was assumed to be the same as cattle manure.

The daily mass of manure generated from cattle and sheep in each sub-basin was approximated and input into the model as follows. Livestock density (i.e., number of cattle and sheep) in each sub-basin was first approximated using agricultural census maps (AgCensus 2020). Based on the kg of manure produced by one livestock unit per day (3.5kg for sheep and 27.5kg for cattle; Coffey et al. 2010a), the number of livestock units in the subbasin and the area of the subbasin (hectares), this was used to approximate the kg/hectare/day in each subbasin of both sheep and cattle (Appendix Table C-5). These values were input into the model through editing land management operations, where in this instance the approximated daily deposition of manure in each sub-basin was fixed for 365 days a year, however further iterations of this model should investigate the annual variability of livestock operations.

Table 5-1 Assumed values of *E. coli* and ESBL *E. coli* in sheep and cattle faeces as model inputs.

Type	***E. coli*** **(CFU/g)**	**ESBL *E. coli*** **(CFU/g)**	**References**
Sheep manure	$1.6x10^7$	10^3	(Coffey et al. 2010a)
Cattle manure	$7.5x10^5$	10^3	(Coffey et al. 2010a; Heuvelink et al. 2019; Gonggrijp et al. 2023)

Wastewater effluent (point source discharges)

The influence of wastewater effluent on *E. coli* and ESBL *E. coli* in the Coquet River was approximated through inputting (1) wastewater effluent sites as a point source input through the ArcSWAT interface and (2) editing the point source input to include concentrations of *E. coli* and ESBL *E. coli* and daily input flow.

Wastewater effluent was input into the model as a point source input, which are loadings into the stream network from sources other than those associated with the land-use. The point sources selected were the measured wastewater effluent sites in Chapter 4 (i.e., BS1, CS1 and DS2; see Figure 5-1). The output of *E. coli* and ESBL *E. coli* from point sources in the river catchment, was approximated using average measured values from Chapter 4 in CFU/100mL (Table 5-2). The daily flow (m^3/day) for each treatment site was calculated based on the approximate water consumption

in England of 150L/per day/per capita (Table 5-2) (DEFRA 2008) and the estimated contributing population (CITY POPULATION 2023).

Table 5-2 Approximate daily flow based on population size (2023), and an average per capita daily flow of 150L/day.

	Population	**Approximate Daily Flow (m^3/day)**	**Average *E. coli* concentration (CFU/100mL)**	**Average ESBL *E. coli* concentration (CFU/100mL)**
BS1	882	132.3	38267	263
CS1	2107	316.05	80533	2733
DS2	1090	163.5	40000	1125

5.2.5 Calibration and sensitivity analysis

Model simulations were conducted to capture conditions in the catchment between 01/01/2017 to the 31/12/2022. A two-year “warmup period” (i.e., 01/01/2015-31/12/2016) was included to allow the hydrology to reach an optimal state. When the SWAT model is first run, internal stores and reservoirs are assumed to be empty, therefore a warmup period can allow these to stabilise and reach optimal values (Kim et al. 2018).

The sensitivity analysis and calibration of the hydrological flux and storage parameters was conducted using R-SWAT (Nguyen et al. 2022). Sensitivity analysis was performed using the Latin hypercube (LH) technique, where nine parameters were selected for calibrating daily flow (Table 5-3). Flow calibration was performed using the dynamically dimensioned search algorithm (Tolson and Shoemaker 2007). This method has been deemed suitable for calibrating complex watershed models and is initiated through a broad search for optimal parameter values, transitioning to a more localised search, whilst dynamically editing the number of parameters being adjusted and automatically calibrating model parameters (Tolson and Shoemaker 2007).

A graphical method that tested ±10% changes of basin parameters and source bacteria concentrations was used for the sensitivity analysis of bacterial outputs. Percentage increases and decreases of output bacteria concentrations were calculated following a parameter adjustment (i.e., ±10% adjustment of the initial basin parameter) to assess the sensitivity of bacteria outputs to that parameter. This

procedure has been used previously in bacterial modelling with SWAT models (Coffey et al. 2010a; Niazi et al. 2015).

5.2.6 Statistical analysis

Model outputs were analysed through R Studio. Graphics were developed with *ggplot2* and finalised with Inkscape (v.0.92.4). Statistics were used to determine the relative accuracy of the model predictions against observed data. Observed daily flow rates were obtained from the national river flow archive (NRFA; https://nrfa.ceh.ac.uk/) in Rothbury and Morwick (see Appendix Table C-6). Observed bacterial concentrations of *E. coli* and ESBL *E. coli* from Chapter 3 and 4 were used to compare with simulated values (Section 4.2.2 and Appendix Table C-7). The coefficient of determination (R^2) was used to assess how consistently the observed and predicted values aligned, where a value of 0 indicates no relationship, and 1.0 indicates a perfect relationship. Generally, R^2 values greater than 0.5 are considered acceptable (Santhi et al. 2001; Van Liew et al. 2007; Coffey et al. 2010a).

The Nash Sutcliffe model efficiency (E) was used to assess the predictive power of the modelled flow rate, where a value closer to 1 is indicative of an accurate model (Nash and Sutcliffe 1970). Classifications of model efficiencies specify $0.75 > E < 1.0$ as very good, $0.65 > E < 0.75$ as good, $0.50 < E < 0.65$ as satisfactory, and $E < 0.5$ as unsatisfactory (Moriasi et al. 2007). To measure the agreement between observed and simulated bacteria, a Wilcoxon Rank test was used to determine if there was a significant difference between outputs.

The contribution of different sources was estimated following the approach of Iqbal and Hofstra (2019), where sources were simulated separately, and each source contribution was then calculated as a percentage of the total *E. coli* and ESBL *E. coli* concentration. This was then graphically compared to the source estimation through FEAST analysis from Chapter 4.3.2.

5.3 Results

5.3.1 Sensitivity analysis and calibration

Flow Rate

The performance of the SWAT model to simulate monthly average flow rates was assessed. Using default parameters prior to calibration, at the sites of Rothbury (midstream) and Morwick (downstream), the SWAT model yielded a good NSE value

of E=0.65 and E=0.65 and coefficient of determination value of R^2=0.61 and R^2=0.63, respectively. Sensitivity analysis indicated the most sensitive parameters in the analysis were (1) CN2 (the initial runoff curve number, which is related to factors such as the soils permeability and land-use), (2) the SURLAG value (i.e., the surface runoff lag coefficient) and (3) GW_DELAY (i.e., the groundwater delay time in days). The final parameter values are shown in Table 5-3.

Following the automatic calibration process through R-SWAT, the estimated monthly average flow rate at Rothbury and Morwick had very good/good NSE values of E=0.76 and E=0.74 and improved R^2 values at R^2=0.69 and R^2=0.75 respectively (Figure 5-3).

Table 5-3 Parameters used for calibrating the modelled flow rates

Parameter	Description	Lower bound	Upper bound	Initial Value	Method	Value chosen
SURLAG.hru	Surface runoff factor (days) (per HRU)	0	5	2.0	Replace	0.1
CH_K2	Effective Hydraulic conductivity in main channel alluvium	0	0.5	0	Replace	0.1
CN2	Initial SCS runoff curve number of moisture condition II	-0.2	1	73	Relative	-0.2
ALPHA_BF	Base flow factor	0	3	0.048	Replace	0.01
SURLAG	Surface runoff factor (for the Basin)	0	5	4	Replace	0.05
GW_DELAY	Ground water delay time (days)	50	500	31	Absolute	81
ESCO	Soil evaporation compensation factor	0.5	0.99	0.95	Replace	0.5
SOL_K	Saturated Hydraulic Conductivity	-0.25	0.5	5	Relative	-0.25
SOL_Z	Depth of soil surface to bottom of layer (mm)	-1	10	Default	Relative	-1

Modelling E. coli and ESBL E. coli

A sensitivity analysis was conducted on *E. coli* and ESBL *E. coli* to identify the most sensitive basin parameters that impact model outputs relative to source bacteria concentrations. Figure 5-4 shows the impact on the model output of modifying these parameters by ±10% from their selected values (see Table 5-4) and the associated percentage increase/decrease of simulated bacterial estimates.

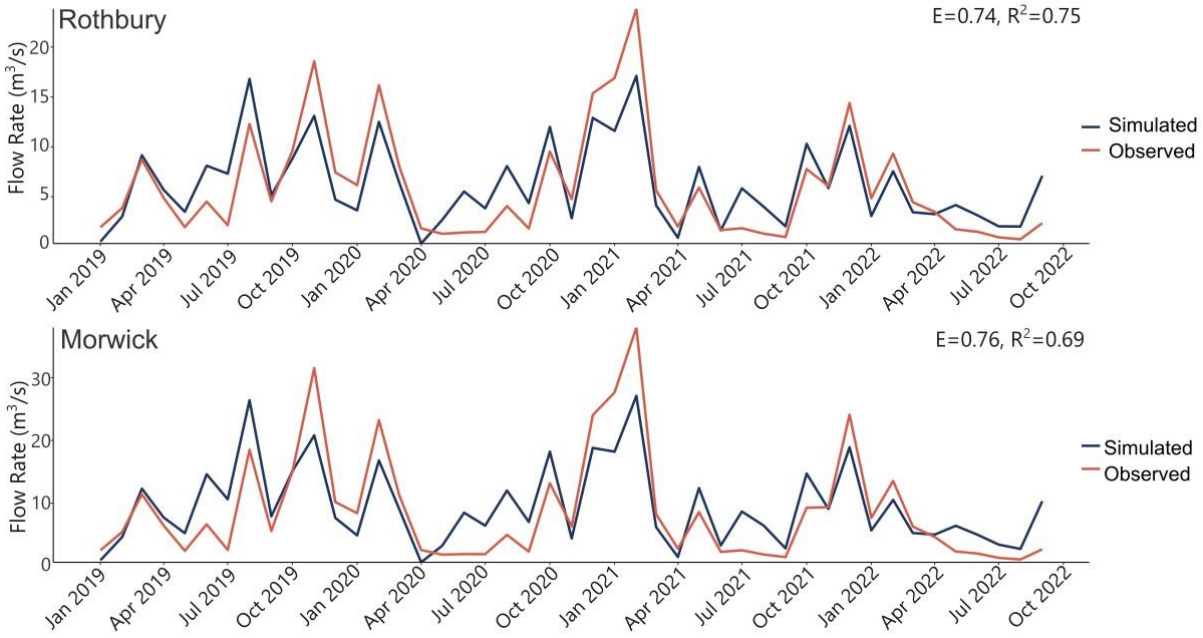


Figure 5-3 Comparison of observed and simulated streamflow in Rothbury (midstream) and Morwick (downstream) in the Coquet

The sensitivity analysis indicated that changing the basin parameters only resulted in a 5-10% increase or decrease in simulated *E. coli* and ESBL *E. coli* levels. From the basin parameters, THBACT (Temperature Adjustment Factor) was found to be particularly sensitive, where a 10% reduction in the parameter resulted in >20% reduction in bacteria concentrations. This particularly impacted *E. coli* outputs. Additional sensitive basin parameters identified included WDPQ and WDLPQ (Die off factor for persistent/less persistent bacteria in soil solution at 20°C) and WGPQ and WGLPQ (Growth factor for persistent/less persistent bacteria in soil solution at 20°C). Altering the source bacteria inputs generally resulted in a 5 to 10% increase or decrease in estimates. For *E. coli*, altering the initial manure concentration by -10% had a greater impact on model outputs compared to +10%. The wastewater effluent bacteria concentration was found to have a near direct 10% increase or

decrease in bacteria outputs when the model inputs were modified. For ESBL *E. coli*, sheep bacteria concentration was found to be the least sensitive and wastewater effluent the most sensitive model inputs.

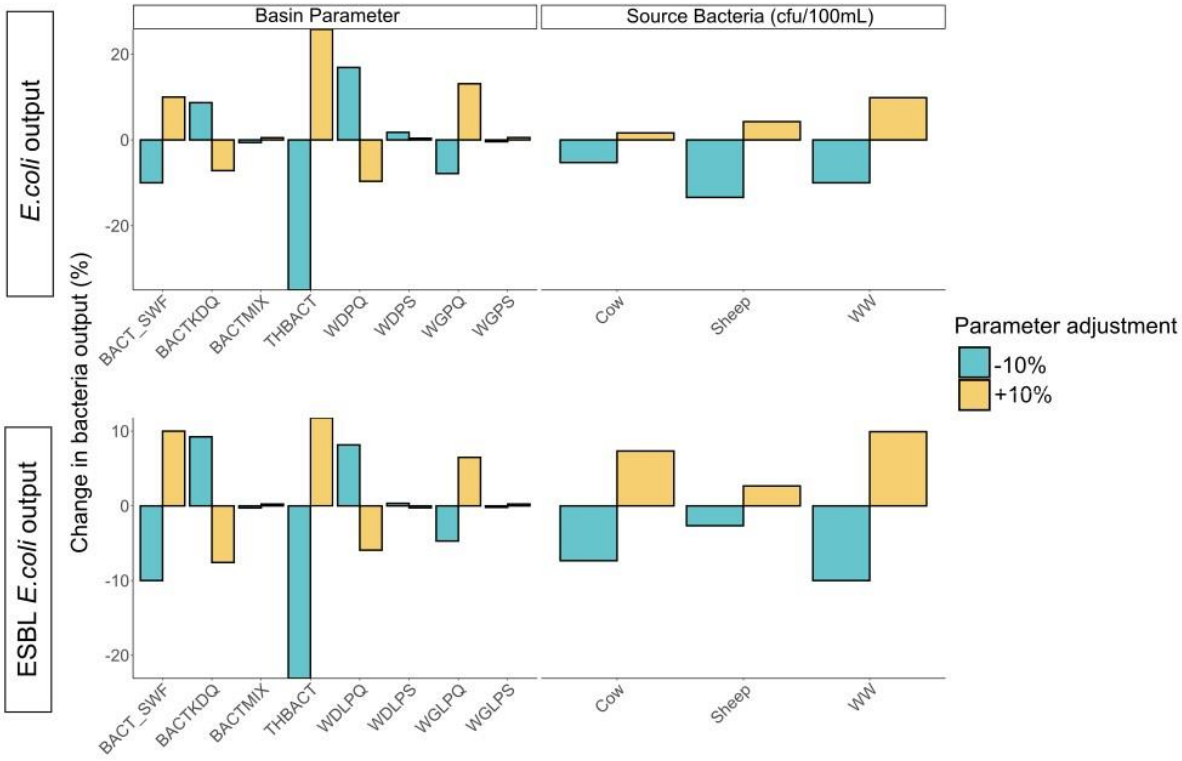


Figure 5-4 Sensitivity of E. coli and ESBL E. coli outputs to basin parameters and source bacteria concentrations (see Table 5-4 for definitions). Source bacteria concentrations (cfu/100mL), include the concentration of E. coli and ESBL E. coli in Cattle/cows, sheep, and wastewater effluent (WW)

5.3.2 Comparison of bacteria data with observed values

The similarity of simulated values and observed bacteria values were determined through the coefficient of determination (R^2) and graphically. The R^2 values (Table 5-5) indicated there was a good relationship between observed and simulated values, achieving an especially good relationship with some sites for *E. coli* in site B1 (R^2=0.6), site B2 (R^2=0.54) and C2 (R^2=0.81). Site B1 achieved a strong relationship with observed values of ESBL *E. coli* (R^2=0.83) however, for most sites R^2 values for ESBL *E. coli* were <0.5. However, the spatial trends are apparent, especially given the limited field data.

Table 5-4 Selected values for basin parameters relevant to modelling bacterial fate

Basin Parameter	Definition	Selected value
WDPQ/WDLPQ	Die off factor for persistent/less persistent bacteria	0.125
WGPQ/WGLPQ	Growth factor for persistent/less persistent bacteria	0.1
WDPS/WDLPS	Die-off factor for persistent/less persistent bacteria adsorbed to soil particles at 20°C	0.2
WGPS/WGLPS	Growth factor for persistent/less persistent bacteria adsorbed to soil particles at 20°C	0.15
BACTKDQ	Bacteria soil partitioning coefficient (m^3/Mg)	175
THBACT	Temperature Adjustment Factor	1.07
WOF_P/ WOF_LP	Wash-off fraction for persistent/less persistent bacteria	0.5
BACT_SWF	Fraction of manure applied to land areas that has active colony forming units	0.75
BACTMIX	Bacteria percolation coefficient (10 m^3/Mg)	10

The results of the SWAT model were viewed graphically and a Wilcoxon Rank test was used to identify if there were significant differences between observed and simulated *E. coli* and ESBL *E. coli* concentrations (Figure 5-5). Results were normalised to log10 values for ease of comparison.

In general, *E. coli* and ESBL *E. coli* observed and simulated values have similar median, means and spread at sampling points particularly in sites B4, C1 and C2. The model also captures spatial trends along the Coquet River, including low *E. coli* concentrations upstream and the increases in the midstream and downstream. However, the Wilcoxon Rank test shows significant differences at site D1 for both *E. coli* and ESBL *E. coli,* at site D2 for *E. coli* and surprisingly at site B1 for ESBL *E. coli* (despite the strong R^2 value) (Wilcoxon test, $p<0.05$). The simulated values at B1, although well correlated to the observed values, were significantly lower than the observed values. The other significant differences at sites were due to the simulated values being higher than the observed values.

Table 5-5 Coefficient of determination values (R^2) at each site. The n values indicate the number of observations and simulations under comparison (see Appendix Table C-7).

	E. coli	**ESBL *E. coli***
Site	**R^2**	**R^2**
A (n=6)	0.007	0.04
B1 (n=5)	0.60	0.83
B2 (n=5)	0.53	0.18
B4 (n=6)	0.0002	0.074
C1 (n=6)	0.20	0.01
C2 (n=6)	0.81	0.38
D1 (n=6)	0.02	$4.5x10^{-5}$
D2 (n=5)	0.57	0.03

5.3.3 *E. coli* and ESBL *E. coli* source allocations

The estimated contributions of the different sources in the model (i.e., sheep, cattle, and wastewater effluent) at different sites along the river are shown in Figure 5-6. The model outputs suggest the input of *E. coli* into the system in the upstream river (i.e., Site A and B1) is primarily due to sheep and cattle faecal matter in agricultural runoff. However, upon inputs of wastewater effluent to the river, wastewater effluent

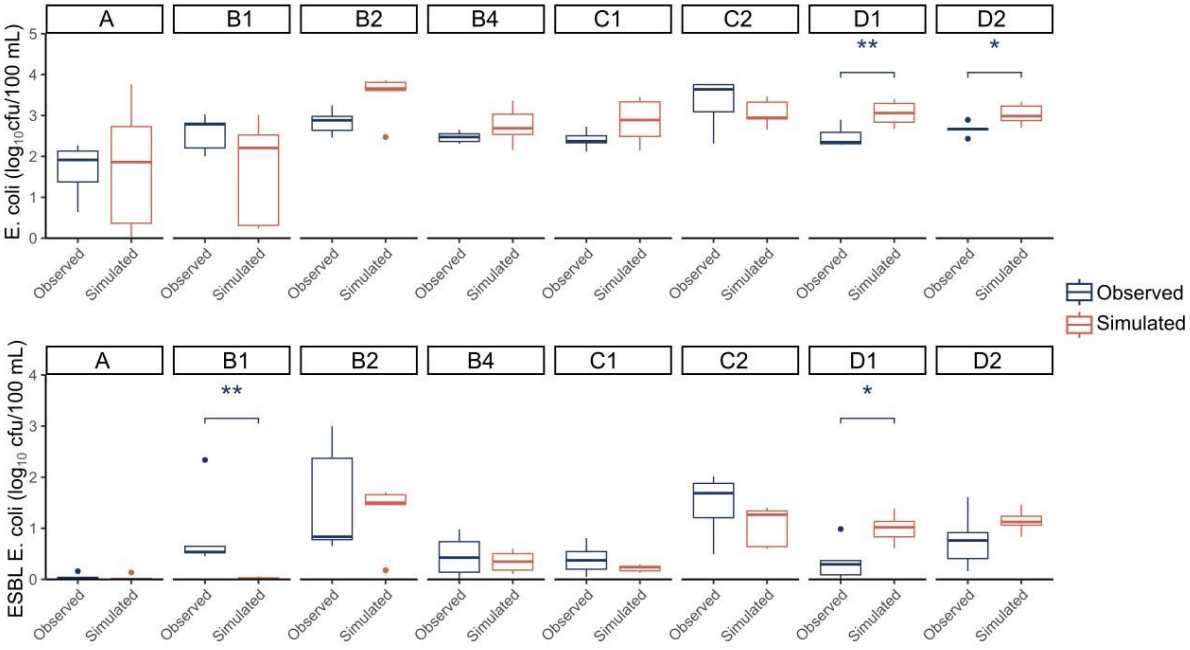


Figure 5-5 General comparison between Observed and Simulated E. coli (top panel) and ESBL E. coli (bottom panel). *A significant difference between observed and simulated values (Wilcoxon test, *p<0.05, **p<0.01)

becomes the dominant source of *E. coli* and ESBL *E. coli* in the system. The concentrations of *E. coli* and ESBL *E. coli* (CFU/100mL) originating from sheep, cattle, and wastewater effluent along the Coquet River are shown in Figure 5-7.

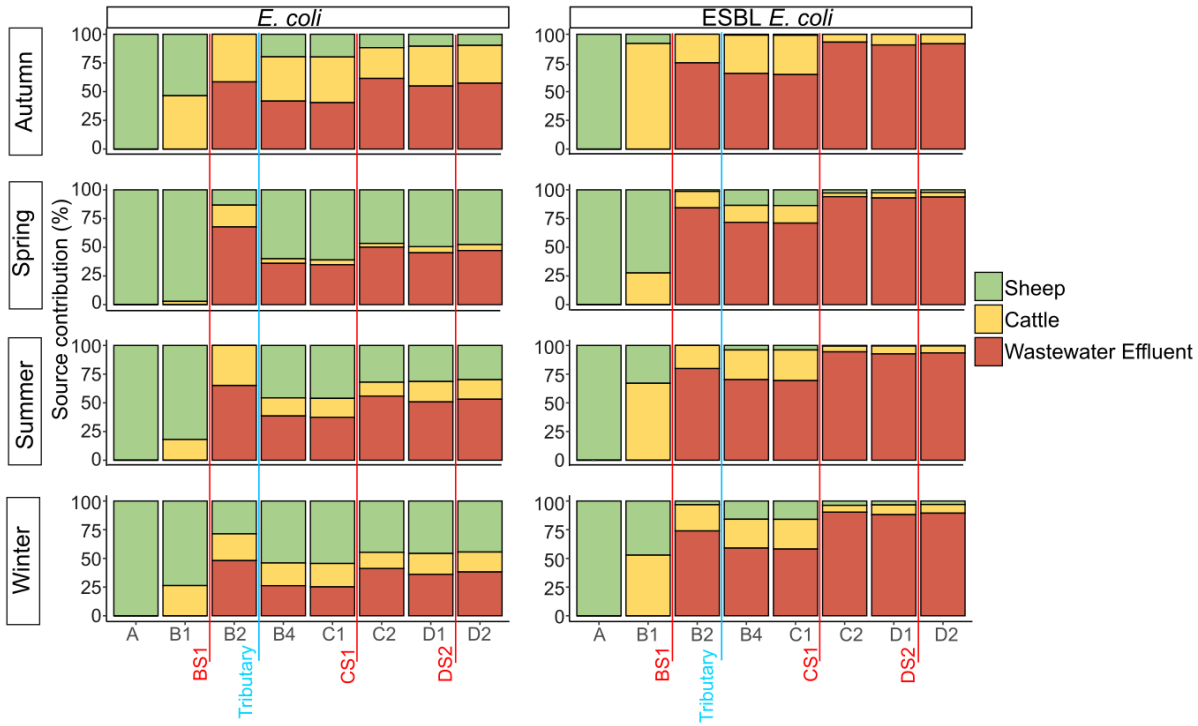


Figure 5-6 Simulated percentage breakdown of sources contributing to downstream concentrations of E. coli and ESBL E. coli. Panels show Autumn (September-November), Spring (March-May), Summer (June-August) and Winter (December-February) over 2019-2022.

Source contributions at downstream sink sites in Clusters B, C and D (site B4, C2 and D2) were estimated through the FEAST approach (see Chapter 4.3.2) and compared to source contributions estimated through the SWAT model over the observed sampling timeframe (Figure 5-8). When comparing the 'unknown' source contributions for the FEAST model estimations (i.e, the diffuse input), with the SWAT model's estimated source contributions for sheep and cattle for *E. coli* (i.e., equivalent of diffuse input), both models reported similar findings. This indicates that when modelling *E. coli*, the SWAT model can estimate diffuse contributions to a sink site. However, the SWAT model estimated point source contribution, (i.e., wastewater contribution) for *E. coli* and ESBL *E. coli* is higher than the FEAST model estimations. It should be considered, however, that the SWAT model cannot separate the impact of the upstream river water in source estimations. It is likely that in the SWAT model, the bacteria input from upstream river water is contributing to

the high percentage of wastewater effluent downstream. The FEAST model shows a high percentage contribution of upstream river water, and it is likely that this is true for the SWAT model as well.

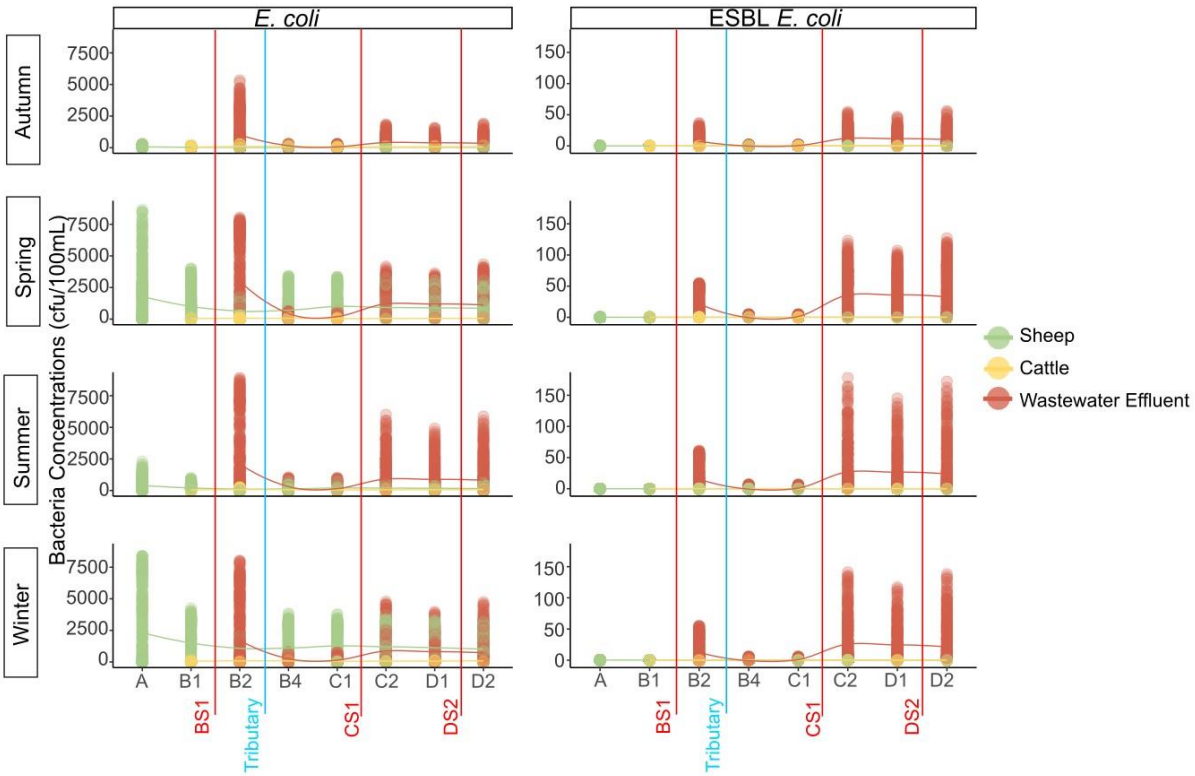


Figure 5-7 Simulated concentrations of E. coli and ESBL E. coli (left and right panel respectively) and their derived sources. Panels show Autumn (September-November), Spring (March-May), Summer (June-August) and Winter (December-February) over 2019-2022.

5.4 Discussion

5.4.1 Use of SWAT model to predict daily concentrations and source loadings

This study shows that the SWAT model can be used to simultaneously estimate *E. coli* and ESBL *E. coli* concentrations in a river system. Through comparing observed and simulated measurements of *E. coli* and ESBL *E. coli* at individual sites graphically, consistency was seen despite the R^2 values sometimes being <0.5. A visual alignment between observed and simulated bacteria data has been observed in previous SWAT studies, where typical indices such as the NSE, which are used to assess model predictive ability against daily observed values, generally show a low ability for the model to match observed data (Baffaut and Sadeghi 2010; Coffey et al. 2013; Frey et al. 2013). Part of the reason for lower quality model performance (i.e., measured through NSE) to predict daily *E. coli* and ESBL *E. coli* is due to the high

temporal variability in observed bacteria measurements, where observations were based on grab samples that will not reflect daily fluctuations (Traister and Anisfeld 2006). In addition, there are likely other processes in the environment that are not being represented in the model which may impact bacteria concentrations in this area, such as the underlying ecological processes that take place when wastewater effluent microbial communities are introduced to the native bacteria instream, co-resistance from metals, or the role of nutrient availability, where the latter could be explored in further iterations (Makarewicz et al. 2015). Whilst river water residence times can be very short, the response of bacteria to stressors such as nutrient limitation can be significant, such as reduced growth rates and selection of faster growing species (Findlay 2010; Romero et al. 2019). Exposure to metals can also quickly result in co-selection, where studies have indicated this could occur in around 6 hours and be maintained for a further 20 hours in absence of selective pressure (Zhang et al. 2018). This could particularly be a factor for river biofilms, which have a longer exposure period (Romero et al. 2019).

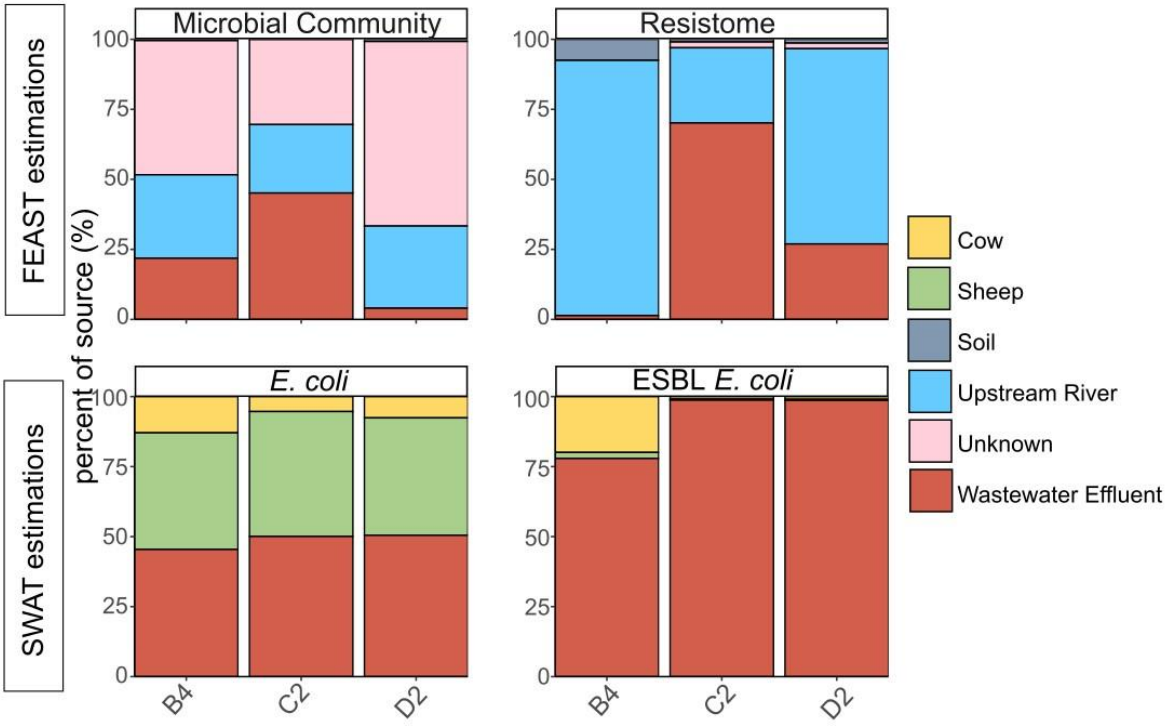


Figure 5-8 Percentage breakdown of sources contributing to E. coli and ESBL E. coli concentrations at sink sites B4, C2 and D2 using the FEAST model from Chapter 4.3.2 (top panel) and the SWAT model predictions (bottom panel), over the observed sampling timeframe

The model was also assessed relative to its ability to determine whether it can predict source loadings. The loadings of *E. coli* and ESBL *E. coli* in the catchment on average mostly aligned with observations from Chapter 4, where diffuse input was dominant in upstream sites, and bacteria sourced from humans was found immediately downstream of wastewater effluent. This and previous studies have also found bacteria concentrations predicted by SWAT directly increased or decreased relative to a change in point source input loads (Parajuli 2007; Frey et al. 2013). Whilst diffuse input of *E.* coli was observed, ESBL *E. coli* was found to be predominantly from wastewater effluents. This is likely due to the low starting concentration of ESBL *E. coli* in manure, and a lack of direct input into the river system. Further, there is no direct defecation from animals included in the model, which would likely lead to different concentrations compared to runoff. Future work could also better characterise the ESBL *E.* coli concentration per gram of sheep and cattle manure with complementary field studies, to allow more accurate estimations.

As this modelling study primarily aimed to test specific source-inputs characterised from Chapter 4, there is a lack of other source inputs in the model, which include wildlife, on-site septic tank systems, birds and other farm or domestic animals. A study in Colorado showed wildlife contributed nearly 50% of bacterial inputs (Jeong et al. 2019). This may occur in the upstream part of the Coquet River, where the Cheviot hills have wild animals such as Roe Deer or Goats. On-site septic systems are also common in rural areas and have been shown to have high concentrations of faecal coliforms and *E. coli* (Richards et al. 2016). In addition, the role of sediment resuspension is not explored in this model. Previous AMR studies have found the role of sediment resuspension in a river to be important in the dissemination of AMR (Knapp et al. 2012; Abia et al. 2016; Heß et al. 2018) and this has previously been modelled with SWAT on *E. coli* concentrations (Kim et al. 2009, 2010). Reducing the error and improving the predictive ability of future model iterations can be achieved through combined laboratory/field experiments to understand overland, sediment and groundwater transport of bacteria (Cho et al. 2016).

5.4.2 Application to assist in AMR monitoring and further work

There is an increased interest in environmental AMR surveillance, but a lack of direction related to the aim of the monitoring (Bengtsson-Palme et al. 2023). While the information derived from the model cannot replace the wealth of information

gained from on-site analysis or be able to replicate the complexity of the river catchment system, modelling can be used to identify relevant AMR settings for monitoring, understand temporal variability and conduct scenario experiments (Jampani et al. 2023). Following improvement of source characterisation and further calibration with observed data, this model could be used to guide further targeted environmental monitoring.

One of the main advantages of using a SWAT model is that it is an open-source tool and has a wide range of studies, tutorials, and information readily available. When designing a model to be used to guide further monitoring of AMR, many factors need to be considered, including the user-friendliness of the model, as this tends to be what the end-users of the model (i.e., policy makers, scientists, river catchment scientists) prefer (Cho et al. 2016). The SWAT model has been shown to be much more user friendly compared to similar software such as HSPF (Gebremariam et al. 2014; Iqbal and Hofstra 2019).

Further, there is an opportunity for this type of modelling to be integrated into current standardised monitoring programmes, where ESBL *E. coli* is already monitored as part of the WHO Tricycle surveillance (WHO 2021b). This would also improve and calibrate this SWAT model, as previous SWAT AMR models, including the one in this study, have been limited by lack of observed data (House 2020; Thilakarathne 2020). *E. coli* is also inexpensive and technically easy to monitor and can be used as an indicator for monitoring environmental AMR (Anjum et al. 2021), which could be an alternative approach in less-contaminated areas such as the Coquet River, where ESBL *E. coli* levels are low in abundance and, therefore, more challenging to detect.

Overall, to improve the model's ability to measure daily/monthly/yearly *E. coli* and ESBL *E. coli*, further observed data is essential. Further, field data is essential to reduce uncertainty in the model, which may improve the accuracy of models so applications can extend to scenario testing, for the purposes of planning mitigation or management approaches (Knight et al. 2019; Bengtsson-Palme et al. 2021; Jampani et al. 2023). For example, combined sewer overflows (CSOs) have recently gained media and public attention and can be harmful to water users (Giakoumis and Voulvoulis 2023). Using modelling tools can enable further understanding of bacteria dissemination as well as AMR and would be used to identify areas which would be at

risk of exposure. SWAT can also be used to test different climate change scenarios (Zhang et al. 2007; Ficklin et al. 2009; Githui et al. 2009) and its application may be useful to address the increased interest in understanding the link between climate change and AMR (Magnano San Lio et al. 2023). There is also the potential to extend this modelling to other AMR bacteria of interest. For example, the carbapenemase-producing *E. coli* has received research interest and has been previously represented in environmental AMR modelling studies (Niebaum et al. 2023).

Whilst SWAT is a promising tool for application to inform in situ monitoring for less experienced end users, it should be noted that environmental AMR modelling is as a field still in its infancy, and depending on the application alternative modelling approaches should continue to be explored. For example, a distributed modelling approach, such as GREAT-ER (Van Heijnsbergen et al. 2022; Niebaum et al. 2023), may be more appropriate when capturing fine spatial resolution and complexity is important (Petrucci and Bonhomme 2014).

5.5 Conclusions

This study successfully developed a SWAT model to simulate the spatiotemporal distribution and dynamics of *E. coli* and ESBL *E. coli* in the Coquet River. Overall, the study found alignment between the simulated and observed *E. coli* and ESBL *E. coli* generated from the model. However, improvements are required in the model's ability to predict daily loadings, which require collecting more observed data and incorporating more bacteria sources into the model. With SWAT being a user-friendly and open-source tool, there is potential for this work to be expanded and the model to be used widely. However as environmental AMR modelling is a field still in its infancy, a variety of modelling approaches should continue to be explored depending on the application.

Chapter 6 Conclusions and Recommendations

6.1 Conclusions

The recent COVID-19 pandemic and the global economic and health cost provided an insight into the potential consequences of AMR. Environmental settings are recognised as a source, sink and driver for AMR and tackling the problem necessitates adopting a One Health approach. River catchment studies have been frequently employed as a method to monitor environmental AMR, as rivers act as conduits for anthropogenic contamination. Current river catchment monitoring is limited to local research studies, often in more contaminated catchments. This includes the UK, where environmental AMR is studied within the context of bathing waters, or contaminated catchments such as the Thames catchment. Analysis of rural catchments may be particularly important for the UK, where there is a substantial level of agricultural activity, and where half of antimicrobials consumed are from livestock in agricultural settings (Manyi-Loh et al. 2018). Analysis of these areas in the UK offers the opportunity to identify the more nebulous drivers of AMR, in rural settings without clear sources. Enhanced surveillance and AMR characterisation can also help to inform tools for monitoring resistance, including predictive models.

This thesis aimed to understand the importance of various drivers and sources of rural AMR in the UK, and the best tools and approaches for surveillance. Firstly, two catchments in the UK with distinct hydrometeorological conditions were used to increase understanding of how external drivers impacted AMR. The FEAST community sourcetracking model was then applied to a more focused study in the Coquet catchment to determine the prevalence and persistence of AMR through point source inputs. Finally, a SWAT model was developed for the Coquet catchment to simulate *E. coli* and ESBL *E. coli* to determine whether this could be used to complement on-site monitoring.

The work presented in Chapters 3-5 fulfilled the following tasks.

1. **Perform a thorough spatial assessment of AMR and microbial communities in the Coquet (Northumberland) and Eden (Cumbria) Rivers.**

River catchment surveillance provides the opportunity to analyse environments surrounding and within the river environment, where rivers act as conduits for anthropogenic contamination. There is a lack of comprehensive studies assessing AMR and microbial communities in environments where the sources of AMR are more nebulous. Such studies may be crucial in uncovering the drivers of AMR dissemination, reveal areas that require on-site monitoring, and may be used to better understand environmental AMR in more natural, less impacted environments (Bengtsson-Palme et al. 2023).

To address this research gap, a comprehensive spatial assessment was conducted in river water in the Coquet and Eden River catchments, as well as a more focused and thorough spatial and seasonal assessment of the Coquet catchment, river water, soil and wastewater effluent. This spatial assessment included sampling in different locations along the course of the river, as well as assessing microbial communities, 96 ARGs and MGEs, water quality indicators (i.e. pH, dissolved oxygen and physiochemical parameters) and microbial source tracking markers for human and ruminant *Bacteroidetes*. Both the Coquet and Eden were characterised by a high input of ruminant *Bacteroidetes* and a lower input of human *Bacteroidetes*. Further assessment and community source tracking analysis of the Coquet catchment revealed most inputs were from unknown sources, indicating that these agricultural catchments have microbial communities and resistomes, which are primarily comprised from diffuse input.

Generally, in both the Coquet and Eden catchments, there was not a consistent upstream to downstream trend in decreasing water quality, and increase in AMR, which has been observed in previous studies in more contaminated catchments (Ho et al. 2021; Ott et al. 2021a). Throughout the river in general, AMR was not correlated with water quality parameters, indicating AMR in these sites was less associated with the decrease in water quality associated with anthropogenic input. However, comparison of the Coquet and Eden catchments demonstrated noteworthy resistance in rural catchments, including clinically relevant bla_{KPC} in the Eden catchment, likely due to faecal inputs.

This comprehensive spatial assessment of these catchments demonstrates the potential insights from increased surveillance in less-contaminated sites. In addition,

the development of an AMR model using open-source software can be used to help guide this further monitoring work and identify sites of potential interest.

2. **Compare the microbial and hydrological signatures of the Coquet and Eden River catchments**

Previous studies which incorporate river catchment analysis as the approach for monitoring environmental AMR are limited to one river catchment. This means external drivers for environmental AMR such as local climate and hydrology are difficult to identify, without the context of additional sites. Comparison across different studies is also challenging due to the differences in methodology for characterising microbiomes and resistomes. This problem was addressed through a comparison of the Coquet and Eden catchments, which both had similar land-uses but different hydrometeorological characteristics.

A comparison of the hydrology showed the Eden was a flashier catchment, with more extensive periods of soil saturation. This enhances the runoff in this catchment. Further analysis of the microbiome showed the Eden had significantly more abundant microbes associated with soils, animal faeces and wastewater, whilst the Coquet had microbes associated with less impacted environments (Wilcoxon test, $p<0.01$). Analysis of the resistome also showed a higher abundance of antibiotic resistance genes (ARGs) in the Eden, as well as significantly higher diversity (Kruskal Wallis, $p<0.05$). ARGs that were unique to the Eden catchment included bla_{KPC}, which is clinically significant, particularly in NW England. ARGs, such as *tetM* which is indicative of pollution caused by livestock were also found to be significantly more abundant in the Eden catchment.

Network analysis also revealed that the order *Bacteroidales* highly correlated with ARGs, however only the Eden catchment showed a strong association with human and ruminant *Bacteroidetes* microbial source tracking markers, indicating an increased association with human/agricultural pollution related resistance.

The results from this comparison revealed the Eden had higher AMR abundance and diversity compared to the Coquet, likely due to the elevated river flows, rainfall and runoff. This association between resistance and local climate would not have been as apparent without the context gained from comparing the two catchments together. With increased runoff and flashiness associated with increased resistance, this must

be considered where more dynamic rainfall events are expected with climate change. The findings from this study emphasise the need for understanding AMR in different environments to provide insight into typical abundance ranges, which could be used to inform on the relative risk to human and animal health (Bengtsson-Palme et al. 2023).

3. **Focusing on the Coquet catchment, use FEAST as an approach for community source tracking and determine the most important sources in the catchment.**

In environments such as the rural UK, rivers are less likely to receive contamination from urbanisation, and more likely to receive inputs from small wastewater treatment works and runoff from agriculture, which both have the potential to enhance the resistome. The results from Chapter 3 indicated the potential influence of non-point source inputs in rural catchments, however to what extent point sources in rural environments is influencing the microbiome and resistome was unclear. Community source tracking is a useful tool to determine the contribution of point source samples into the structure of the microbial community. Recently, the FEAST method has been found to be a good approach for environmental water bodies where the faecal input is low (Xu et al. 2022).

For the Coquet River, generally point source contribution such as wastewater treatment plant effluent had a low influence on the microbial community structure and resistome of downstream sites (e.g., Cluster B and D). Therefore, it is likely that for the most part, resistance in this environment is driven from the input of diffuse sources or the native microbial communities. The exception to this was one site immediately downstream wastewater effluent input from the point source site CS1 (Cluster C: C2), where diffuse sources and native microbial communities were overridden by the communities of the wastewater effluent. Interestingly, whilst this site had identical sampling conditions to another site further down the catchment (Cluster D: D2), there was only an increase in resistance genes associated with faecal contamination in C2. Higher flow rate further down the catchment could result in reducing the coalescence duration and reduce the potential for horizontal gene transfer, where conjugation processes require direct cell to cell contact (Michaelis and Grohmann 2023) and the transfer of genetic material can take several minutes

to hours (Raleigh and Low 2013). In addition, the placement of the wastewater effluent pipe in the centre of the river as oppose the side of the river, may result in temporary input of faecal contamination being attenuated more quickly, where the river flow is faster and more laminar. Whilst the FEAST model showed that cluster C was uniquely characterised by a high contribution of human sources which impacted the downstream resistome, the SWAT model simulations indicated a greater impact of wastewater effluent on downstream resistomes in cluster B and D as well. It should be considered however that the SWAT model does not separate bacteria input from upstream river water, therefore it is likely that the bacteria input from upstream sources is contributing to the high percentage of wastewater effluent source contribution downstream.

The FEAST analysis was also completed for resistance gene data, which was obtained through HT-qPCR. Previous studies have only used FEAST with resistance gene data obtained through metagenomics (Zhang et al. 2022; Chen et al. 2023). Metagenomics requires more computational knowledge and has a lower detection limit than HT-qPCR (Waseem et al. 2019). In addition, HT-qPCR can be completed from the extraction to analysis phase commercially, through companies like Resistomap Oy. Whilst there are benefits to obtaining resistome data through HT-qPCR, this study indicated a large discrepancy between the unknown source contribution for microbiome data compared to resistome data. The unknown source contribution for resistome data was measured as very low, giving the appearance of a low contribution of unknown or diffuse sources to the downstream resistome. However, it is more likely that measured sources (i.e. wastewater) contributed highly to the abundance of the measured 96 ARGs in this study, whilst the unknown or diffuse sources were contributing to other ARGs that were not measured in the HT-qPCR assay. Further studies should repeat FEAST analysis for a greater number of ARGs/MGEs to obtain a richer number of contributing sources.

4. **Identify the sources and drivers for AMR in rural river catchments in the UK.**

The environmental AMR in the Eden and Coquet catchments was interpreted to be primarily derived from diffuse agricultural sources. In Chapter 3, the comparison between physiochemical markers such as dissolved oxygen and AMR indicators

such as ARG abundance, showed there were no significant correlations, which is contradictory to previous studies in more contaminated catchments (Ho et al. 2021; Ott et al. 2021a). This indicates that in rural environments which are less impacted by anthropogenic activity and direct point sources, diffuse sources may be more important drivers of environmental AMR. This was supported through the work in Chapter 4, where the FEAST approach and volcano plot analysis showed that in general the point sources in the Coquet catchment did not shape the downstream resistome.

Whilst understanding the sources of AMR in a river system is important, understanding the factors which dictate the survival of non-native bacteria and ARGs is highly important to best implement approaches for mitigation. In Chapter 3, it was found that difference in hydrometeorological conditions in the Eden and the Coquet led to differences in the abundance and diversity of river resistomes. In Chapter 4, the impact of water quality on microbial composition and AMR was explored through investigating nutrient limitation, represented through N:P ratios. The N:P ratios were calculated at each river site, which indicated the Coquet was a predominately phosphorous limited environment, except for immediately downstream wastewater effluent discharges, but then returned to phosphorous limited conditions. Mirroring this pattern, following the temporary influx of ARGs into a river system, resistance was generally not maintained down the course of the river, possibly due to the host microbes of non-native bacteria being not well adapted to the nutrient limited system. The nutrient limited conditions and lack of selective pressure therefore meant that point-source input of wastewater effluent was not maintained down the course of the river. The diffuse input of AMR from the Coquet River was attempted to be characterised through a SWAT model, however this showed the diffuse input of ESBL *E.* coli (which was selected to model AMR in the river system), was quite low. Further improvements and iterations of this model could be used to understand the diffuse inputs of AMR into a river system and assess how different drivers can dictate persistence. The results of this work indicate that maintaining natural water quality, effective wastewater management and reducing runoff from agricultural sites is important to reduce ARG dissemination and persistence in a rural catchment.

5. **Demonstrate the ability of the open-source SWAT modelling tool to predict AMR hotspots in a river catchment.**

The catchment scale analysis in Chapters 3 and 4 provided an insight into AMR in rural environments in the UK. However continued monitoring of this scale and in different environments is challenging, due to the cost and time involved. In addition, due to the wide range of methodologies to measure AMR, as well as the various sampling locations, knowing what and where to monitor is difficult (Bengtsson-Palme et al. 2023). Watershed models such as the Soil and Water Assessment tool (SWAT) can be used to simulate the point and diffuse inputs of bacteria into the river system, and Chapter 5 demonstrates the first attempt at using this tool to simulate both *E. coli* and ESBL *E. coli* simultaneously. When compared to observed data, the SWAT model simulated *E. coli* and ESBL *E. coli* values which visually aligned with observed data. In terms of predicting daily loadings, the coefficient of determination values (R^2) were mostly <0.5, indicating the model could not be used to reliably predict daily loadings. However, this is likely due to the variability of the observed values, which were based on grab samples. When predicting source loadings of *E. coli*, the model could characterise diffuse and point source inputs, that align with FEAST predictions from Chapter 4. However, when source loadings of ESBL *E. coli* were characterised, this was mostly found to be sourced from wastewater, likely due to the low input values in sheep and cows. Further characterisation of sources of bacteria, including wildlife could improve these estimations. In addition, further on-site sampling and quantification of *E. coli* and ESBL *E. coli* could be used to calibrate the model, where in this study, the lack of observed bacterial data made robust calibration and validation methods challenging and a longer term bacterial data set would allow for more rigorous calibration and validation techniques to be employed. Ideally, daily concentrations of *E. coli* and ESBL *E. coli* would be gathered over 2-3 sites over 1-2 years, which could capture seasonal variations, daily fluctuations, reveal long term trends and reduce uncertainty.

One of the main advantages of using the SWAT model is it is an open-source tool which is user friendly, and has a wide range of studies, tutorials and information readily available. Therefore, the user-friendliness of SWAT makes it an attractive option for less experienced end-users to inform on-site monitoring. The results from this thesis will hopefully initiate further work on refining the SWAT model to be used in the field of environmental AMR surveillance.

6. Suggest tools and approaches to continue monitoring environmental AMR in rural river catchments

Surveillance of environmental AMR in rural settings is not common and identifying the tools, locations and techniques to conduct such surveillance is ongoing. Part of this thesis aimed to utilise the information gathered from the studies to suggest approaches for continued monitoring of AMR in rural settings, which have been highly understudied in the UK.

From the results in Chapter 3, understanding the spatial variability of AMR is important to provide context for the scale of the problem and identify alternative drivers for AMR, including heavy rainfall and runoff. Therefore, the first suggestion (1) is to conduct surveillance across a wider spatial scale rather than isolated catchment studies. The thesis also utilised more accurate and quantitative approaches for analysing community sequencing data (i.e. Quantitative Microbial Profiling (QMP); Vandeputte et al. 2017; Ott et al. 2021b), which provided more insight into the different environmental conditions in the Eden and Coquet catchments through the different taxa abundances. In line with recommendations by Ott et al. (2021b), this thesis also suggests (2) that the QMP approach is employed for analysing sequencing data in further environmental studies.

The findings from Chapter 4 suggested that the temporary influx of AMR from a point source dissipates through the river, likely due to the environmental conditions such as the natural nutrient limited conditions. Within the context of AMR monitoring this thesis suggests (3) this work would be expanded to other catchments to identify the extent to which nutrient availability impacts the persistence of AMR. Such findings would be impactful for informing monitoring approaches, where the stoichiometric ratio of nutrients could potentially be used as a ‘proxy marker’ for environmental AMR in rural settings and/or provide insight into the environmental factors dictating the survival of AMR in a river.

The final suggestion (4) is that continued surveillance in rural settings should harmonise with existing surveillance approaches to enable comparison between environments. As there are currently no standardised approaches for genomic monitoring, initiating rural environmental AMR monitoring in-line with the WHO GLASS or the WHO Tricycle Protocol (WHO 2021b, a) would ensure comparability

amongst studies and allow identification of areas where more in depth monitoring may be required. Regular collection of this data is also necessary to calibrate AMR models, such as the model developed in Chapter 5, where there is often insufficient observed data to complete this step.

6.2 Application of findings to map AMR risk in river catchments

The findings of this thesis addressed research questions related to the characterisation of AMR in rural river catchments and provided tools which could be utilised in routine monitoring. The wider impact of this work is how the findings and tools proposed in this research could be used to inform environmental AMR risk within a One Health context.

Based on the findings of this study, for an understanding of environmental AMR, we require a large spatial distribution of surveillance to target river catchments of different sizes, hydrometeorological conditions and climates. The results of this thesis indicated catchments with high levels of rainfall and run-off contributed to enhanced environmental AMR, which could have implications for AMR risk in such locations. AMR risk in the environment would also be likely elevated during and after hydrological events, including storms and floods. Understanding external drivers would be crucial for estimating the risk associated with climate change and environmental AMR.

Despite the focus on rural sites in this thesis, the importance of highly urbanised catchments for enhancing environmental AMR cannot be discounted. Urbanised sites are particularly linked to high levels of clinically relevant environmental AMR, due to the various anthropogenic inputs. Therefore, as a priority for surveillance, we should aim to capture different land-uses, including less impacted/pseudo-pristine sites, rural sites and urbanised catchments. This would also help in mapping the levels of AMR in contaminated sites, against the context of the pristine/less impacted sites, thereby separating intrinsic resistance from anthropogenic resistance, as well as the types of resistance found in rural and urban sites.

This thesis also demonstrated the importance of the application of modelling tools to support on-site sampling. Whilst modelling data cannot replace on-site data, the development of robust models can permit simulations of scenarios including

changing climate, land-use and weather events to be tested. This could help to map AMR risk in catchments and reduce the need for expensive on-site monitoring.

Finally, to fully assess environmental AMR risk within a One Health context, we should look to identify areas of potential human and animal exposure. The includes agricultural sites, human and animal drinking sites, fisheries and official and unofficial bathing water sites. This could be further linked to clinical data and allow for an assessment of environmental AMR risk. Exposure routes may also differ in rural and urbanised catchments, further supporting the requirement for monitoring of AMR in both land-uses.

6.3 Recommendations for future work

1. *Further AMR surveillance of rural catchments:* Rural catchments which are not influenced by urban contamination are highly understudied. This study shows the importance of non-point sources and nutrient availability as drivers for AMR in rural locations. Further work should continue to measure AMR in rural catchments, targeting agricultural catchments and pristine catchments to understand background AMR and the influence of animals and crop production. Understanding environmental AMR in these land-uses can also provide useful context when assessing highly urbanised catchments. Therefore, future research should aim to diversify the types and land-uses of river catchments that are targeted for AMR surveillance.
2. *Improving the temporal accuracy of this study:* Due to the impact of the COVID-19 pandemic, this study was unable to capture the temporal accuracy that may have been captured if field studies were less restricted. Future work should aim to sample more frequently (e.g. daily or weekly) over a longer sampling duration (e.g. several months to a year), in order to limit any temporal variability and to gather an accurate impression of AMR in the river system.
3. *Impact of sediment resuspension:* The impact of sediment resuspension (i.e. where sediment from the river-bed gets resuspended and entrained into the river at high flow), in the river system could be an important source of AMR, particularly downstream of point sources where sediment is subject to more contamination. Sediment resuspension however could not be investigated in

this study, as it was not possible to incorporate the extra time to access the river safely to sample the bottom sediment into the sampling timeframe. This could however be achieved through a more focused study of one site.

4. *Regular E. coli and ESBL E. coli data is needed to calibrate models*: The long-term daily flow data from the Coquet River made it possible to easily validate and calibrate the SWAT model for estimating flow rate. Regular determination of *E. coli* and ESBL *E. coli* at a site on a daily/semi-daily basis over a long duration in the Coquet would make it possible to improve the accuracy of the SWAT model. Determination of *E. coli* and ESBL *E. coli* is technically easy to do, and validation of this model against observed data would allow further experiments to be conducted.
5. *Further work on characterising the SWAT sources:* The SWAT model in this study ran simulations based on a few point sources in the catchment, and diffuse sources from cattle and sheep. However further characterisation of different AMR sources, including wildlife and septic tank systems could improve the accuracy of the model, and allow further investigation of the impact of different sources into instream AMR. As well as this, further field monitoring data and determination of *E. coli* and ESBL *E. coli* concentrations in sheep and cattle manure would improve model predictions.

References

Abbaspour KC, Vaghefi SA, Srinivasan R (2017) A guideline for successful calibration and uncertainty analysis for soil and water assessment: a review of papers from the 2016 international SWAT conference. Water (Basel) 10:6. https://doi.org/10.3390/w10010006

Abdulsattar BO, Abdulsattar JO, Rasool KH, et al (2020) Study of antimicrobial resistance pattern of Escherichia coli and Klebsiella strains and multivariate analysis for water quality assessment of Tigris river, Baghdad, Iraq. Nature Environment and Pollution Technology 19:1327–1334. https://doi.org/10.46488/NEPT.2020.v19i03.050

Abia ALK, Ubomba-Jaswa E, Genthe B, Momba MNB (2016) Quantitative microbial risk assessment (QMRA) shows increased public health risk associated with exposure to river water under conditions of riverbed sediment resuspension. Science of the Total Environment 566:1143–1151. https://doi.org/10.1016/j.scitotenv.2016.05.155

Abushaheen MA, Fatani AJ, Alosaimi M, et al (2020) Antimicrobial resistance, mechanisms and its clinical significance. Disease-a-Month 66:100971. https://doi.org/10.1016/j.disamonth.2020.100971

Acharya K, Khanal S, Pantha K, et al (2019) A comparative assessment of conventional and molecular methods, including MinION nanopore sequencing, for surveying water quality. Sci Rep 9:1–11. https://doi.org/10.1038/s41598-019-51997-x

AgCensus E (2020) EDINA AgCensus 2020. https://agcensus.edina.ac.uk/. Accessed 7 Nov 2023

Ahammad ZS, Sreekrishnan TR, Hands CL, et al (2014) Increased waterborne bla NDM-1 resistance gene abundances associated with seasonal human pilgrimages to the Upper Ganges River. Environ Sci Technol 48:3014–3020. https://doi.org/10.1021/es405348h

Ahmad I, Malak HA, Abulreesh HH (2021) Environmental antimicrobial resistance and its drivers: a potential threat to public health. J Glob Antimicrob Resist 27:101–111. https://doi.org/10.1016/j.jgar.2021.08.001

Ahmed W, Zhang Q, Lobos A, et al (2018) Precipitation influences pathogenic bacteria and antibiotic resistance gene abundance in storm drain outfalls

in coastal sub-tropical waters. Environ Int 116:308–318. https://doi.org/10.1016/j.envint.2018.04.005

Alaoui Mdarhri H, Benmessaoud R, Yacoubi H, et al (2022) Alternatives Therapeutic Approaches to Conventional Antibiotics: Advantages, Limitations and Potential Application in Medicine. Antibiotics 11:. https://doi.org/10.3390/antibiotics11121826

Alavez-Ramírez J, Castellanos JRA, Esteva L, et al (2007) Within-host population dynamics of antibiotic-resistant M. tuberculosis. Mathematical Medicine and Biology 24:35–56 https://doi.org/10.7554/eLife.61805

Alberdi A, Gilbert MTP (2019a) A guide to the application of Hill numbers to DNA-based diversity analyses. Mol Ecol Resour 19:804–817. https://doi.org/10.1111/1755-0998.13014

Alberdi A, Gilbert MTP (2019b) hilldiv: an R package for the integral analysis of diversity based on Hill numbers. Biorxiv 545665 https://doi.org/10.7554/eLife.61805

Alcock BP, Raphenya AR, Lau TTY, et al (2020) CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517–D525. https://doi.org/10.1093/nar/gkz935

Almakki A, Jumas-Bilak E, Marchandin H, Licznar-Fajardo P (2019) Antibiotic resistance in urban runoff. Science of the Total Environment 667:64–76. https://doi.org/10.1016/j.scitotenv.2019.02.183

Amos GC, Gozzard E, Carter CE, et al (2015) Validated predictive modelling of the environmental resistome. ISME Journal 9:1467–1476. https://doi.org/10.1038/ismej.2014.237

Amos GCA, Hawkey PM, Gaze WH, Wellington EM (2014) Waste water effluent contributes to the dissemination of CTX-M-15 in the natural environment. Journal of Antimicrobial Chemotherapy 69:1785–1791. https://doi.org/10.1093/jac/dku079

An U, Shenhav L, Olson CA, et al (2022) STENSL: Microbial Source Tracking with ENvironment SeLection. mSystems 7:. https://doi.org/10.1128/msystems.00995-21

Anjum MF, Schmitt H, Börjesson S, et al (2021) The potential of using E. coli as an indicator for the surveillance of antimicrobial resistance (AMR) in the environment. Curr Opin Microbiol 64:152–158. https://doi.org/10.1016/j.mib.2021.09.011

Archundia D, Boithias L, Duwig C, et al (2018) Environmental fate and ecotoxicological risk of the antibiotic sulfamethoxazole across the Katari catchment (Bolivian Altiplano): Application of the GREAT-ER model. Science of the Total Environment 622–623:1046–1055. https://doi.org/10.1016/j.scitotenv.2017.12.026

Arnold BJ, Huang I-T, Hanage WP (2022) Horizontal gene transfer and adaptive evolution in bacteria. Nat Rev Microbiol 20:206–218. https://doi.org/10.1038/ s41579-021-00650-4

Arzanlou M, Chai WC, Venter H (2017) Intrinsic, adaptive and acquired antimicrobial resistance in Gram-negative bacteria. Essays Biochem. https://doi.org/10.1042/EBC20160063

Ashiru-Oredope D, Cunningham N, Casale E, et al (2023) Reporting England's progress towards the ambitions in the UK action plan for antimicrobial resistance: the English surveillance programme for antimicrobial utilisation and resistance (ESPAUR). Journal of Antimicrobial Chemotherapy dkad248. https://doi.org/10.1093/jac/dkad248

Ashiru-Oredope D, Hopkins S, ESPAUR Oversight Group Writing Committee . (2015) English surveillance programme for antimicrobial utilisation and resistance (ESPAUR) from 2010 to 2014: 2015 Report. Public Health England Publications 14–15

Attal M (2017) Structural Analysis of Rocks and Regions 2017 - Topographic analysis. 1–22

Avis EL, Harrop SE (1983) Soils of England and Wales. Sheet 1. Northern England. Scale 1: 250 000. Ordnance Survey

Baffaut C, Sadeghi A (2010) Bacteria modeling with SWAT for assessment and remediation studies: A review. Trans ASABE 53:1585–1594. https://doi.org/10.13031/2013.34907

Baker-Austin C, Wright MS, Stepanauskas R, McArthur J V. (2006) Co-selection of antibiotic and metal resistance. Trends Microbiol 14:176–182 https://doi.org/10.1016/j.tim.2006.02.006

Ballesté E, Demeter K, Masterson B, et al (2020) Implementation and integration of microbial source tracking in a river watershed monitoring plan. Science of The Total Environment 736:139573. https://doi.org/10.1016/j.scitotenv.2020.139573

Baral D, Dvorak BI, Admiraal D, et al (2018a) Tracking the Sources of Antibiotic Resistance Genes in an Urban Stream during Wet Weather using Shotgun Metagenomic Analyses. Environ Sci Technol 52:9033–9044. https://doi.org/10.1021/acs.est.8b01219

Baral D, Speicher A, Dvorak B, et al (2018b) Quantifying the relative contributions of environmental sources to the microbial community in an urban stream under dry and wet weather conditions. Appl Environ Microbiol 84:e00896-18. https://doi.org/10.1128/aem.00896-18

Bartosch S, Fite A, Macfarlane GT, McMurdo MET (2004) Characterization of bacterial communities in feces from healthy elderly volunteers and hospitalized elderly patients by using real-time PCR and effects of antibiotic treatment on the fecal microbiota. Appl Environ Microbiol 70:3575–3581. https://doi.org/10.1128/AEM.70.6.3575-3581.2004

Bastian M, Heymann S, Jacomy M (2009) Gephi: An open source software for exploring and manipulating networks. BT - International AAAI Conference on Weblogs and Social. In: International AAAI Conference on Weblogs and Social Media https://doi.org/10.1609/icwsm.v3i1.13937

Batjes NH (1997) A world dataset of derived soil properties by FAO-UNESCO soil unit for global modelling. Soil Use Manag 13:9–16. https://doi.org/10.1111/j.1475-2743.1997.tb00550.x

Bengtsson-Palme J, Abramova A, Berendonk TU, et al (2023) Towards monitoring of antimicrobial resistance in the environment: For what reasons, how to implement it, and what are the data needs? Environ Int 108089. https://doi.org/10.1016/j.envint.2023.108089

Bengtsson-Palme J, Boulund F, Fick J, et al (2014) Shotgun metagenomics reveals a wide array of antibiotic resistance genes and mobile elements in a polluted lake in India. Front Microbiol 5:1–14. https://doi.org/10.3389/fmicb.2014.00648

Bengtsson-Palme J, Jonsson V, Heß S (2021) What is the role of the environment in the emergence of novel antibiotic resistance genes? A modeling approach. Environ Sci Technol 55:15734–15743. https://doi.org/10.1021/acs.est.1c02977

Bengtsson-Palme J, Kristiansson E, Larsson DGJ (2018) Environmental factors influencing the development and spread of antibiotic resistance. FEMS Microbiol Rev 42:68–80. https://doi.org/10.1093/femsre/fux053

Benjamini Y, Hochberg Y (1995) Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B. 1995:p. 289--300. Journal of the Royal Statistical Society B. https://doi.org/10.2307/2346101

Bennett PM (2008) Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. Br J Pharmacol 153:S347–S357. https://doi.org/10.1038/sj.bjp.0707607

Bernhard AE, Field KG (2000a) A PCR assay to discriminate human and ruminant feces on the basis of host differences in Bacteroides-Prevotella genes encoding 16S rRNA. Appl Environ Microbiol 66:4571–4574. https://doi.org/10.1128/AEM.66.10.4571-4574.2000

Bernhard AE, Field KG (2000b) Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes. Appl Environ Microbiol 66:1587–1594. https://doi.org/10.1128/AEM.66.4.1587-1594.2000

Bhandari V, Gupta RS (2012) Molecular signatures for the phylum Synergistetes and some of its subclades. Antonie Van Leeuwenhoek 102:517–540. https://doi.org/10.1007/s10482-012-9759-2

Birkegård AC, Halasa T, Toft N, et al (2018) Send more data: A systematic review of mathematical models of antimicrobial resistance 01 Mathematical Sciences 0102 Applied Mathematics. Antimicrob Resist Infect Control 7:1–12. https://doi.org/10.1186/s13756-018-0406-1

Boehm AB, Van De Werfhorst LC, Griffith JF, et al (2013) Performance of forty-one microbial source tracking methods: A twenty-seven lab evaluation study. Water Res 47:6812–6828. https://doi.org/10.1016/j.watres.2012.12.046

Boeken B, Shachak M (2006) Linking community and ecosystem processes: the role of minor species. Ecosystems 9:119–127. https://doi.org/10.1007/s10021-004-0079-x

Brown CM, Mathai PP, Loesekann T, et al (2018) Influence of library composition on SourceTracker predictions for community-based microbial source tracking. Environ Sci Technol 53:60–68. https://doi.org/10.1021/acs.est.8b04707

Brown CT, Hug LA, Thomas BC, et al (2015) Unusual biology across a group comprising more than 15% of domain Bacteria. Nature 523:208–211. https://doi.org/10.1038/nature14486

Burch TR, Newton RJ, Kimbell LK, et al (2022) Targeting current and future threats: recent methodological trends in environmental antimicrobial resistance research and their relationships to risk assessment. Environ Sci (Camb) 8:1787–1802. https://doi.org/10.1039/d2ew00087c

Burns EE, Carter LJ, Kolpin DW, et al (2018) Temporal and spatial variation in pharmaceutical concentrations in an urban river system. Water Res 137:72–85. https://doi.org/10.1016/j.watres.2018.02.066

Calero-Cáceres W, Ye M, Balcázar JL (2019) Bacteriophages as environmental reservoirs of antibiotic resistance. Trends Microbiol 27:570–577. https://doi.org/10.1016/j.tim.2019.02.008

Callahan BJ, McMurdie PJ, Holmes SP (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME Journal 11:2639–2643. https://doi.org/10.1038/ismej.2017.119

Callahan BJ, McMurdie PJ, Rosen MJ, et al (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583. https://doi.org/10.1038/nmeth.3869

Caneschi A, Bardhi A, Barbarossa A, Zaghini A (2023) The Use of Antibiotics and Antimicrobial Resistance in Veterinary Medicine, a Complex Phenomenon: A Narrative Review. Antibiotics 12:487. https://doi.org/10.3390/antibiotics12030487

Carr A, Diener C, Baliga NS, Gibbons SM (2019) Use and abuse of correlation analyses in microbial ecology. ISME Journal 13:2647–2655. https://doi.org/10.1038/s41396-019-0459-z

Carter KM, Lu M, Jiang H, An L (2019) Microbial Community Dissimilarity for Source Tracking with Application in Forensic Studies. bioRxiv 749671. https://doi.org/10.1101/749671

CEH UC for E and H (2023) Flood Estimation Handbook: Corrigenda and FAQs. https://www.ceh.ac.uk/flood-estimation-handbook-corrigenda-and-faqs. Accessed 28 Jul 2023

Chao A, Chiu CH, Jost L (2014) Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through hill numbers. Annu Rev Ecol Evol Syst 45:297–324. https://doi.org/10.1146/annurev-ecolsys-120213-091540

Chaudhari NM, Overholt WA, Figueroa-Gonzalez PA, et al (2021) The economical lifestyle of CPR bacteria in groundwater allows little preference for environmental drivers. Environ Microbiome 16:1–18. https://doi.org/10.1186/s40793-021-00395-w

Chen B, Liang X, Huang X, et al (2013) Differentiating anthropogenic impacts on ARGs in the Pearl River Estuary by using suitable gene indicators. Water Res 47:2811–2820. https://doi.org/10.1016/j.watres.2013.02.042

Chen H, Bai X, Jing L, et al (2019a) Characterization of antibiotic resistance genes in the sediments of an urban river revealed by comparative metagenomics analysis. Science of the Total Environment 653:1513–1521. https://doi.org/10.1016/j.scitotenv.2018.11.052

Chen H, Bai X, Li Y, et al (2019b) Source identification of antibiotic resistance genes in a peri-urban river using novel crAssphage marker genes and metagenomic signatures. Water Res 167:115098. https://doi.org/10.1016/j.watres.2019.115098

Chen H, Li Y, Sun W, et al (2020) Characterization and source identification of antibiotic resistance genes in the sediments of an interconnected river-lake system. Environ Int 137:105538. https://doi.org/10.1016/j.envint.2020.105538

Chen J, Chen H, Liu C, et al (2023) Evaluation of FEAST for metagenomics-based source tracking of antibiotic resistance genes. J Hazard Mater 442:130116. https://doi.org/10.1016/j.jhazmat.2022.130116

Chen K, Pachter L (2005) Bioinformatics for whole-genome shotgun sequencing of microbial communities. PLoS Comput Biol 1:e24. https://doi.org/10.1371/journal.pcbi.0010024

Cho KH, Pachepsky YA, Oliver DM, et al (2016) Modeling fate and transport of fecally-derived microorganisms at the watershed scale: State of the science and future opportunities. Water Res 100:38–56. https://doi.org/10.1016/j.watres.2016.04.064

Chung SW, Ko IH, Kim YK (2008) Effect of reservoir flushing on downstream river water quality. J Environ Manage 86:139–147. https://doi.org/10.1016/j.jenvman.2006.11.031

CITY POPULATION (2023) CITY POPULATION WEBSITE. https://www.citypopulation.de/. Accessed 5 Nov 2023

Cloutier DD, McLellan SL (2017) Distribution and differential survival of traditional and alternative indicators of fecal pollution at freshwater beaches. Appl Environ Microbiol 83:e02881-16. https://doi.org/10.1128/AEM.02881-16

Coe S, Balogun B, Sutherland N (2023) The use of antibiotics on healthy farm animals and antimicrobial resistance. House of Commons Library 1–27

Coffey R, Cummins E, Bhreathnach N, et al (2010a) Development of a pathogen transport model for Irish catchments using SWAT. Agric Water Manag 97:101–111. https://doi.org/10.1016/j.agwat.2009.08.017

Coffey R, Cummins E, O'Flaherty V, Cormican M (2010b) Analysis of the soil and water assessment tool (SWAT) to model Cryptosporidium in surface water sources. Biosyst Eng 106:303–314. https://doi.org/10.1016/j.biosystemseng.2010.04.003

Coffey R, Dorai-Raj S, O'Flaherty V, et al (2013) Modeling of pathogen indicator organisms in a small-scale agricultural catchment using SWAT. Human and Ecological Risk Assessment: An International Journal 19:232–253. https://doi.org/10.1080/10807039.2012.701983

Comeau AM, Douglas GM, Langille MGI (2017) Microbiome helper: a custom and streamlined workflow for microbiome research. mSystems 2:e00127-16. https://doi.org/10.1128/msystems.00127-16

Cox G, Wright GD (2013) Intrinsic antibiotic resistance: mechanisms, origins, challenges and solutions. International Journal of Medical Microbiology 303:287–292. https://doi.org/10.1016/j.ijmm.2013.02.009

Dantas G, Sommer MOA (2014) How to fight back against antibiotic resistance. Am Sci 102:42–51

Dantas G, Sommer MOA (2011) Ecological and Clinical Consequences of Antibiotic Subsistence by Environmental Microbes. Antimicrobial Resistance in the Environment 29–41. https://doi.org/10.1002/9781118156247.ch4

Datta N, Hughes VM (1983) Plasmids of the same Inc groups in Enterobacteria before and after the medical use of antibiotics. Nature 306:616–617. https://doi.org/10.1038/306616a0

David WP (1988) Soil and water conservation planning: policy issues and recommendations. Philippine Journal of Development XV No 1-c:

Davies J, Davies D (2010) Origins and evolution of antibiotic resistance. Microbiologia 12:9–16. https://doi.org/10.1128/mmbr.00016-10

Dcosta VM, King CE, Kalan L, et al (2011) Antibiotic resistance is ancient. Nature 477:457–461. https://doi.org/10.1038/nature10388

DEFRA (2008) Future Water: The Government's Water Strategy for England. HM Government

Derx J, Kılıç HS, Linke R, et al (2023) Probabilistic fecal pollution source profiling and microbial source tracking for an urban river catchment. Science of The Total Environment 857:159533. https://doi.org/10.1016/j.scitotenv.2022.159533

Devarajan N, Laffite A, Graham ND, et al (2015) Accumulation of clinically relevant antibiotic-resistance genes, bacterial load, and metals in freshwater lake sediments in Central Europe. Environ Sci Technol 49:6528–6537. https://doi.org/10.1021/acs.est.5b01031

Dhanji H, Murphy NM, Akhigbe C, et al (2011) Isolation of fluoroquinolone-resistant O25b: H4-ST131 Escherichia coli with CTX-M-14 extended-spectrum β-lactamase from UK river water. Journal of Antimicrobial Chemotherapy 66:512–516. https://doi.org/10.1093/jac/dkq472

Di Cesare A, Eckert EM, Rogora M, Corno G (2017) Rainfall increases the abundance of antibiotic resistance genes within a riverine microbial community. Environmental Pollution 226:473–478 https://doi.org/10.1016/j.envpol.2017.04.036

Dionisi HM, Harms G, Layton AC, et al (2003) Power analysis for real-time PCR quantification of genes in activated sludge and analysis of the variability introduced by DNA extraction. Appl Environ Microbiol 69:6597–6604. https://doi.org/10.1128/AEM.69.11.6597-6604.2003

Do CB, Batzoglou S (2008) What is the expectation maximization algorithm? Nat Biotechnol 26:897–899. https://doi.org/10.1038/nbt1406

Doretto A, Piano E, Larson CE (2020) The River Continuum Concept: lessons from the past and perspectives for the future. Canadian Journal of Fisheries and Aquatic Sciences 77:1853–1864. https://doi.org/10.1139/cjfas-2020-003

Ensz AP, Knapp CW, Graham DW (2003) Influence of autochthonous dissolved organic carbon and nutrient limitation on alachlor biotransformation in aerobic aquatic systems. Environ Sci Technol 37:4157–4162. https://doi.org/10.1021/es0303790

Environment Agency (2008) Microbial Source-Tracking Project SC070003. Environment Agency ISBN: 978-1-84432-965-6

ESDB (2006) The European soil database v.2.0 (vector and attribute data). In: GEO: connexion. https://esdac.jrc.ec.europa.eu/content/european-soil-database-v20-vector-and-attribute-data. Accessed 12 Oct 2023

ESRI (2018) ArcMap 10.6.1. ESRI

Estaki M, Jiang L, Bokulich NA, et al (2020) QIIME 2 Enables Comprehensive End-to-End Analysis of Diverse Microbiome Data and Comparative Studies with Publicly Available Data. Curr Protoc Bioinformatics 70:. https://doi.org/10.1002/cpbi.100

Fair RJ, Tor Y (2014) Antibiotics and bacterial resistance in the 21st century. Perspect Medicin Chem 6:PMC-S14459. https://doi.org/10.4137/pmc.s14459

Feijtel T, Boeije G, Matthies M, et al (1997) Development of a geography-referenced regional exposure assessment tool for European rivers-GREAT-ER contribution to GREAT-ER# 1. Chemosphere 34:2351–2373. https://doi.org/10.1016/S0045-6535(97)00048-9

Ficklin DL, Luo Y, Luedeling E, Zhang M (2009) Climate change sensitivity assessment of a highly agricultural watershed using SWAT. J Hydrol (Amst) 374:16–29. https://doi.org/10.1016/j.jhydrol.2009.05.016

Findlay S (2010) Stream microbial ecology. J North Am Benthol Soc 29:170–181. https://doi.org/10.1899/09-023.1

Fleming A (1945) The Nobel Prize in Physiology or Medicine 1945. HARRISONS PRINCIPLES OF INTERNAL MEDICINE 1:1252

Forsberg KJ, Patel S, Gibson MK, et al (2014) Bacterial phylogeny structures soil resistomes across habitats. Nature 509:612–616. https://doi.org/10.1038/nature13377

Fox KK, Daniel M, Morris G, Holt MS (2000) The use of measured boron concentration data from the GREAT-ER UK validation study (1996–1998) to generate predicted regional boron concentrations. Science of the total environment 251:305–316. https://doi.org/10.1016/S0048-9697(00)00392-2

Frankenberger JR, Daneshvar F (2019) Soil and Water Assessment Tool (SWAT) Instructional Videos. https://purr.purdue.edu/publications/3237/1. Accessed 1 Oct 2023

Frey SK, Topp E, Edge T, et al (2013) Using SWAT, Bacteroidales microbial source tracking markers, and fecal indicator bacteria to predict waterborne pathogen occurrence in an agricultural watershed. Water Res 47:6326–6337. https://doi.org/10.1016/j.watres.2013.08.010

Gebremariam SY, Martin JF, DeMarchi C, et al (2014) A comprehensive approach to evaluating watershed models for predicting river flow regimes

critical to downstream ecosystem services. Environmental modelling & software 61:121–134. https://doi.org/10.1016/j.envsoft.2014.07.004

Ghoul M, Mitri S (2016) The ecology and evolution of microbial competition. Trends Microbiol 24:833–845. https://doi.org/10.1016/j.tim.2016.06.011

Giakoumis T, Voulvoulis N (2023) Combined sewer overflows: relating event duration monitoring data to wastewater systems' capacity in England. Environ Sci (Camb) 9:707–722. https://doi.org/10.1039/d2ew00637e

Gillings MR (2014) Integrons: past, present, and future. Microbiology and molecular biology reviews 78:257–277. https://doi.org/10.1128/mmbr.00056-13

Gillings MR, Gaze WH, Pruden A, et al (2015) Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. ISME Journal 9:1269–1279. https://doi.org/10.1038/ismej.2014.226

Githui F, Gitau W, Mutua F, Bauwens W (2009) Climate change impact on SWAT simulated streamflow in western Kenya. Int J Climatol 29:1823–1834. https://doi.org/10.1002/joc.1828

Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ (2017) Microbiome datasets are compositional: And this is not optional. Front Microbiol 8:1–6. https://doi.org/10.3389/fmicb.2017.02224

Gonggrijp MA, Velthuis AGJ, Heuvelink AE, et al (2023) Prevalence of extended-spectrum and AmpC β-lactamase-producing Escherichia coli in young calves on Dutch dairy farms. J Dairy Sci 106:4257–4265. https://doi.org/10.3168/jds.2022-22362

Gonzalez A, Ronce O, Ferriere R, Hochberg ME (2013) Evolutionary rescue: An emerging focus at the intersection between ecology and evolution. Philosophical Transactions of the Royal Society B: Biological Sciences 368:. https://doi.org/10.1098/rstb.2012.0404

González-Fernández A, Symonds EM, Gallard-Gongora JF, et al (2021) Relationships among microbial indicators of fecal pollution, microbial source tracking markers, and pathogens in Costa Rican coastal waters. Water Res 188:116507. https://doi.org/10.1016/j.watres.2020.116507

Gothwal R, Thatikonda S (2020) Modeling transport of antibiotic resistant bacteria in aquatic environment using stochastic differential equations. Sci Rep 10:. https://doi.org/10.1038/s41598-020-72106-3

Gothwal R, Thatikonda S (2018) Mathematical model for the transport of fluoroquinolone and its resistant bacteria in aquatic environment. Environmental Science and Pollution Research 25:20439–20452. https://doi.org/10.1007/s11356-017-9848-x

GovUK (2019) Tackling antimicrobial resistance 2019–2024 – The UK's five-year national action plan. London: HM Government 2019 101:426–427. https://doi.org/10.1016/j.jhin.2019.02.019

GovUK (2022) Tackling antimicrobial resistance 2019 to 2024: addendum to the UK's 5-year national action plan. January 2019. London: HM Government 2022 1–40

Graham DW, Bergeron G, Bourassa MW, et al (2019) Complexities in understanding antimicrobial resistance across domesticated animal, human, and environmental systems. Ann N Y Acad Sci 1441:17–30. https://doi.org/10.1111/nyas.14036

Greenberg J, Price B, Ware A (2010) Alternative estimate of source distribution in microbial source tracking using posterior probabilities. Water Res 44:2629–2637. https://doi.org/10.1016/j.watres.2010.01.018

Grudzinski B, Fritz K, Dodds W (2020) Does riparian fencing protect stream water quality in cattle-grazed lands? Environ Manage 66:121–135 https://doi.org/10.1007/s00267-020-01297-2

Gunnigle E, Ramond J-B, Guerrero LD, et al (2015) Draft genomic DNA sequence of the multi-resistant Sphingomonas sp. strain AntH11 isolated from an Antarctic hypolith. FEMS Microbiol Lett 362:fnv037. https://doi.org/10.1093/femsle/fnv037

Gupta S, Graham DW, Sreekrishnan TR, Ahammad SZ (2022) Effects of heavy metals pollution on the co-selection of metal and antibiotic resistance in urban rivers in UK and India. Environmental Pollution 306:119326. https://doi.org/10.1016/j.envpol.2022.119326

Haberstroh CJ, Arias ME, Yin Z, Wang MC (2021) Effects of hydrodynamics on the cross-sectional distribution and transport of plastic in an urban coastal

river. Water Environment Research 93:186–200. https://doi.org/10.1002/wer.1386

HACH (2022) What is the factor to convert from PO4 to PO4-P? https://support.hach.com. Accessed 15 Nov 2023

Hagedorn C, Blanch A, Harwood V (2011) Microbial source tracking methods, applications, and case studies. Springer Science & Business Media

Hamilton KA, Garner E, Joshi S, et al (2020) Antimicrobial resistant microorganisms and their genetic determinants in stormwater: A systematic review. Curr Opin Environ Sci Health. https://doi.org/10.1016/j.coesh.2020.02.012

Han Z, Chen G, Li Y, et al (2015) Exploring the velocity distribution of debris flows: an iteration algorithm based approach for complex cross-sections. Geomorphology 241:72–82 https://doi.org/10.1016/j.geomorph.2015.03.043

Harnisz M, Kiedrzyńska E, Kiedrzyński M, et al (2020) The impact of WWTP size and sampling season on the prevalence of antibiotic resistance genes in wastewater and the river system. Science of the Total Environment 741:. https://doi.org/10.1016/j.scitotenv.2020.140466

Hart A, Warren J, Wilkinson H, Schmidt W (2023) Environmental surveillance of antimicrobial resistance (AMR), perspectives from a national environmental regulator in 2023. Eurosurveillance 28:2200367 https://doi.org/10.2807/1560-7917.ES.2023.28.11.2200367

Harwood VJ, Staley C, Badgley BD, et al (2014) Microbial source tracking markers for detection of fecal contamination in environmental waters: Relationships between pathogens and human health outcomes. FEMS Microbiol Rev 38:1–40. https://doi.org/10.1111/1574-6976.12031

Hassoun-Kheir N, Stabholz Y, Kreft J-U, et al (2021) EMBRACE-WATERS statement: Recommendations for reporting of studies on antimicrobial resistance in wastewater and related aquatic environments. One Health 13:100339. https://doi.org/10.1016/j.onehlt.2021.100339

He LY, He LK, Liu YS, et al (2019) Microbial diversity and antibiotic resistome in swine farm environments. Science of the Total Environment 685:197–207. https://doi.org/10.1016/j.scitotenv.2019.05.369

Hellweger FL (2013) Simple Model of Tetracycline Antibiotic Resistance in Aquatic Environment: Accounting for Metal Coselection. Journal of Environmental Engineering 139:913–921. https://doi.org/10.1061/(asce)ee.1943-7870.0000696

Hellweger FL, Ruan X, Sanchez S (2011) A simple model of tetracycline antibiotic resistance in the aquatic environment (with application to the Poudre River). Int J Environ Res Public Health 8:480–497. https://doi.org/10.3390/ijerph8020480

Hennessey M, Whatford L, Payne-Gifford S, et al (2020) Antimicrobial & antiparasitic use and resistance in British sheep and cattle: a systematic review. Prev Vet Med 185:105174. https://doi.org/10.1016/j.prevetmed.2020.105174

Hernandez-Suarez JS, Woznicki SA, Nejadhashemi AP (2020) Multi-site watershed model calibration for evaluating best management practice effectiveness in reducing fecal pollution. Human and Ecological Risk Assessment 26:2690–2715. https://doi.org/10.1080/10807039.2019.1680526

Herrmann M, Wegner C-E, Taubert M, et al (2019) Predominance of Cand. Patescibacteria in groundwater is caused by their preferential mobilization from soils and flourishing under oligotrophic conditions. Front Microbiol 10:1407. https://doi.org/10.3389/fmicb.2019.01407

Heß S, Berendonk TU, Kneis D (2018) Antibiotic resistant bacteria and resistance genes in the bottom sediment of a small stream and the potential impact of remobilization. FEMS Microbiol Ecol 94:128. https://doi.org/10.1093/femsec/fiy128

Heuvelink AE, Gonggrijp MA, Buter RGJ, et al (2019) Prevalence of extended-spectrum and AmpC β-lactamase-producing Escherichia coli in Dutch dairy herds. Vet Microbiol 232:58–64. https://doi.org/10.1016/j.vetmic.2019.04.005

Hibbing ME, Fuqua C, Parsek MR, Peterson SB (2010) Bacterial competition: surviving and thriving in the microbial jungle. Nat Rev Microbiol 8:15–25. https://doi.org/10.1038%2Fnrmicro2259

Ho JY, Jong MC, Acharya K, et al (2021) Multidrug-resistant bacteria and microbial communities in a river estuary with fragmented suburban waste

management. J Hazard Mater 405:. https://doi.org/10.1016/j.jhazmat.2020.124687

Hocquet D, Llanes C, Thouverez M, et al (2012) Evidence for induction of integron-based antibiotic resistance by the SOS response in a clinical setting. PLoS Pathog 8:e1002778. https://doi.org/10.1371/journal.ppat.1002778

Honda R, Tachi C, Yasuda K, et al (2020) Estimated discharge of antibiotic-resistant bacteria from combined sewer overflows of urban sewage system. NPJ Clean Water 3:15 https://doi.org/10.1038/s41545-020-0059-5

Hooban B, Fitzhenry K, Cahill N, et al (2021) A Point Prevalence Survey of Antibiotic Resistance in the Irish Environment, 2018–2019. Environ Int 152:. https://doi.org/10.1016/j.envint.2021.106466

Hooban B, Joyce A, Fitzhenry K, et al (2020) The role of the natural aquatic environment in the dissemination of extended spectrum beta-lactamase and carbapenemase encoding genes: A scoping review. Water Res 180:115880. https://doi.org/10.1016/j.watres.2020.115880

House G (2020) Modeling Fecal Indicator Bacteria and Antibiotic Resistance in Diverse Aquatic Environments. Masters dissertation, Virginia Tech

Hu Y, Yang X, Qin J, et al (2013) Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. Nat Commun 4:. https://doi.org/10.1038/ncomms3151

Hüffmeyer N, Klasmeier J, Matthies M (2009) Geo-referenced modeling of zinc concentrations in the Ruhr river basin (Germany) using the model GREAT-ER. Science of the total environment 407:2296–2305. https://doi.org/10.1016/j.scitotenv.2008.11.055

Hwengwere K, Paramel Nair H, Hughes KA, et al (2022) Antimicrobial resistance in Antarctica: is it still a pristine environment? Microbiome 10:1–13. https://doi.org/10.1186/s40168-022-01250-x

Ibekwe AM, Ma J, Murinda SE (2016) Bacterial community composition and structure in an Urban River impacted by different pollutant sources. Science of the Total Environment 566:1176–1185. https://doi.org/10.1016/j.scitotenv.2016.05.168

Im H, Kwon H, Cho G, et al (2019) Viscosity has dichotomous effects on Bdellovibrio bacteriovorus HD100 predation. Environ Microbiol 21:4675–4684. https://doi.org/10.1111/1462-2920.14799

Iqbal MS, Hofstra N (2019) Modeling Escherichia coli fate and transport in the Kabul River Basin using SWAT. Human and Ecological Risk Assessment: An International Journal 25:1279–1297. https://doi.org/10.1080/10807039.2018.1487276

Jampani M, Mateo-Sagasta J, Chandrasekar A, et al (2023) Fate and Transport Modelling for Evaluating Antibiotic Resistance in Aquatic Environments: Current Knowledge and Research Priorities. J Hazard Mater 132527. https://doi.org/10.1016/j.jhazmat.2023.132527

Janda JM, Abbott SL (2007) 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. J Clin Microbiol 45:2761–2764. https://doi.org/10.1128/jcm.01228-07

Jardé E, Jeanneau L, Harrault L, et al (2018) Application of a microbial source tracking based on bacterial and chemical markers in headwater and coastal catchments. Science of The Total Environment 610:55–63. https://doi.org/10.1016/j.scitotenv.2017.07.235

Jarvie HP, Pallett DW, Schäfer SM, et al (2020) Biogeochemical and climate drivers of wetland phosphorus and nitrogen release: Implications for nutrient legacies and eutrophication risk. Wiley Online Library https://doi.org/10.1002/jeq2.20155.PMID: 33459392

Jarvie HP, Smith DR, Norton LR, et al (2018) Phosphorus and nitrogen limitation and impairment of headwater streams relative to rivers in Great Britain: A national perspective on eutrophication. Science of the Total Environment 621:849–862. https://doi.org/10.1016/j.scitotenv.2017.11.128

Jeong J, Wagner K, Flores JJ, et al (2019) Linking watershed modeling and bacterial source tracking to better assess E. coli sources. Science of the Total Environment 648:164–175. https://doi.org/10.1016/j.scitotenv.2018.08.097

Jowett IG (1997) Instream flow methods: a comparison of approaches. Regulated Rivers: Research & Management: An International Journal Devoted to River Research and Management 13:115–127. https://doi.org/10.1002/(SICI)1099-1646(199703)13:2<115::AID-RRR440>3.0.CO;2-6

Karkman A, Pärnänen K, Larsson DGJ (2019) Fecal pollution can explain antibiotic resistance gene abundances in anthropogenically impacted environments. Nat Commun 10:1–8. https://doi.org/10.1038/s41467-018-07992-3

Keen PL, Knapp CW, Hall KJ, Graham DW (2018) Seasonal dynamics of tetracycline resistance gene transport in the Sumas River agricultural watershed of British Columbia, Canada. Science of the Total Environment 628–629:490–498. https://doi.org/10.1016/j.scitotenv.2018.01.278

Keen PL, Montforts MHMM (2011) Antimicrobial Resistance in the Environment, 1st edn. Wiley-Blackwell

Kim J-W, Pachepsky YA, Shelton DR, et al (2009) Including sediment-associated bacteria resuspension and settling in SWAT predictions of microbial water quality. In: 2009 International SWAT Conference. Available at: http://ssl-video. tamu. edu/video/D3-Kim. pdf

Kim J-W, Pachepsky YA, Shelton DR, Coppock C (2010) Effect of streambed bacteria release on E. coli concentrations: Monitoring and modeling with the modified SWAT. Ecol Modell 221:1592–1604. https://doi.org/10.1016/j.ecolmodel.2010.03.005

Kim KB, Kwon H-H, Han D (2018) Exploration of warm-up period in conceptual hydrological modelling. J Hydrol (Amst) 556:194–210. https://doi.org/10.1016/j.jhydrol.2017.11.015

Klappenbach JA (2001) rrndb: the Ribosomal RNA Operon Copy Number Database. Nucleic Acids Res. https://doi.org/10.1093/nar/29.1.181

Klare I, Badstübner D, Konstabel C, et al (1999) Decreased incidence of VanA-type vancomycin-resistant enterococci isolated from poultry meat and from fecal samples of humans in the community after discontinuation of avoparcin usage in animal husbandry. Microbial drug resistance 5:45–52. https://doi.org/10.1089/mdr.1999.5.45

Knapp CW, Callan AC, Aitken B, et al (2017) Relationship between antibiotic resistance genes and metals in residential soil samples from Western Australia. Environmental Science and Pollution Research 24:2484–2494. https://doi.org/10.1007/s11356-016-7997-y

Knapp CW, Dolfing J, Ehlert PAI, Graham DW (2010) Evidence of increasing antibiotic resistance gene abundances in archived soils since 1940. Environ Sci Technol 44:580–587. https://doi.org/10.1021/es901221x

Knapp CW, Lima L, Olivares-Rieumont S, et al (2012) Seasonal variations in antibiotic resistance gene transport in the Almendares River, Havana, Cuba. Front Microbiol 3:1–11. https://doi.org/10.3389/fmicb.2012.00396

Knapp CW, McCluskey SM, Singh BK, et al (2011) Antibiotic resistance gene abundances correlate with metal and geochemical conditions in archived Scottish soils. PLoS One 6:. https://doi.org/10.1371/journal.pone.0027300

Kneis D, Berendonk TU, Forslund SK, Hess S (2022) Antibiotic resistance genes in river biofilms: a metagenomic approach toward the identification of sources and candidate hosts. Environ Sci Technol 56:14913–14922. https://doi.org/10.1021/acs.est.2c00370

Knight GM, Davies NG, Colijn C, et al (2019) Mathematical modelling for antibiotic resistance control policy: Do we know enough? BMC Infect Dis 19:. https://doi.org/10.1186/s12879-019-4630-y

Knights D, Kuczynski J, Charlson ES, et al (2011) Bayesian community-wide culture-independent microbial source tracking. Nat Methods 8:761–765. https://doi.org/10.1038/nmeth.1650

Koczura R, Mokracka J, Taraszewska A, Łopacinska N (2016) Abundance of class 1 integron-integrase and sulfonamide resistance genes in river water and sediment is affected by anthropogenic pressure and environmental factors. Microb Ecol 72:909–916. https://doi.org/10.1007%2Fs00248-016-0843-4

Koerselman W, Meuleman AFM (1996) The vegetation N: P ratio: a new tool to detect the nature of nutrient limitation. Journal of applied Ecology 1441–1450. https://doi.org/10.2307/2404783

Kouyos RD, Abel zur Wiesch P, Bonhoeffer S (2011) On being the right size: the impact of population size and stochastic effects on the evolution of drug resistance in hospitals and the community. PLoS Pathog 7:e1001334 https://doi.org/10.1371/journal.ppat.1001334

Kozich JJ, Westcott SL, Baxter NT, et al (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon

sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol 79:5112–5120. https://doi.org/10.1128/aem.01043-13

Kümmerer K (2009) Antibiotics in the aquatic environment–a review–part I. Chemosphere 75:417–434. https://doi.org/10.1016/j.chemosphere.2008.11.086

LandIS (2023) LandIS: The Land Information System. https://www.landis.org.uk/ Accessed 18/10/2023

Larsson DGJ, Gaze WH, Laxminarayan R, Topp E (2023) AMR, One Health and the environment. Nat Microbiol 8:754–755 https://doi.org/10.1038/s41564-023-01351-9

Lawrence JG (2005) Horizontal and vertical gene transfer: the life history of pathogens. Contrib Microbiol 12:255–271. https://doi.org/10.1159/000081699

Léger A, Lambraki I, Graells T, et al (2021) AMR-Intervene: a social–ecological framework to capture the diversity of actions to tackle antimicrobial resistance from a One Health perspective. Journal of Antimicrobial Chemotherapy 76:1–21 https://doi.org/10.1093/jac/dkaa394

Leonard AFC, Zhang L, Balfour AJ, et al (2018) Exposure to and colonisation by antibiotic-resistant E. coli in UK coastal water users: Environmental surveillance, exposure assessment, and epidemiological study (Beach Bum Survey). Environ Int 114:326–333 https://doi.org/10.1016/j.envint.2017.11.003

Levin BR, Lipsitch M, Perrot V, et al (1997) The Population Genetics of Antibiotic Resistance. Clinical Infectious Diseases. https://doi.org/10.1093/clinids/24.supplement_1.s9

Li B, Yang Y, Ma L, et al (2015) Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. ISME Journal 9:2490–2502. https://doi.org/10.1038/ismej.2015.59

Li LG, Yin X, Zhang T (2018) Tracking antibiotic resistance gene pollution from different sources using machine-learning classification. Microbiome 6:93. https://doi.org/10.1186/s40168-018-0480-x

Li Q-M, Zhou Y-L, Wei Z-F, Wang Y (2021) Phylogenomic insights into distribution and adaptation of Bdellovibrionota in marine waters. Microorganisms 9:757. https://doi.org/10.3390/microorganisms9040757

Li R, Zhu L, Cui L, Zhu Y-G (2022a) Viral diversity and potential environmental risk in microplastic at watershed scale: Evidence from metagenomic analysis of plastisphere. Environ Int 161:107146. https://doi.org/10.1016/j.envint.2022.107146

Li W, Mao F, Ng C, et al (2022b) Population-based variations of a core resistome revealed by urban sewage metagenome surveillance. Environ Int 163:107185. https://doi.org/10.1016/j.envint.2022.107185

Ling Z, Yang Y, Huang Y, et al (2013) A preliminary investigation on the occurrence and distribution of antibiotic resistance genes in the Beijiang River, South China. Journal of Environmental Sciences 25:1656–1661 https://doi.org/10.1016/S1001-0742(12)60223-X

Liu Y-Y, Wang Y, Walsh TR, et al (2016) Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. Lancet Infect Dis 16:161–168. https://doi.org/10.1016/s1473-3099(15)00424-7

Luo Y, Mao D, Rysz M, et al (2010) Trends in antibiotic resistance genes occurrence in the Haihe River, China. Environ Sci Technol 44:7220–7225 https://doi.org/10.1021/es100233w

MacFadden DR, McGough SF, Fisman D, et al (2018) Antibiotic resistance increases with local temperature. Nat Clim Chang 8:510–514 https://doi.org/10.1038/s41558-018-0161-6

Magnano San Lio R, Favara G, Maugeri A, et al (2023) How antimicrobial resistance is linked to climate change: an overview of two intertwined global challenges. Int J Environ Res Public Health 20:1681. https://doi.org/10.3390%2Fijerph20031681

Mahaney AP, Franklin RB (2022) Persistence of wastewater-associated antibiotic resistant bacteria in river microcosms. Science of the Total Environment 819:153099. https://doi.org/10.1016/j.scitotenv.2022.153099

Maidment DR (1993) Handbook of hydrology. McGraw-Hill New York

Makarewicz JC, Lewis TW, Rea E, et al (2015) Using SWAT to determine reference nutrient conditions for small and large streams. J Great Lakes Res 41:123–135. https://doi.org/10.1016/j.jglr.2014.12.022

Mandal S, Van Treuren W, White RA, et al (2015) Analysis of composition of microbiomes: a novel method for studying microbial composition. Microb Ecol Health Dis 26:. https://doi.org/10.3402/mehd.v26.27663

Manyi-Loh C, Mamphweli S, Meyer E, Okoh A (2018) Antibiotic use in agriculture and its consequential resistance in environmental sources: potential public health implications. Molecules 23:795. https://doi.org/10.3390%2Fmolecules23040795

Mardis ER (2011) A decade's perspective on DNA sequencing technology. Nature 470:198–203. https://doi.org/10.1038/nature09796

Mathai PP, Staley C, Sadowsky MJ (2020) Sequence-enabled community-based microbial source tracking in surface waters using machine learning classification: A review. J Microbiol Methods 177:106050. https://doi.org/10.1016/j.mimet.2020.106050

Mayer RE, Reischer GH, Ixenmaier SK, et al (2018) Global distribution of human-associated fecal genetic markers in reference samples from six continents. Environ Sci Technol 52:5076–5084. https://doi.org/10.1021%2Facs.est.7b04438

McCann CM, Christgen B, Roberts JA, et al (2019) Understanding drivers of antibiotic resistance genes in High Arctic soil ecosystems. Environ Int 125:497–504. https://doi.org/10.1016/j.envint.2019.01.034

McGhee JJ, Rawson N, Bailey BA, et al (2019) Meta-SourceTracker: Application of Bayesian source tracking to shotgun metagenomics. https://doi.org/10.7287/peerj.preprints.27869v1

McLain JE, Cytryn E, Durso LM, Young S (2016) Culture-based methods for detection of antibiotic resistance in agroecosystems: Advantages, challenges, and gaps in knowledge. J Environ Qual 45:432–440. https://doi.org/10.2134/jeq2015.06.0317

McMurdie PJ, Holmes S (2014) Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. PLoS Comput Biol 10:. https://doi.org/10.1371/journal.pcbi.1003531

Mengistu TS, Garcias B, Castellanos G, et al (2022) Occurrence of multidrug resistant Gram-negative bacteria and resistance genes in semi-aquatic wildlife-Trachemys scripta, Neovison vison and Lutra lutra-as sentinels of environmental health. Science of the total environment 830:154814. https://doi.org/10.1016/j.scitotenv.2022.154814

Meshesha T, Wang J, Demelash Melaku N (2020) A modified hydrological model for assessing effect of pH on fate and transport of Escherichia coli in the Athabasca River basin. J Hydrol (Amst) 582:. https://doi.org/10.1016/j.jhydrol.2019.124513

Met Office (2023) Historic Station Data. https://www.metoffice.gov.uk/research/climate/maps-and-data/historic-station-data. Accessed 18 Oct 2023

Michaelis C, Grohmann E (2023) Horizontal gene transfer of antibiotic resistance genes in biofilms. Antibiotics 12:328. https://doi.org/10.3390/antibiotics12020328

Michan C, Blasco J, Alhama J (2021) High-throughput molecular analyses of microbiomes as a tool to monitor the wellbeing of aquatic environments. Microb Biotechnol 14:870–885. https://doi.org/10.1111/1751-7915.13763

Michaud JP, Wierenga M (2005) Estimating discharge and stream flows. A guide for sand and gravel operators. Ecology Publication 70:

Milewski A, Seyoum WM, Elkadiri R, Durham M (2019) Multi-scale hydrologic sensitivity to climatic and anthropogenic changes in Northern Morocco. Geosciences (Basel) 10:13. https://doi.org/10.3390/geosciences10010013

Mills MC, Lee J (2019) The threat of carbapenem-resistant bacteria in the environment: Evidence of widespread contamination of reservoirs at a global scale. Environmental Pollution 255:113143. https://doi.org/10.1016/j.envpol.2019.113143

Mohsen A, Park J, Chen YA, et al (2019) Impact of quality trimming on the efficiency of reads joining and diversity analysis of Illumina paired-end reads in the context of QIIME1 and QIIME2 microbiome analysis frameworks. BMC Bioinformatics 20:1–10. https://doi.org/10.1186/s12859-019-3187-5

Moon TK (1996) The expectation-maximization algorithm. IEEE Signal Process Mag 13:47–60. https://doi.org/10.1109/79.543975

Moriasi DN, Arnold JG, Van Liew MW, et al (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans ASABE 50:885–900. https://doi.org/10.13031/2013.23153

Morton JT, Marotz C, Washburne A, et al (2019) Establishing microbial composition measurement standards with reference frames. Nat Commun 10:. https://doi.org/10.1038/s41467-019-10656-5

Mujakić I, Piwosz K, Koblížek M (2022) Phylum Gemmatimonadota and its role in the environment. Microorganisms 10:151. https://doi.org/10.3390/microorganisms10010151

Mukherjee M, Gentry T, Mjelde H, et al (2020) Escherichia coli Antimicrobial Resistance Variability in Water Runoff and Soil from a Remnant Native Agricultural Watershed. 1–13 https://doi.org/10.3390/w12051251

Munck C, Albertsen M, Telke A, et al (2015) Limited dissemination of the wastewater treatment plant core resistome. Nat Commun 6:2–11. https://doi.org/10.1038/ncomms9452

Murray CJL, Ikuta KS, Sharara F, et al (2022) Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. The Lancet 399:629–655. https://doi.org/10.1016/S0140-6736(21)02724-0

Muurinen J, Stedtfeld R, Karkman A, et al (2017) Influence of Manure Application on the Environmental Resistome under Finnish Agricultural Practice with Restricted Antibiotic Use. Environ Sci Technol 51:5989–5999. https://doi.org/10.1021/acs.est.7b00551

Nash JE, Sutcliffe J V (1970) River flow forecasting through conceptual models part I—A discussion of principles. J Hydrol (Amst) 10:282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Neher TP, Lanying M, Moorman TB, et al (2020) Catchment-scale export of antibiotic resistance genes and bacteria from an agricultural watershed in central Iowa. 1–18. https://doi.org/10.1371/journal.pone.0227136

Neitsch SL, Arnold JG, Kiniry JR, Williams JR (2011) Soil and water assessment tool theoretical documentation version 2009. Texas Water Resources Institute

Nguyen T V, Dietrich J, Dang TD, et al (2022) An interactive graphical interface tool for parameter calibration, sensitivity analysis, uncertainty analysis, and visualization for the Soil and Water Assessment Tool. Environmental Modelling & Software 156:105497. https://doi.org/10.1016/j.envsoft.2022.105497

Niazi M, Obropta C, Miskewitz R (2015) Pathogen transport and fate modeling in the Upper Salem River Watershed using SWAT model. J Environ Manage 151:167–177. https://doi.org/10.1016/j.jenvman.2014.12.042

Niebaum G, Berlekamp J, Schmitt H, et al (2023) Geo-referenced simulations of E. coli in a sub-catchment of the Vecht River using a probabilistic approach. Science of the Total Environment 868:. https://doi.org/10.1016/j.scitotenv.2023.161627

Niewiadomska AM, Jayabalasingham B, Seidman JC, et al (2019) Population-level mathematical modeling of antimicrobial resistance: A systematic review. BMC Med 17:1–20. https://doi.org/10.1186/s12916-019-1314-9

O'Neill J (2016) Review on antimicrobial resistance: tackling drug-resistant infections globally: final report and recommendations. Wellcome Trust 2016:80

Oo HT, Zin WW, Kyi CCT (2020) Analysis of streamflow response to changing climate conditions using SWAT model. Civil Engineering Journal 6:194–209. https://doi.org/10.28991/cej-2020-03091464

Opatowski L, Guillemot D, Boëlle P-Y, Temime L (2011) Contribution of mathematical modeling to the fight against bacterial antibiotic resistance. Curr Opin Infect Dis 24:279–287. https://doi.org/10.1097/QCO.0b013e3283462362

Ott A, O'Donnell G, Tran NH, et al (2021a) Developing Surrogate Markers for Predicting Antibiotic Resistance "hot Spots" in Rivers Where Limited Data Are Available. Environ Sci Technol 55:7466–7478. https://doi.org/10.1021/acs.est.1c00939

Ott A, Quintela-Baluja M, Zealand AM, et al (2021b) Improved quantitative microbiome profiling for environmental antibiotic resistance surveillance. Environmental Microbiomes 16:1–14. https://doi.org/10.1186/s40793-021-00391-0

Pal C, Bengtsson-Palme J, Kristiansson E, Larsson DGJ (2015) Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. BMC Genomics 16:1–14. https://doi.org/10.1186/s12864-015-2153-5

Parajuli P (2007) SWAT bacteria sub-model evaluation and application. Doctor of Philosophy, Kansas State University

Peng F, Guo Y, Isabwe A, et al (2020) Urbanization drives riverine bacterial antibiotic resistome more than taxonomic community at watershed scale. Environ Int 137:105524. https://doi.org/10.1016/j.envint.2020.105524

Petrucci G, Bonhomme C (2014) The dilemma of spatial representation for urban hydrology semi-distributed modelling: Trade-offs among complexity, calibration and geographical data. J Hydrol (Amst) 517:997–1007. https://doi.org/10.1016/j.jhydrol.2014.06.019

Pray LA (2008) Transposons: The jumping genes. In: Nature education. https://www.nature.com/scitable/topicpage/transposons-the-jumping-genes. Accessed 4 May 2023

Pruden A, Arabi M, Storteboom HN (2012) Correlation between upstream human activities and riverine antibiotic resistance genes. Environ Sci Technol 46:11541–11549. https://doi.org/10.1021/es302657r

Pruden A, Pei R, Storteboom H, Carlson K (2006) Antibiotic resistance genes as emerging contaminants: Studies in Northern Colorado". Environ Sci Technol 40:7445–7450. https://doi.org/10.1021/es0680156

Pruden A, Vikesland PJ, Davis BC, de Roda Husman AM (2021) Seizing the moment: now is the time for integrated global surveillance of antimicrobial resistance in wastewater environments. Curr Opin Microbiol 64:91–99 https://doi.org/10.1016/j.mib.2021.09.01

Queenan AM, Bush K (2007) Carbapenemases: the versatile β-lactamases. Clin Microbiol Rev 20:440–458. https://doi.org/10.1128/cmr.00001-07

Quintela-Baluja M, Abouelnaga M, Romalde J, et al (2019) Spatial ecology of a wastewater network defines the antibiotic resistance genes in downstream receiving waters. Water Res. https://doi.org/10.1016/j.watres.2019.06.075

Quintela-Baluja M, Frigon D, Abouelnaga M, et al (2021) Dynamics of integron structures across a wastewater network – Implications to resistance gene transfer. Water Res 206:117720. https://doi.org/10.1016/j.watres.2021.117720

R Core Team (2018) A Language and Environment for Statistical Computing. R Foundation for Statistical Computing

Raleigh EA, Low KB (2013) Conjugation. Brenner's Encyclopedia of Genetics

Reller LB, Weinstein M, Jorgensen JH, Ferraro MJ (2009) Antimicrobial susceptibility testing: a review of general principles and contemporary practices. Clinical infectious diseases 49:1749–1755. https://doi.org/10.1086/647952

Richards S, Paterson E, Withers PJA, Stutter M (2016) Septic tank discharges as multi-pollutant hotspots in catchments. Science of the Total Environment 542:854–863. https://doi.org/10.1016/j.scitotenv.2015.10.160

Roberto AA, Van Gray JB, Engohang-Ndong J, Leff LG (2019) Distribution and co-occurrence of antibiotic and metal resistance genes in biofilms of an anthropogenically impacted stream. Science of the Total Environment 688:437–449. https://doi.org/10.1016/j.scitotenv.2019.06.053

Robins K, McCann CM, Zhou X-Y, et al (2022) Bioavailability of potentially toxic elements influences antibiotic resistance gene and mobile genetic element abundances in urban and rural soils. Science of the Total Environment 847:157512. https://doi.org/10.1016/j.scitotenv.2022.157512

Robinson TP, Bu DP, Carrique-Mas J, et al (2016) Antibiotic resistance is the quintessential One Health issue. Trans R Soc Trop Med Hyg

Romero F, Acuña V, Font C, et al (2019) Effects of multiple stressors on river biofilms depend on the time scale. Sci Rep 9:15810. https://doi.org/10.1038/s41598-019-52320-4

Rowland CS, Morton RD, Carrasco L, et al (2017) Land Cover Map 2015. NERC Environmental Information Data Centre. https://doi.org/10.5285/505d1e0c-ab60-4a60-b448-68c5bbae403e

Sabaliunas D, Webb SF, Hauk A, et al (2003) Environmental fate of triclosan in the River Aire Basin, UK. Water Res 37:3145–3154. https://doi.org/10.1016/s0043-1354(03)00164-7

Sabri Na, Schmitt H, Van der Zaan B, et al (2020) Prevalence of antibiotics and antibiotic resistance genes in a wastewater effluent-receiving river in the Netherlands. J Environ Chem Eng 8:102245. https://doi.org/10.1016/j.jece.2018.03.004

Sadeghi AM, Arnold JG (2002) A SWAT/microbial sub-model for predicting pathogen loadings in surface and groundwater at watershed and basin scales. Total Maximum Daily Load (TMDL): Environmental Regulations, Proceedings of 2002 Conference 56. https://doi.org/10.13031/2013.7529

Saleh A, Du B (2004) Evaluation of SWAT and HSPF within BASINS program for the upper North Bosque River watershed in central Texas. Transactions of the ASAE 47:1039–1049. https://doi.org/10.13031/2013.10387

Santhi C, Arnold JG, Williams JR, et al (2001) Validation of the SWAT Model on a Large River Basin with Point and Nonpoint Sources. JAWRA, 37, 1169-1188

Schowanek D, Webb S (2002) Exposure simulation for pharmaceuticals in European surface waters with GREAT-ER. Toxicol Lett 131:39–50. https://doi.org/10.1016/s0378-4274(02)00064-4

Schulze C, Matthies M (2001) Georeferenced aquatic fate simulation of cleaning agent and detergent ingredients in the river Rur catchment (Germany). Science of the total environment 280:55–77. https://doi.org/10.1016/S0048-9697(01)00814-2

Scott TM, Rose JB, Jenkins TM, et al (2002) Microbial source tracking: current methodology and future directions. Appl Environ Microbiol 68:5796–5803. https://doi.org/10.1128%2FAEM.68.12.5796-5803.2002

Seiler C, Berendonk TU (2012) Heavy metal driven co-selection of antibiotic resistance in soil and water bodies impacted by agriculture and aquaculture. Front Microbiol 3:. https://doi.org/10.3389/fmicb.2012.00399

Serwecińska L (2020) Antimicrobials and antibiotic-resistant bacteria: a risk to the environment and to public health. Water (Basel) 12:3313. https://doi.org/10.3390/w12123313

Shenhav L, Thompson M, Joseph TA, et al (2019) FEAST: fast expectation-maximization for microbial source tracking. Nat Methods 16:627–632. https://doi.org/10.1038/s41592-019-0431-x

Šimatović A, Udiković-Kolić N (2020) Antibiotic resistance in pharmaceutical industry effluents and effluent-impacted environments. Antibiotic resistance in the environment: A worldwide overview 101–122. https://doi.org/10.1007/698_2019_389

Singh R, Singh AP, Kumar S, et al (2019) Antibiotic resistance in major rivers in the world: A systematic review on occurrence, emergence, and management strategies. J Clean Prod 234:1484–1505. https://doi.org/10.1016/j.jclepro.2019.06.243

Sipos R, Székely AJ, Palatinszky M, et al (2007) Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. FEMS Microbiol Ecol 60:341–350. https://doi.org/10.1111/j.1574-6941.2007.00283.x

Slatko BE, Gardner AF, Ausubel FM (2018) Overview of next-generation sequencing technologies. Curr Protoc Mol Biol 122:e59. https://doi.org/10.1002%2Fcpmb.59

Smillie CS, Smith MB, Friedman J, et al (2011) Ecology drives a global network of gene exchange connecting the human microbiome. Nature 480:241–244. https://doi.org/10.1038/nature10571

Sockett RE, Lambert C (2004) Bdellovibrio as therapeutic agents: a predatory renaissance? Nat Rev Microbiol 2:669–675. https://doi.org/10.1038/nrmicro959

Son DI, Aleta P, Park M, et al (2018) Seasonal changes in antibiotic resistance genes in rivers and reservoirs in South Korea. J Environ Qual 47:1079–1085. https://doi.org/10.2134/jeq2017.12.0493

Sowah RA, Bradshaw K, Snyder B, et al (2020) Evaluation of the soil and water assessment tool (SWAT) for simulating E. coli concentrations at the

watershed-scale. Science of the Total Environment 746:. https://doi.org/10.1016/j.scitotenv.2020.140669

Spicknall IH, Foxman B, Marrs CF, Eisenberg JNS (2013) A modeling framework for the evolution and spread of antibiotic resistance: literature review and model categorization. Am J Epidemiol 178:508–520 https://doi.org/10.1093/aje/kwt017

Staley C, Gould TJ, Wang P, et al (2015) Species sorting and seasonal dynamics primarily shape bacterial communities in the Upper Mississippi River. Science of the Total Environment 505:435–445. https://doi.org/10.1016/j.scitotenv.2014.10.012

Staley C, Kaiser T, Lobos A, et al (2018) Application of SourceTracker for accurate identification of fecal pollution in recreational freshwater: a double-blinded study. Environ Sci Technol 52:4207–4217. https://doi.org/10.1021/acs.est.7b05401

Stapleton CM, Wyer MD, Kay D, et al (2007) Microbial source tracking: a forensic technique for microbial source identification? Journal of Environmental Monitoring 9:427–439. https://doi.org/10.1039/b617059e

Stedtfeld RD, Guo X, Stedtfeld TM, et al (2018) Primer set 2.0 for highly parallel qPCR array targeting antibiotic resistance genes and mobile genetic elements. FEMS Microbiol Ecol 94:fiy130. https://doi.org/10.1093/FEMSEC/FIY130

Stepanauskas R, Glenn TC, Jagoe CH, et al (2006) Coselection for microbial resistance to metals and antibiotics in freshwater microcosms. Environ Microbiol. https://doi.org/10.1111/j.1462-2920.2006.01091.x

Sterner RW, Elser JJ (2017) Ecological stoichiometry: the biology of elements from molecules to the biosphere. In: Ecological stoichiometry. Princeton university press

Stewart EJ (2012) Growing unculturable bacteria. J Bacteriol 194:4151–4160. https://doi.org/10.1128/jb.00345-12

Stewart JR, Santo Domingo JW, Wade TJ (2006) Fecal pollution, public health, and microbial source tracking. Microbial source tracking 1–32. https://doi.org/10.1128/9781555815769.ch1

Stoesser N, Phan HTT, Seale AC, et al (2020) Genomic epidemiology of complex, multispecies, plasmid-borne bla KPC carbapenemase in Enterobacterales in the United Kingdom from 2009 to 2014. Antimicrob Agents Chemother 64:10–1128. https://doi.org/10.1128/aac.02244-19

Strange JES, Leekitcharoenphon P, Møller FD, Aarestrup FM (2021) Metagenomics analysis of bacteriophages and antimicrobial resistance from global urban sewage. Sci Rep 11:1600. https://doi.org/10.1038/s41598-021-80990-6

Swift BMC, Bennett M, Waller K, et al (2019) Anthropogenic environmental drivers of antimicrobial resistance in wildlife. Science of the Total Environment 649:12–20 https://doi.org/10.1016/j.scitotenv.2018.08.180

Szekeres E, Chiriac CM, Baricz A, et al (2018) Investigating antibiotics, antibiotic resistance genes, and microbial contaminants in groundwater in relation to the proximity of urban areas. Environmental Pollution 236:734–744. https://doi.org/10.1016/j.envpol.2018.01.107

Tacão M, Laço J, Teixeira P, Henriques I (2022) CTX-M-Producing Bacteria Isolated from a Highly Polluted River System in Portugal. Int J Environ Res Public Health 19:11858. https://doi.org/10.3390/ijerph191911858

Taylor SD, He Y, Hiscock KM (2016) Modelling the impacts of agricultural management practices on river water quality in Eastern England. J Environ Manage 180:147–163. https://doi.org/10.1016/j.jenvman.2016.05.002

Thilakarathne BMMN (2020) Modeling the dissemination of antibiotic resistance in aquatic environments. Doctoral dissertation, Virginia Tech

Thompson LR, Sanders JG, McDonald D, et al (2017) A communal catalogue reveals Earth's multiscale microbial diversity. Nature 551:457–463. https://doi.org/10.1038/nature24621

Tian R, Ning D, He Z, et al (2020) Small and mighty: adaptation of superphylum Patescibacteria to groundwater environment drives their genome simplicity. Microbiome 8:1–15. https://doi.org/10.1186/s40168-020-00825-w

Tilman D (1977) Resource competition between plankton algae: an experimental and theoretical approach. Ecology 58:338–348. https://doi.org/10.2307/1935608

Tilman D (1982) Resource competition and community structure. Princeton university press

Tolson BA, Shoemaker CA (2007) Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. Water Resour Res 43:. https://doi.org/10.1029/2005WR004723

Traister E, Anisfeld SC (2006) Variability of indicator bacteria at different time scales in the upper Hoosic River watershed. Environ Sci Technol 40:4990–4995. https://doi.org/10.1021/es0601437

Ukuhor HO (2021) The interrelationships between antimicrobial resistance, COVID-19, past, and future pandemics. J Infect Public Health 14:53–60. https://doi.org/10.1016/j.jiph.2020.10.018

UNEP (2023) Bracing for Superbugs: Strengthening Environmental Action in the One Health Response to Antimicrobial Resistance. https://www.unep.org/resources/superbugs/environmental-action. Accessed 5 Jan 2024

Van Heijnsbergen E, Niebaum G, Lämmchen V, et al (2022) (Antibiotic-Resistant) E. coli in the Dutch-German Vecht Catchment─Monitoring and Modeling. Environ Sci Technol 56:15064–15073. https://doi.org/10.1021/acs.est.2c00218

van Leeuwen CJ, Vermeire TG (2007) Risk assessment of chemicals: an introduction. Springer

Van Liew MW, Veith TL, Bosch DD, Arnold JG (2007) Suitability of SWAT for the conservation effects assessment project: comparison on USDA agricultural research service watersheds. J Hydrol Eng Vol: 12:173–189. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:2(173)

Van Wijk WR, Scholte Ubing DW (1963) Radiation

Vandeputte D, Kathagen G, D'Hoe K, et al (2017) Quantitative microbiome profiling links gut community variation to microbial load. Nature 551:507–511. https://doi.org/10.1038/nature24460

Vaz-Moreira I, Nunes OC, Manaia CM (2014) Bacterial diversity and antibiotic resistance in water habitats: Searching the links with the human microbiome. FEMS Microbiol Rev 38:761–778. https://doi.org/10.1111/1574-6976.12062

Ventola CL (2015) The antibiotic resistance crisis: part 1: causes and threats. Pharmacy and therapeutics 40:277–83

VMD (2022) UK Veterinary Antibiotic Resistance and Sales Surveillance Report – UK-VARSS 2021, Published November 2022. In: Veterinary Medicines Directorate . www.nationalarchives.gov.uk/doc/open-government-licence/version/3/oremailPSI@nationalarchives.gov.uk. Accessed 18 Nov 2023

von Wintersdorff CJH, Penders J, van Niekerk JM, et al (2016) Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. Front Microbiol. 2016 Feb 19; 7. https://doi.org/10.3389/fmicb.2016.00173

Walsh TR (2018) A one-health approach to antimicrobial resistance. Nat Microbiol 3:854–855. https://doi.org/10.1038/s41564-018-0208-5

Wang H, Qi J-F, Qin R, et al (2023) Intensified livestock farming increases antibiotic resistance genotypes and phenotypes in animal feces. Commun Earth Environ 4:123 https://doi.org/10.1038/s43247-023-00790-w

Wang JY, An XL, Huang FY, Su JQ (2020) Antibiotic resistome in a landfill leachate treatment plant and effluent-receiving river. Chemosphere 242:. https://doi.org/10.1016/j.chemosphere.2019.125207

Wang X, Liu S, Guo H (2017) A nosocomial-pathogens-infections model with impulsive antibiotics treatment on multiple bacteria. Appl Math Comput 296:64–87 https://doi.org/10.1016/j.amc.2016.10.013

Waseem H, Jameel S, Ali J, et al (2019) Contributions and challenges of high throughput qPCR for determining antimicrobial resistance in the environment: A critical review. Molecules. https://doi.org/10.3390/molecules24010163

Watkins RR, Smith TC, Bonomo RA (2016) On the path to untreatable infections: colistin use in agriculture and the end of ‘last resort’antibiotics. Expert Rev Anti Infect Ther 14:785–788. https://doi.org/10.1080/14787210.2016.1216314

Watts G, Battarbee RW, Bloomfield JP, et al (2015) Climate change and water in the UK–past changes and future prospects. Prog Phys Geogr 39:6–28. https://doi.org/10.1177/0309133314542957

Wegener HC (2003) Antibiotics in animal feed and their role in resistance development. Curr Opin Microbiol 6:439–445. https://doi.org/10.1016/j.mib.2003.09.009

Weiss S, Xu ZZ, Peddada S, et al (2017) Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 5:1–18. https://doi.org/10.1186/s40168-017-0237-y

Wen T, Niu G, Chen T, et al (2023) The best practice for microbiome analysis using R. Protein Cell 2:1–13. https://doi.org/10.1093/procel/pwad024

Wepking C, Avera B, Badgley B, et al (2017) Exposure to dairy manure leads to greater antibiotic resistance and increased mass-specific respiration in soil microbial communities. Proceedings of the Royal Society B: Biological Sciences 284:. https://doi.org/10.1098/rspb.2016.2233

WHO (2016) Antimicrobial resistance: a manual for developing national action plans. World Health Organisation and the Food and Agriculture Organization of the United Nations and the World Organisation for Animal Health

WHO (2020) Global antimicrobial resistance surveillance system (GLASS) report: early implementation 2020. World Health Organization

WHO (2021a) Global antimicrobial resistance and use surveillance system (GLASS) report: 2021. World Health Organization

WHO (2021b) WHO integrated global surveillance on ESBL-producing E. coli using a “One Health” approach: Implementation and opportunities. World Health Organization

Wiedenbeck J, Cohan FM (2011) Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. FEMS Microbiol Rev 35:957–976 https://doi.org/10.1111/j.1574-6976.2011.00292.x

Wilkinson JL, Boxall ABA, Kolpin DW, et al (2022) Pharmaceutical pollution of the world's rivers. Proceedings of the National Academy of Sciences 119:e2113947119. https://doi.org/10.1073/pnas.2113947119

Woolhouse M, Ward M, Van Bunnik B, Farrar J (2015) Antimicrobial resistance in humans, livestock and the wider environment. Philosophical Transactions of the Royal Society B: Biological Sciences 370:20140083. https://doi.org/10.1098/rstb.2014.0083

Wright GD (2010) Antibiotic resistance in the environment: a link to the clinic? Curr Opin Microbiol 13:589–594 https://doi.org/10.1016/j.mib.2010.08.005

Wu D, Zhao J, Su Y, et al (2023) Explaining the resistomes in a megacity's water supply catchment: Roles of microbial assembly-dominant taxa, niched environments and pathogenic bacteria. Water Res 228:119359. https://doi.org/10.1016/j.watres.2022.119359

Xiang Q, Chen QL, Zhu D, et al (2018) Spatial and temporal distribution of antibiotic resistomes in a peri-urban area is associated significantly with anthropogenic activities. Environmental Pollution 235:525–533. https://doi.org/10.1016/j.envpol.2017.12.119

Xiao L, Estellé J, Kiilerich P, et al (2016) A reference gene catalogue of the pig gut microbiome. Nat Microbiol 1:1–6. https://doi.org/10.1038/nmicrobiol.2016.161

Xie H, Lian Y (2013) Uncertainty-based evaluation and comparison of SWAT and HSPF applications to the Illinois River Basin. J Hydrol (Amst) 481:119–131. https://doi.org/10.1016/j.jhydrol.2012.12.027

Xu X, Wang N, Lipson D, et al (2020) Microbial macroecology: In search of mechanisms governing microbial biogeographic patterns. Global Ecology and Biogeography 29:1870–1886. https://doi.org/10.1111/geb.13162

Xu Y, Han G, Zhang H, et al (2022) Application of fast expectation-maximization microbial source tracking to discern fecal contamination in

rivers exposed to low fecal inputs. Journal of Microbiology 60:594–601. https://doi.org/10.1007/s12275-022-1651-9

Yan C, Yang Y, Zhou J, et al (2013) Antibiotics in the surface water of the Yangtze Estuary: Occurrence, distribution and risk assessment. Environmental Pollution 175:22–29. https://doi.org/10.1016/j.envpol.2012.12.008

Yang J, Xiang J, Xie Y, et al (2022) Dynamic distribution and driving mechanisms of antibiotic resistance genes in a human-intensive watershed. Water Res 222:118841 https://doi.org/10.1016/j.watres.2022.118841

Yang T, Shi Y, Zhu J, et al (2021) The spatial variation of soil bacterial community assembly processes affects the accuracy of source tracking in ten major Chinese cities. Sci China Life Sci 1–14 https://doi.org/10.1007/s11427-020-1843-6

Yang X, Liu L, Liu X, et al (2023) The responding mechanism of indigenous bacteria in municipal wastewater inoculated with different concentrations of exogenous microalgae. J Environ Manage 345:118547. https://doi.org/10.1016/j.jenvman.2023.118547

Yu K, Li P, Chen Y, et al (2020) Antibiotic resistome associated with microbial communities in an integrated wastewater reclamation system. Water Res 173:115541. https://doi.org/10.1016/j.watres.2020.115541

Zainab SM, Junaid M, Xu N, Malik RN (2020) Antibiotics and antibiotic resistant genes (ARGs) in groundwater: A global review on dissemination, sources, interactions, environmental and human health risks. Water Res 187:116455. https://doi.org/10.1016/j.watres.2020.116455

Zhang B, Qin S, Guan X, et al (2021) Distribution of antibiotic resistance genes in Karst River and its ecological risk. Water Res 203:117507 https://doi.org/10.1016/j.watres.2021.117507

Zhang L, Xue M, Wang M, et al (2014) The spatiotemporal distribution of dissolved inorganic and organic carbon in the main stem of the Changjiang (Yangtze) River and the effect of the Three Gorges Reservoir. J Geophys Res Biogeosci 119:741–757. https://doi.org/10.1002/2012JG002230

Zhang M, Chen L, Ye C, Yu X (2018) Co-selection of antibiotic resistance via copper shock loading on bacteria from a drinking water bio-filter. Environmental Pollution 233:132–141. https://doi.org/10.1016/j.envpol.2017.09.084

Zhang S-X, Zhang Q-Q, Liu Y-S, et al (2020) Emission and fate of antibiotics in the Dongjiang River Basin, China: Implication for antibiotic resistance risk. Science of The Total Environment 712:136518. https://doi.org/10.1016/j.scitotenv.2020.136518

Zhang X, Srinivasan R, Hao F (2007) Predicting hydrologic response to climate change in the Luohe River basin using the SWAT model. Trans ASABE 50:901–910. https://doi.org/10.13031/2013.23154

Zhang Y, Hao W, Minghui HU, et al (2023a) Heavy metals potentially drive co-selection of antibiotic resistance genes by shifting soil bacterial communities in paddy soils along middle and lower Yangtze River. Pedosphere. https://doi.org/10.1016/j.pedsph.2023.01.012

Zhang Y, Liu C, Chen H, et al (2022) Metagenomic insights into resistome coalescence in an urban sewage treatment plant-river system. Water Res 224:119061. https://doi.org/10.1016/j.watres.2022.119061

Zhang Y, Zhu Z, Jiang Y, et al (2023b) Addition of Bdellovibrio to aquaculture water can significantly alter the distribution of microbial community on the gills and enhance the survival rate of Carassius auratus gibelio. Aquaculture 576:739820. https://doi.org/10.1016/j.aquaculture.2023.739820

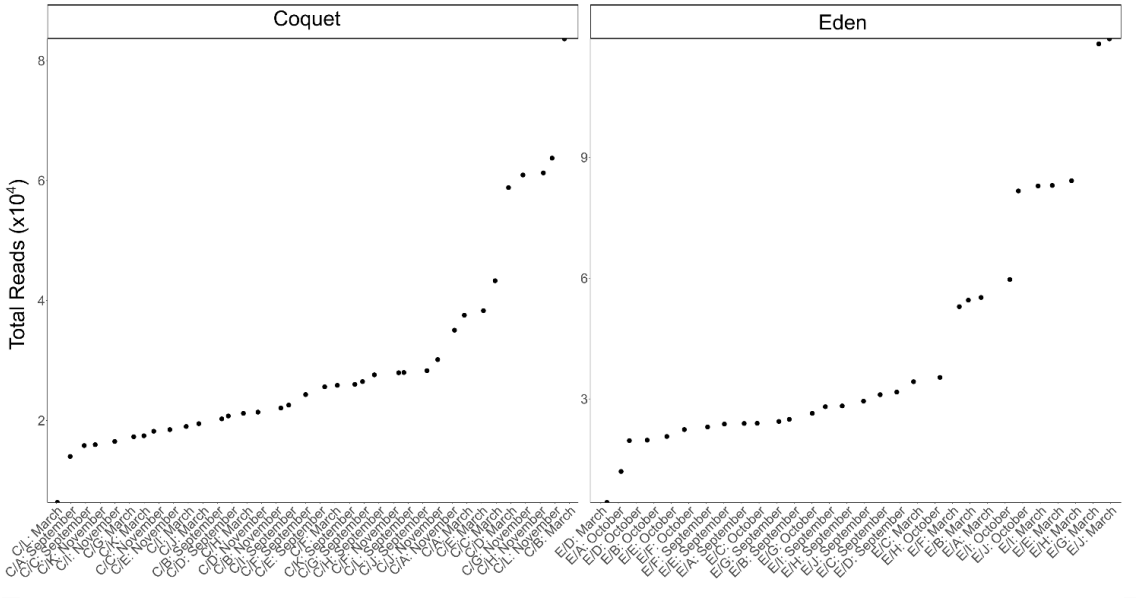
Zhou L, Liu L, Chen W-Y, et al (2020) Stochastic determination of the spatial variation of potentially pathogenic bacteria communities in a large subtropical river. Environmental Pollution 264:114683. https://doi.org/10.1016/j.envpol.2020.114683

Appendix

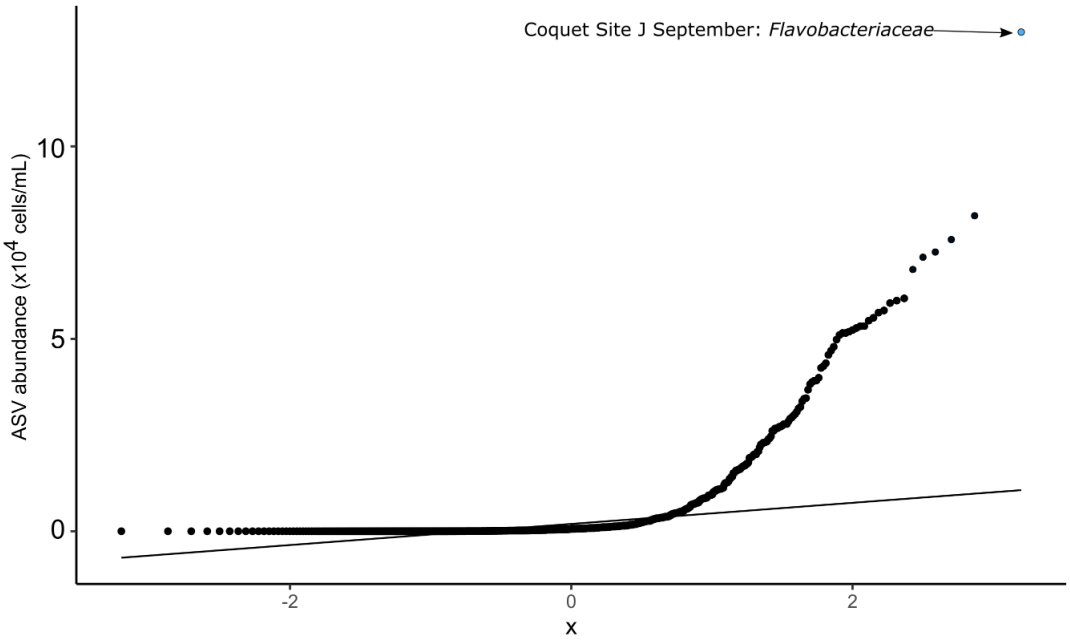
List of appendix figures

List of appendix tables

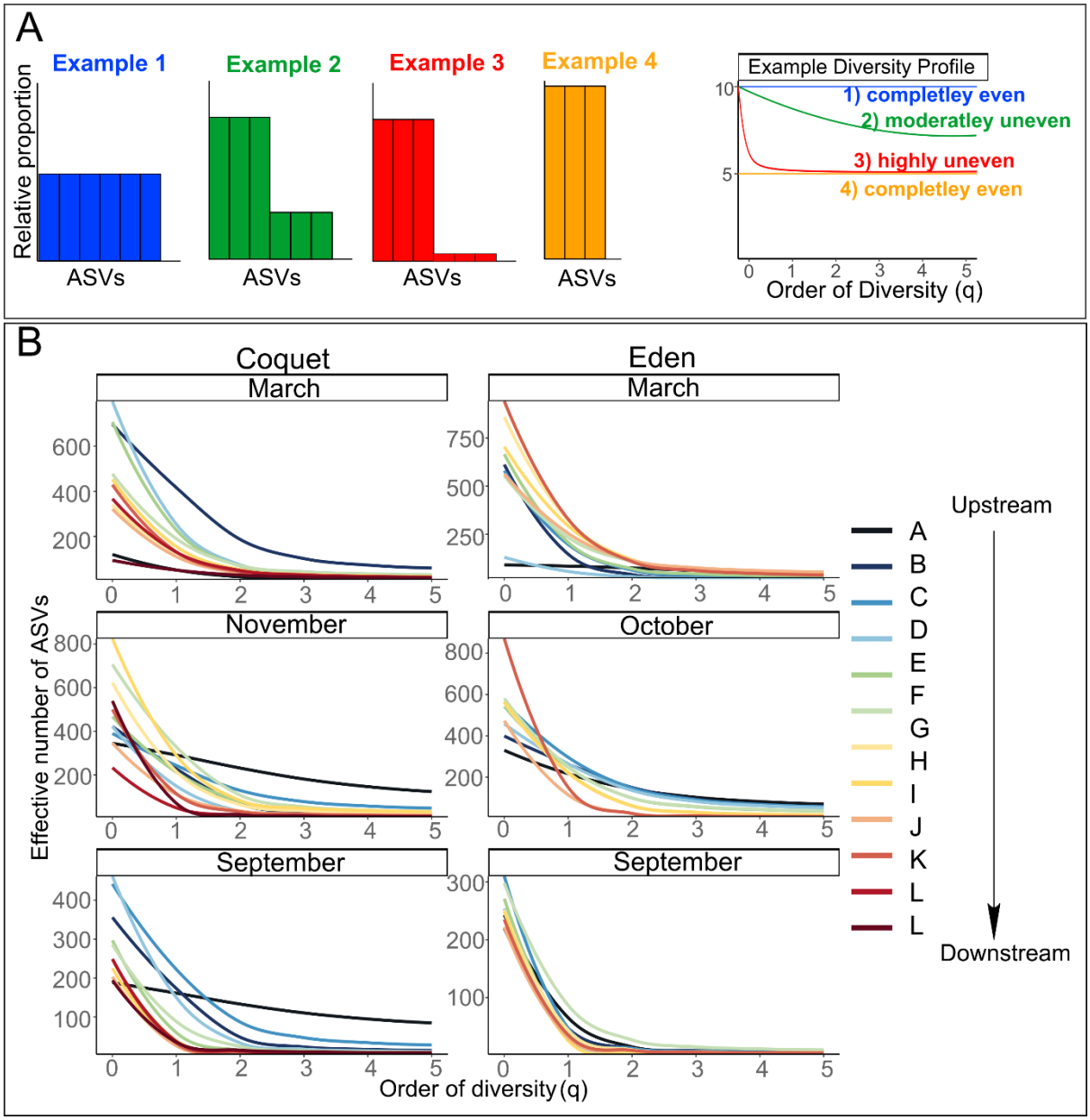
Appendix A



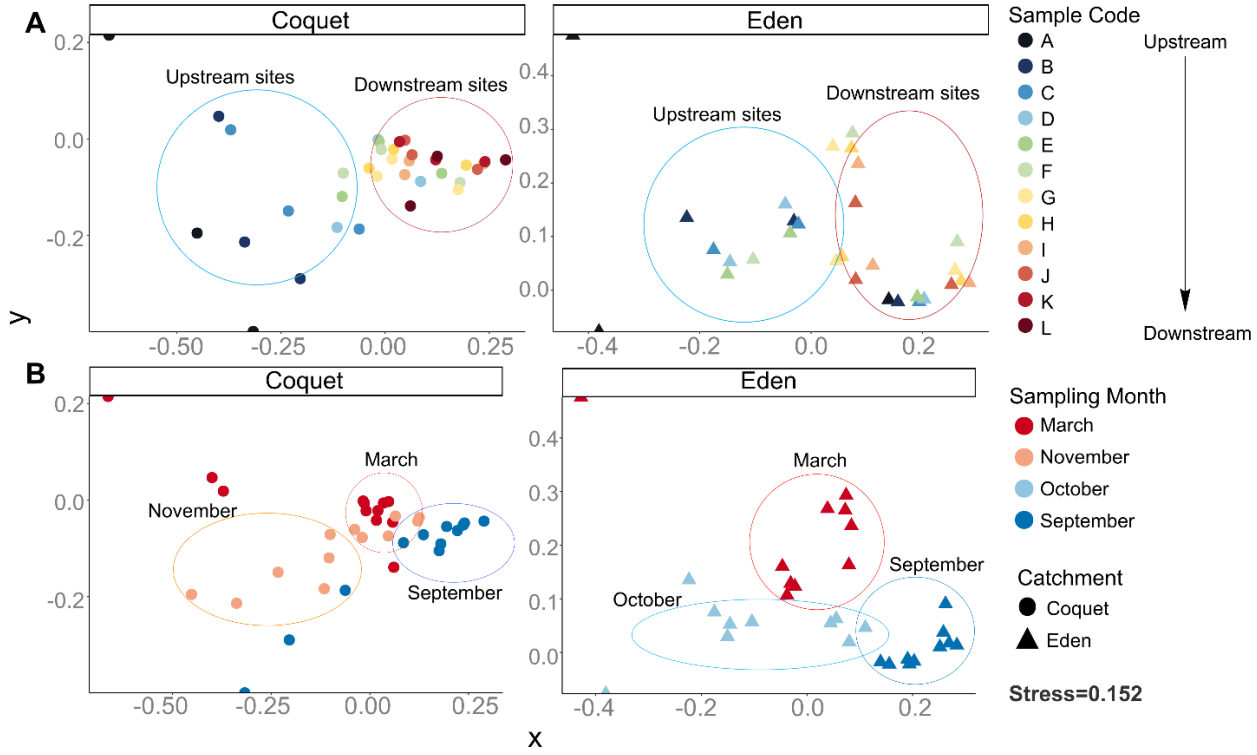
Appendix Figure A-1 Read depth amongst samples from lowest read depth to highest. X-axis shows the Catchment Code (C/E), the Site Code (A:L) and the sampling month. For example, Coquet Site A sampled in September would be C/A: September



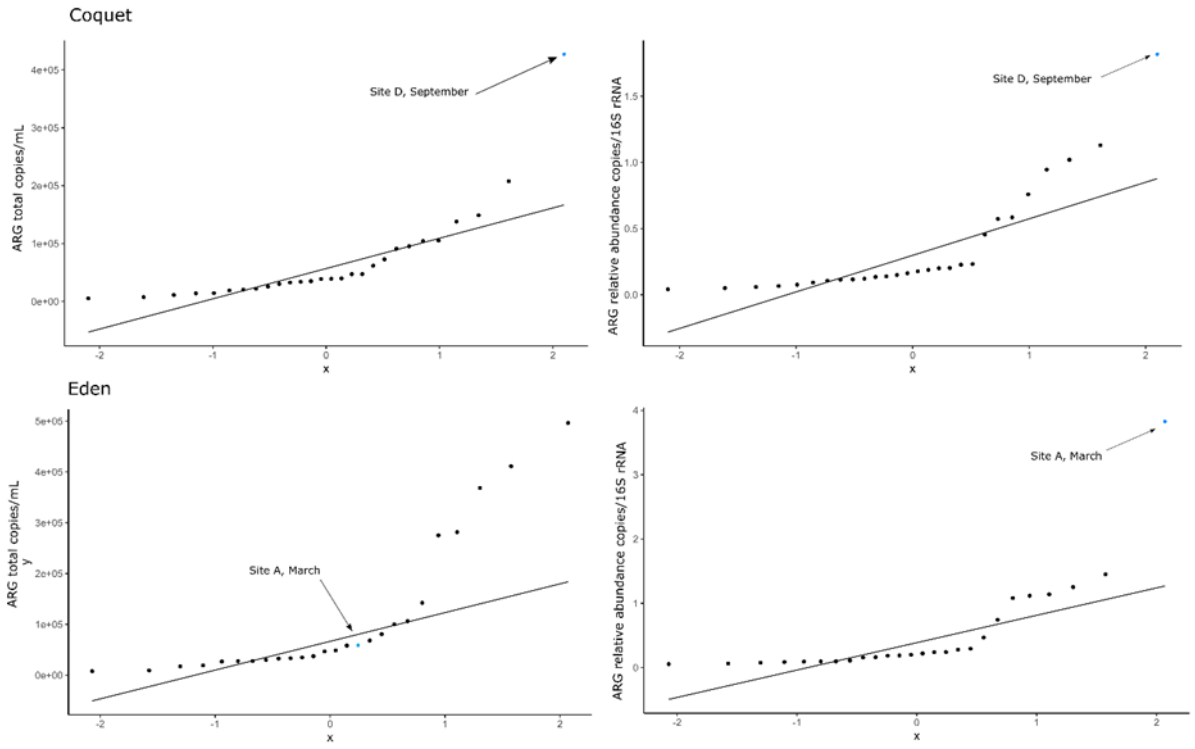
Appendix Figure A-2 QQ plot to identify outlier samples based on ASV abundance using QMP analysis.



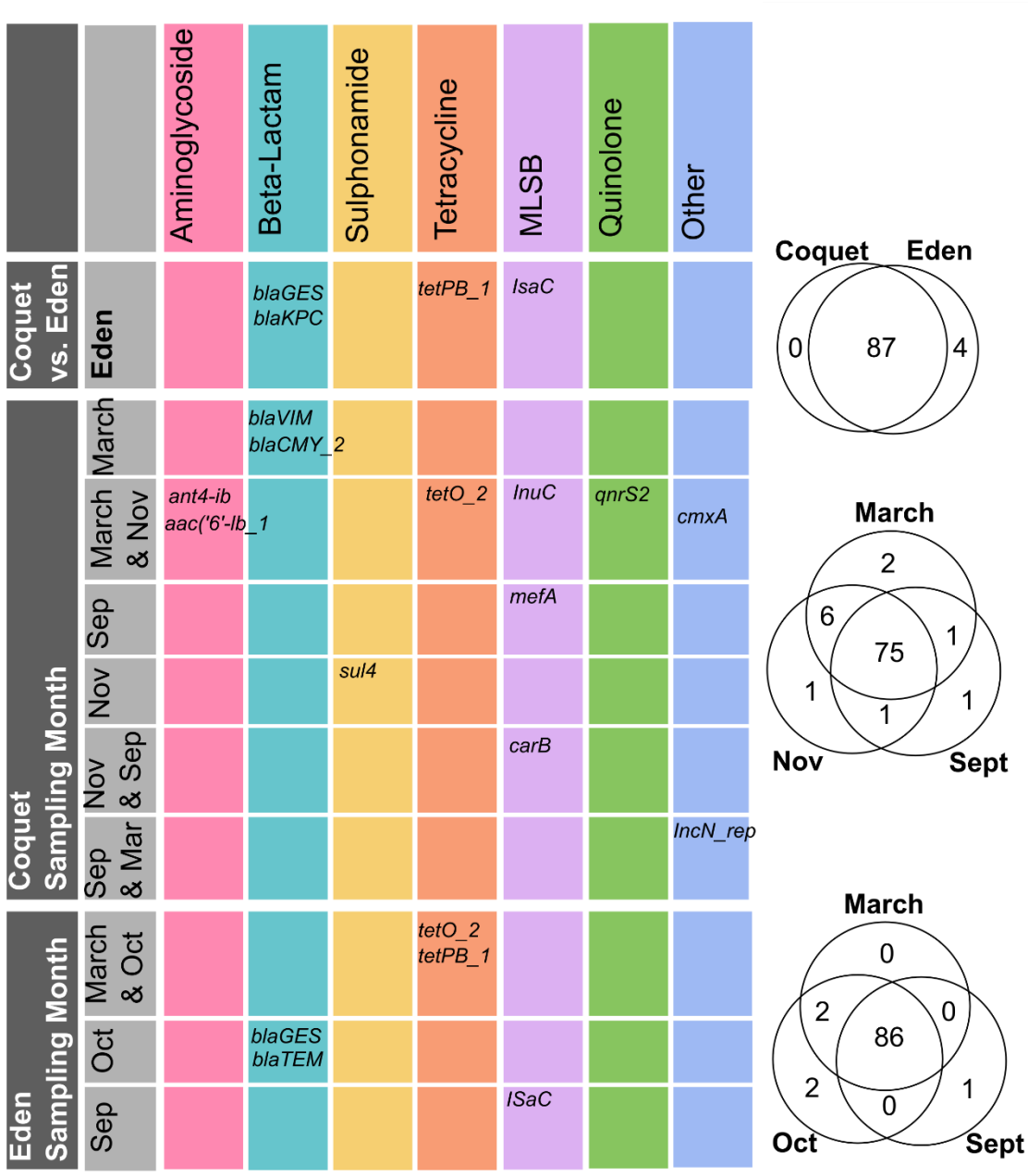
Appendix Figure A-3 A) Example diversity profile adapted from Alberdi et al, 2019 and Ott et al, 2021b used to provide context to how profiles can show differences in community evenness and richness. Example 1 and Example 4 show equal amplicon sequence variants (ASVs), therefore they are both completely even, however as example 4 has half the number of ASVs in a sample, the hill numbers are also half. Their evenness is reflected with their straight diversity profiles. Example 2 and 3 show examples of less evenly distributed systems, which are reflected in the shape of the profile plots. B) Diversity profiles for the Coquet and the Eden for sampling sites up to downstream during different sampling months, based on the quantitative microbial profiling (QMP) analysis.



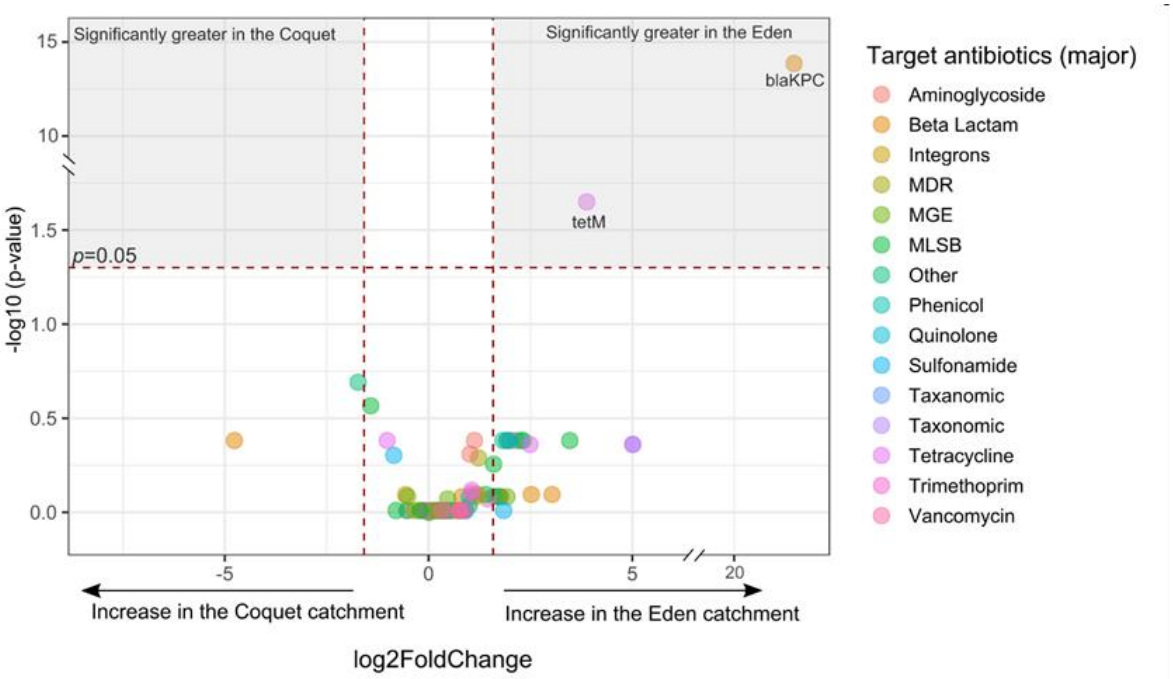
Appendix Figure A-4 NMDS plot displaying the Sorenson Overlap dissimilarity measure at q=1, divided by catchment and coloured according to (A) Sample code, where upstream site and downstream site groups are highlighted, and (B) sampling month. The stress value refers to the 'goodness of fit' of the NMDS ordination.



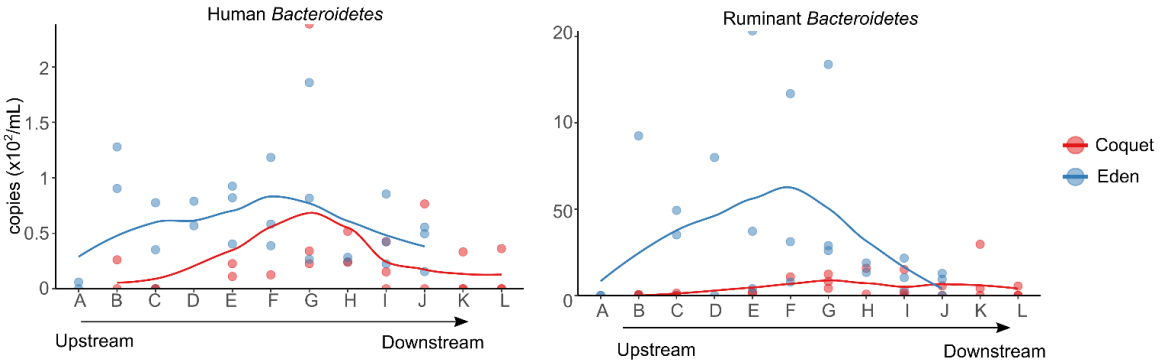
Appendix Figure A-5 QQ-plot to identify outlier samples based on absolute (copies/mL) (left column) and relative abundance (copies/16S rRNA) (right column) of ARGs and MGEs per sample.



Appendix Figure A-6 Number of shared ARGs/MGEs between the Coquet and Eden and sampling months. Based on HT-qPCR data, the catchments had similar resistomes, where the Eden had four unique genes. The different in sampling months can also be observed, where the Coquet had at least one unique gene in all sampling months, as well as six unique genes observed in the winter months of March and November, conferring resistance to Aminoglycoside, Tetracyclines, MLSB, quinolone and other types of antibiotics. In the Eden, the months of October and September had two and one unique gene, respectively. Two tetracycline ARGs were also observed in March and October, but not September.



Appendix Figure A-7 Volcano plot showing the log2FoldChange in ARGs and MGEs between catchments (Wald test). The p-value is adjusted according to Benjamini Hochbery ($p<0.05$). Whilst most genes had similar absolute abundance between catchments, there were higher abundances of some genes in the Eden catchment, with two genes being significantly more abundant in the Eden catchment, the tetracycline gene tetM and the beta-lactam gene bla_{KPC}.



Appendix Figure A-8 Scatter plot with fitted line showing the abundance (copies/mL) of human Bacteriodetes (HuBac) MST marker and Ruminant Bacteroidetes (RuBac) MST marker from upstream to downstream. Line colour indicates the abundance in the Coquet and Eden catchments.

Appendix Table A-1 Grouping of LCM 2015 land-use classifications into 'urban', 'rural, and 'pristine'.

Urban	Rural	Pristine
Suburban Urban	Improved grassland, Arable and horticulture	Acid grassland, Bog, Broadleaved woodland, Calcareous grassland, Coniferous woodland, Fen Marsh Swamp, Heather, Heather Grassland, Inland rock, Littorial Rock, Littorial sediment, Saltmarsh, Supra-littorial sediment

Appendix Table A-2 Coordinates of Sampling Sites for this study.

Catchment	Code	Sample Site	Coordinates
Coquet	A	Ullswater	55.36368,-2.18461
	B	Sharperton	55.32703, -2.07515
	C	Hepple	55.29616,-2.03117
	D	Rothbury U/S	55.30531-1.93459
	E	Rothbury D/S	55.30839,-1.9086
	F	Rothbury STP D/S	55.30347,-1.88921
	G	Pauperhaugh U/S	55.28955-1.84313
	H	Todstead Farm STW D/S	55.2805-1.78434
	I	Felton U/S	55.296727, -1.709756
	J	Felton STW D/S	55.3207,-1.67626
	K	Coquet U/S Warkworth	55.33836,-1.63139
	L	Warkworth Medieval Bridge	55.3496,-1.60958

Eden	A	Tommy Road	54.420143, -2.338734
	B	Home St-Appleby	54.576459, -2.494481,
	C	Bolton	54.605556, -2.556082,
	D	Farmland near Temple_Sowerby	54.621344, -2.599331
	E	Temple Sowerby_Park D/S	54.647181, -2.615354
	F	Langwathby D/S	54.694876, -2.674814
	G	Lazonby D/S	54.756359, -2.699388
	H	Armathwaite	54.806792, -2.768275
	I	Warwick Bridge_D/S Wetheral STP	54.902496, -2.829087
	J	Carlisle U/S	54.899173, -2.920526

Appendix Table A-3: Sampling Dates for the Coquet and the Eden

Coquet	**Eden**
7th of September 2020	2nd of September 2020
19th of November 2020	8th of October 2020
15th of March 2021	1st of March 2021

Appendix Table A-4: Information on primers/probes using in qPCR reactions. HuBac and RuBac primer and probe sequences are considered intellectual property of the Environment Agency therefore could not be provided.

		Primer		**Amplicon Size**	**Annealing (°C)**	**Reference**
16S rRNA	F	F (1055f)	ATGGYTGTCGTCAGCT		60	(Dionisi et al. 2003)
	R	R (1392r)	ACGGGCGGTGTGTAC			
	probe	Probe (Total_bac)	TxRed-CAACGAGCGCAACCC[BHQ-2]			

Faecal Coliforms	F	Eco1457F	CATTGACGTTACCCG CAGAAGAAGC	190	58	(Bartosch et al. 2004)
	R	Eco1652R	CTCTACGAGACTCAA GCTTGC			

Appendix Table A-5 Samples not sent for HT-qPCR.

Catchment	**Sampling Month**	**Site Code**
Eden	September 2020	A
	September 2020	B
	October 2020	H
	March 2021	D
Coquet	September 2020	A
	November 2020	A
	November 2020	D
	March 2021	A

Appendix Table A-6 Representative samples sent for HT-qPCR for analysis of 384 ARGs and MGEs, before narrowing to 96 genes. Samples were chosen from the present study, as well as a later sampling campaign which is unpublished.

Catchment	**Sampling Month**	**Site Code**
Samples from this study		
Eden	September 2020	E
Eden	September 2020	G
Eden	September 2020	J
Coquet	September 2020	C
Coquet	September 2020	D
Coquet	September 2020	H
Coquet	September 2020	J
Coquet	September 2020	L
Eden	October 2020	A
Eden	October 2020	B
Eden	October 2020	E
Eden	October 2020	G
Eden	October 2020	J
Coquet	November 2020	B
Coquet	November 2020	H
Coquet	November 2020	J
Coquet	November 2020	L
Eden	March 2021	B
Eden	March 2021	E
Coquet	March 2021	B
Coquet	March 2021	D
Coquet	March 2021	F
Coquet	March 2021	J
Coquet	March 2021	L
Samples from separate sampling campaign *(unpublished)*		
Coquet	March 2022	BS1 (wastewater effluent)

Coquet	April 2022	DS2 (wastewater effluent)
Coquet	July 2022	CS1 (wastewater effluent)
Coquet	August 2022	BS1 (wastewater effluent)
Coquet	March 2022	A soil
Coquet	April 2022	D2 soil
Coquet	July 2022	C2 soil
Coquet	August 2022	B4 soil

Appendix Table A-7 Ninety-six assays used for HT-qPCR.

Assay	Gene	Target antibiotics (major)	Forward Primer	Reverse Primer
AY1	16S rRNA	16S rRNA	GGGTTGCGCTCGTTGC	ATGGYTGTCGTCAGCTCGTG
AY8	aac(6')-Ib_1	Aminoglycoside	CGTCGCCGAGCAACTTG	CGGTACCTTGCCTCTCAAACC
AY396	aac(6)-ig	Aminoglycoside	GCGATGTTAGAAGCCTCAATTCG	CACACTTCGGCCTGTCGAA
AY397	aac(6)-iic	Aminoglycoside	CAGTCTTTGGCTAATCCATCACAG	AACGAACCCGGCCTTCTC
AY404	aac(6)-iz	Aminoglycoside	TGCGCCATGACTACGTGAAC	GACTGTCCGAAGCCAGTTCG
AY3	aacC4	Aminoglycoside	CGGCGTGGGACACGAT	AGGGAACCTTTGCCATCAACT
AY10	aadA_1	Aminoglycoside	GTTGTGCACGACGACATCATT	GGCTCGAAGATACCTGCAAGAA
AY408	aadA10	Aminoglycoside	ACAGGCACTCAACGTCATCG	CGCGGAGAACTCTGCTTTGA
AY411	aadA6	Aminoglycoside	CCATCGAGCGTCATCTGGAA	CCCGTCTGGCCGGATAAC
AY412	aadA7	Aminoglycoside	CACTCCGCGCCTTGGA	TGTGGCGGGCTCGAAG

AY414	ant4-ib	Aminoglycoside	GATGGCCGCTGACACATG	TCAACATTGCGCCATAGTGG
AY24	strB	Aminoglycoside	GCTCGGTCGTGAGAACAATCT	CAATTTCGGTCGCCTGGTAGT
AY444	blaACT	Beta Lactam	AAGCCGCTCAAGCTGGA	GCCATATCCTGCACGTTGG
AY446	blaCARB	Beta Lactam	TGATTTGAGGGATACGACAACTCC	CTGTAATACTCCGAGCACCAA
AY339	blaCMY_2	Beta Lactam	AAAGCCTCATGGGTGCATAAA	ATAGCTTTTGTTTGCCAGCATCA
AY432	blaCTX-M	Beta Lactam	CGTACCGAGCCGACGTTAA	CAACCCAGGAAGCAGGCA
AY125	blaGES	Beta Lactam	GCAATGTGCTCAACGTTCAAG	GTGCCTGAGTCAATTCTTTCAAAG
AY447	blaGOB	Beta Lactam	CTTGGGCTTGAATGCTCAGGTA	TGTATGGTCGTAGTGAGCCTGA
AY440	blaKPC	Beta Lactam	GCCGCCAATTTGTTGCTGAA	GCCGGTCGTGTTTCCCTTT
AY452	blaMIR	Beta Lactam	CGGTCTGCCGTTACAGGTG	AAAGACCCGCGTCGTCATG
AY152	blaNDM	Beta Lactam	GGCCACACCAGTGACAATATCA	CAGGCAGCCACCAAAAGC
AY601	blaOXA48	Beta Lactam	TGTTTTTGGTGGCATCGAT	GTAAMRATGCTTGGTTCGC
AY108	blaOXY	Beta Lactam	CGTTCAGGCGGCAGGTT	GCCGCGATATAAGATTTGAGAATT
AY126	blaSFO	Beta Lactam	CCGCCGCCATCCAGTA	GGGCCGCCAAGATGCT
AY438	blaSHV11	Beta Lactam	TTGACCGCTGGGAAACGG	TCCGGTCTTATCGGCGATAAAC
AY439	blaTEM	Beta Lactam	CGCCGCATACACTATTCTCAG	GCTTCATTCAGCTCCGGTTC
AY129	blaVIM	Beta Lactam	GCACTTCTCGCGGAGATTG	CGACGGTGATGCGTACGTT
AY114	cfxA	Beta Lactam	TCATTCCTCGTTCAAGTTTTCAGA	TGCAGCACCAAGAGGAGATGT
AY111	cphA_1	Beta Lactam	GCGAGCTGCACAAGCTGAT	CGGCCCAGTCGCTCTTC

AY138	penA	Beta Lactam	AGACGGTAACGTATAACTTTTTGAAAGA	GCGTGTAGCCGGCAATG
AY293	intI1_1	Integrons	CGAACGAGTGGCGGAGGGTG	TACCCGAGAGCTTGGCACCCA
AY493	czcA	MDR	GCCTTGTTCATCGGCGAAC	GGCAATGTCGCCTTCGTTC
AY485	mdtA	MDR	ACAAGCCCAGGGCCAAC	CCTTAATGGTGCCTTCGGTTTC
AY486	mdtH	MDR	ATGCTGGCTGTACAAGTGATG	CACTCCAGCGGGCGATA
AY227	mepA	MDR	ATCGGTCGCTCTTCGTTCAC	ATAAATAGGATCGAGCTGCTGGAT
AY224	oprD	MDR	ATGAAGTGGAGCGCCATTG	GGCCACGGCGAACTGA
AY489	qacF/H	MDR	CTGAAGTCTAGCCATGGATTCACTAG	CAAGCAATAGCTGCCACAAGC
AY316	IncN_rep	MGE	AGTTCACCACCTACTCGCTCCG	CAAGTTCTTCTGTTGGGATTCCG
AY318	IncP_oriT	MGE	CAGCCTCGCAGAGCAGGAT	CAGCCGGGCAGGATAGGTGAAGT
AY506	IS1247_1	MGE	CGGCCGTCACTGACCAA	TCGGCAGGTTGGTGACG
AY514	IS5/IS1182	MGE	TTCTCGAAGAATCGCCATGGC	GCTTTGGATCGCTCCAATCGA
AY311	ISAba3	MGE	TCAGAGGCAGCGGTATACGA	GGTTGATTCAGTTAAAGTACGTAAAACTTT
AY309	ISPps	MGE	CACACTGCAAAAACGCATCCT	TGTCTTTGGCGTCACAGTTCTC
AY524	Tn5403	MGE	AAGCGAATGGCGCGAAC	CGCGCAGGGTAAACTGC
AY299	tnpA_1	MGE	GCCGCACTGTCGATTTTTATC	GCGGGATCTGCCACTTCTT
AY302	tnpA_4	MGE	CATCATCGGACGGACAGAATT	GTCGGAGATGTGGGTGTAGAAAGT
AY303	tnpA_5	MGE	GAAACCGATGCTACAATATCCAATTT	CAGCACCGTTTGCAGTGTAAG
AY92	carB	MLSB	GGAGTGAGGCTGACCGTAGAAG	ATCGGCGAAACGCACAAA

AY531	erm35	MLSB	CCTTCAGTCAGAACCGGCAA	GCTGATTTGACAGTTGGTGGTG
AY54	erm36	MLSB	GGCGGACCGACTTGCAT	TCTGCGTTGACGACGGTTAC
AY547	ermB_3	MLSB	TGAAAGCCATGCGTCTGAC	TTCAGCTGGCAGCTTAAGC
AY546	ermX_2	MLSB	TGATGACGGCTCAGTGG	GTGCACCAGCGCCTGA
AY537	lnuC	MLSB	GGGTGTAGATGCTCTTCTTGGA	CTTTACCCGAAAGAGTTTCTACCG
AY550	lsaC	MLSB	AAACGGCGTGAAAGTATCAGG	TTGTGGTGATGTAACGGATGC
AY538	mefA	MLSB	TAATTATCGCAGCAGCTGGTTC	GTTCCCAAACGGAGTATAAGAGTG
AY539	mphA	MLSB	TCAGCGGGATGATCGACTG	GAGGGCGTAGAGGGCGTA
AY553	msrE	MLSB	CGGCAGATGGTCTGAGCTTAAA	CGCACTCTTCCTGCATAAAGGA
AY91	oleC	MLSB	CCCGGAGTCGATGTTCGA	GCCGAAGACGTACACGAACAG
AY554	vat(A)	MLSB	ATGAACGGAGCGAATCATCGG	CCATACCGATCCAAACGTCATTTC
AY470	arr3	Other	GATCGTCTTCGAACGGTCCTG	TTTGGCGATTGGTGACTTGCT
AY465	bacA	Other	ATCCGCGGCACCCTGA	CCTGCTTGATGGACTTGATGAAGA
AY198	crAss64	Other	TGTATAGATGCTGCTGCAACTGTACTC	CGTTGTTTTCATCTTTATCTTGTCCAT
AY466	mcr1	Other	CACATCGACGGCGTATTCTG	CAACGAGCATACCGACATCG
AY191	merA	Other	GTGCCGTCCAAGATCATG	GGTGGAAGTCCAGTAGGGTGA
AY558	catA3	Phenicol	CTGATTGCTCAGGCCGTGAA	ATGAGTATGGGCAACTCAGTGC
AY563	cmlV	Phenicol	GCCCTCATCACCGTCTTCG	GGACGTTGGCGATGGAGAG
AY37	cmxA	Phenicol	GCGATCGCCATCCTCTGT	TCGACACGGAGCCTTGGT

AY32	floR_1	Phenicol	ATTGTCTTCACGGTGTCCGTTA	CCGCGATGTCGTCGAACT
AY456	qepA	Quinolone	GGGCATCGCGCTGTTC	GCGCATCGGTGAAGCC
AY96	qnrB	Quinolone	GCGACGTTCAGTGGTTCAGA	GCTGCTCGCCAGTCGAA
AY457	qnrB4	Quinolone	TCACCACCCGCACCTG	GGATATCTAAATCGCCCAGTTCC
AY461	qnrS2	Quinolone	TCCCGAGCAAACTTTGCCAA	GGTGAGTCCCTATCCAGCGA
AY245	sul1_2	Sulfonamide	GCCGATGAGATCAGACGTATTG	CGCATAGCGCTGGGTTTC
AY365	sul2_2	Sulfonamide	TCATCTGCCAAACTCGTCGTTA	GTCAAAGAACGCCGCAATGT
AY244	sul3_1	Sulfonamide	TCCGTTCAGCGAATTGGTGCAG	TTCGTTCACGCCTTACACCAGC
AY241	sul4	Sulfonamide	TCAACGTCACTCCAGACAGC	TGGAAATAACGACGTCCACA
AY473	A. baumannii	Taxonomic	TCTTGGTGGTCACTTGAAGC	ACTCTTGTGGTTGTGGAGCA
AY476	Enterococci	Taxonomic	AGAAATTCCAAACGAACTTG	CAGTGCTCTACCTCCATCATT
AY478	K. pneumoniae	Taxonomic	ACGGCCGAATATGACGAATTC	AGAGTGATCTGCTCATGAA
AY479	P. aeruginosa	Taxonomic	AGCGTTCGTCCTGCACAAGT	TCCACCATGCTCAGGGAGAT
AY568	tet39	Tetracycline	TATAGCGGGTCCGGTAATAGGTG	CCATAACGATCCTGCCCATAGATAAC
AY254	tetA_2	Tetracycline	CTCACCAGCCTGACCTCGAT	CACGTTGTTATAGAAGCCGCATAG
AY572	tetG	Tetracycline	TCGCGTTCCTGCTTGCC	CCGCGAGCGACAAACCA
AY574	tetM	Tetracycline	GGAGCGATTACAGAATTAGGAAGC	TCCATATGTCCTGGCGTGTC
AY264	tetO_2	Tetracycline	CAACATTAACGGAAAGTTTATTGTATACCA	TTGACGCTCCAAATTCATTGTATC

AY274	tetPB_1	Tetracycline	TGGGCGACAGTAGGCTTAGAA	TGACCCTACTGAAACATTAGAAATATACCT
AY259	tetQ	Tetracycline	CGCCTCAGAAGTAAGTTCATACACTAAG	TCGTTCATGCGGATATTATCAGAAT
AY577	tetR	Tetracycline	CCGTCAATGCGCTGATGAC	GCCAATCCATCGACAATCACC
AY263	tetW	Tetracycline	ATGAACATTCCCACCGTTATCTTT	ATATCGGCGGAGAGCTTATCC
AY585	dfrA25	Trimethoprim	TCAAACTGGACAGCGGCTA	GTCGATTGTCGACACATGCA
AY586	dfrA27	Trimethoprim	GCCGCTCAGGATCGGTA	GTCGAGATATGTAGCGTGTCG
AY595	vanA	Vancomycin	GGGCTGTGAGGTCGGTTG	TTCAGTACAATGCGGCCGTTA
AY159	vanB_1	Vancomycin	TTGTCGGCGAAGTGGATCA	AGCCTTTTTCCGGCTCGTT
AY163	vanHB	Vancomycin	GAGGTTTCCGAGGCGACAA	CTCTCGGCGGCAGTCGTAT
AY381	vanTC_2	Vancomycin	ACAGTTGCCGCTGGTGAAG	CGTGGCTGGTCGATCAAAA
AY183	vanYD_1	Vancomycin	AAGGCGATACCCTGACTGTCA	ATTGCCGGACGGAAGCA

Appendix Table A-8 Sites from the UK National River Flow Archive (NRFA) used to access catchment descriptor data and from the DEFRA Hydrology Data Explorer used to measure rainfall.

	Catchment		Coordinates
Catchment Descriptor Sites (National River Flow Archive)	Coquet	Coquet at Usway Burn at Shillmoor	55.363168 , -2.1813847
		Coquet at Morwick	55.333099 , -1.6326910
		Coquet at Rothbury	55.308447 , -1.8959997
	Eden	Eden at Kirkby Stephen	54.481233 , -2.3534001
		Eden at Temple Sowerby	54.648240 , -2.6152002
		Eamount at Udford	54.666874, -2.660445
		Eden at Sheepmount	54.904876 , -2.9528598
Rainfall Guage Sites (DEFRA Hydrology Data Explorer)	Coquet	Linbriggs	55.350245 , -2.1703339
		Rothbury	55.308447 , -1.8959052
		Warkworth	55.346569 , -1.6298228
	Eden	Barras	54.504210 , -2.2413626
		Brackenber	54.569779 , -2.4315335
		Harsceugh Castle	54.779356 , -2.6082029
		Willow Holme	54.899425 , -2.9535108
		Linbriggs	55.350245 , -2.1703339

Appendix Table A-9 Downloaded Rainfall data from DEFRA Hydrology Data Explorer and calculated API5.

Catchment	**Site**	**Date**	Measured Rainfall (mm) at Days 1 to 5 before the time of sampling (Day of sampling =Pd)					**Calculated API5 (mm)**
			Pd-1	Pd-2	Pd-3	Pd-4	Pd-5	
Coquet	DEFRA (Linbriggs)	07/09/2020	3.4	0.4	3.8	0.4	9	1.860779
		19/11/2020	1	2.4	0.8	2.6	3.4	1.561249
		15/03/2020	0.2	0.2	8.2	2.8	11.2	1.81659
	DEFRA (Rothbury)	07/09/2020	4.2	1.2	0.4	0	7.2	1.802776
		19/11/2020	2	0.2	0.4	3.8	5.4	1.418626
		15/03/2020	0	0	7.6	3	15	1.792345
	DEFRA (Warkworth)	07/09/2020	1.6	0.4	0.8	0	5	1.229837
		19/11/2020	1.4	0	0.2	2	1.4	1.042833
		15/03/2020	0	0	3.4	1	5.2	1.140175
Eden	DEFRA (Barras)	02/09/2020	0	0	0	0	0.2	0.111803
		08/10/2020	11.8	1.2	0.8	1.8	41	3.080179
		01/03/2021	0.2	0	0	0	5.2	0.65192
	DEFRA (Brackenber)	02/09/2020	0	0	0	0	0.2	0.111803
		08/10/2020	11.2	1.8	2	2.2	29.6	3.020761
		01/03/2021	0	0	0.2	0	6	0.65192
		02/09/2020	0.09	0	0.28	0	0.41	0.375

	DEFRA (Haresceugh Castle)	08/10/2020	8.71	0.12	1.67	2.4	35.74	2.714084
		01/03/2021	0	0	0	0	3.89	0.493077
	DEFRA (Willow Holme)	02/09/2020	0	0	0	0	0.2	0.111803
		08/10/2020	7.6	0	1.8	1.2	43.8	2.67161
		01/03/2021	0	0	0	0	1.8	0.33541

Appendix Table A-10: Summary concentrations of bacteria cells (based on 16SrRNA data), faecal coliforms, human and ruminant derived bacteriodetes (HuBac and RuBac, respectively) (mean+/- standard deviation).

	Coquet				**Eden**			
	Cells	**Faecal Coliforms**	**HuBac**	**RuBac**	**Cells**	**Faecal Coliforms**	**Human Bacteria**	**Ruminant Bacteria**
	Copies/mL				**Copies/mL**			
A	6.37e+03± 1.37e+03	0.000299± 0.000175	0± 0	12.3± 4.62	1.60e+04± 7.48e+03	0.0686± 0.0225	20± 17.1	0± 0
B	2.27e+04± 3.30e+03	0.0394± 0.0216	8.69± 8.69	5.96± 2.98	4.41e+04± 2.20e+04	0.133± 0.0389	100± 13.9	641± 604
C	3.27e+04± 6.81e+03	0.0231± 0.00759	0± 0	8.58± 8.58	8.21e+04± 1.38e+04	0.1± 0.0368	37.6± 22.4	563± 293
D	5.89e+04± 1.57e+04	0.615± 0.511	32.2± 15.8	233± 118	5.88e+04± 1.17e+04	0.14± 0.0494	72.6± 7.95	817± 462
E	5.54e+04± 1.91e+04	0.121± 0.0672	16.5± 3.35	23.8± 9.52	8.79e+04± 1.46e+04	0.138± 0.0381	71.6± 15.9	1300± 906
F	4.69e+04± 1.35e+04	0.13± 0.056	170± 144	125± 64.6	7.26e+04± 1.47e+04	0.125± 0.0483	71.8± 23.9	1040± 664
G	4.87e+04± 1.11e+04	0.22± 0.102	98.4± 70.2	163± 47.4	6.73e+04± 2.77e+03	0.126± 0.0443	98± 46.7	1260± 710
H	7.20e+04± 9.06e+03	0.209± 0.0765	34.9± 8.55	232± 110	6.51e+04± 2.01e+04	0.119± 0.0734	144± 118	595± 275
I	5.71e+04± 2.00e+04	0.0817± 0.0372	19.2± 12.4	110± 95.5	3.75e+04± 4.35e+03	0.0865± 0.034	50± 18.6	233± 108
J	1.87e+05± 7.97e+04	0.128± 0.0213	49.2± 24.6	84.6± 43.4	8.79e+04± 3.17e+04	0.197± 0.126	40.2± 12.5	147± 76.2

K	1.05e+05± 3.19e+04	0.114± 0.00619	11.1± 11.1	225± 186				
L	9.65e+04± 4.63e+04	0.0901± 0.0475	12.1± 12.1	35.9± 35.9				

Appendix Table A-11 Bacteria cells/mL, MST markers human derived bacteriodetes (HuBac), and ruminant derived bacteriodetes (RuBac) and Faecal coliforms. Statistically significant differences between sampling month/between catchment are highlighted green.

	Coquet				**Eden**				**Eden vs. Coquet**
	March	September	November	*p-value*	March	September	October	*p-value*	*p-value*
Cells/mL	55400± 10500	88000± 7700	54000 ±10700	0.8473	60700± 9450	62200± 9740	62900 ±12700	0.9524	0.5493
HuBac/mL	13.9±15.3	76.1±13.8	23.0±22.4	0.537	45.9±35.2	50.2±28.8	115.9±104.3	0.0356	6.83_{x10-5}
RuBac/mL	13.8±24	155±197	146.2±122.4	0.00319	598.3±525.3	435±942	944±988	0.1755	0.00257
Faecal coliforms/mL	0.05320 ± 0.009740	0.125± 0.039	0.06540± 0.010100	0.02747	0.06540± 0.010100	0.0928± 0.009	0.21200 ± 0.033100	0.005818	0.02939

Appendix Table A-12 Measured Onsite parameters for the Coquet and Eden (mean +/- standard deviation)

	Coquet					**Eden**				
	Flow Rate (m^3/s) n=2	**Temperature (°C) n=3**	**pH n=3**	**DO (mg/L) N=3**	**Conductivity (n=3)**	**Flow Rate (m^3/s) n=2**	**Temperature (°C) n=3**	**pH n=3**	**DO (mg/L) N=3**	**Conductivity (n=3)**
A	3.70± 0.77	8.37± 2.07	7.5± 0.312	11.7± 0.496	112± 13.9	3.1± 2.7	10.7± 1.5	7.75± 0.174	11.4± 0.491	117± 30.7
B	8.35± 3.05	8.3± 2.29	7.64± 0.272	12± 0.574	121± 13.4	37.9± 32.1	10.5± 1.88	7.98± 0.0603	10.9± 0.507	390± 56
C	6.25± 2.08	8.83± 1.93	7.33± 0.12	11.2± 0.717	139± 15	27.1± 18.6	11.2± 1.83	8.24± 0.0722	11.1± 0.559	408± 42.5

D	5.42± 1.87	8.7± 2.15	7.54± 0.058 1	11± 0.693	194± 10.9	92.9± 84.1	11± 1.96	8.15± 0.0289	11.1± 0.5	413± 51.7
E	7.21± 1.72	9.23± 2.13	7.65± 0.017 3	11.1± 0.626	198± 11.6	73.6± 43.9	10.7± 1.71	8.1± 0.0578	11.1± 0.552	434± 57.5
F	NA	9.6± 2.25	7.78± 0.014 5	11.7± 0.928	203± 11.1	124.4± 47.9	11.4± 1.86	8.02± 0.102	11.1± 0.686	296± 36.8
G	4.79± 1.59	9.77± 1.94	7.77± 0.062 4	11.6± 0.722	204± 12.6	109.7± 40.9	11.4± 2.07	8± 0.0872	11± 0.554	273± 15.5
H	18.69 ± 8.87	9.57± 2.22	7.84± 0.063 5	11.3± 0.588	210± 10.6	87.1± 45.4	11.5± 1.99	7.93± 0.0867	11.1± 0.589	261± 18.7
I	5.92± 0.17	10.2± 2.25	8± 0.032 1	11.5± 0.522	212± 6.92	121.4± 45.4	11.4± 2.22	7.96± 0.0555	11± 0.714	266± 16.9
J	19.76 ± 8.57	9.77± 2.47	7.98± 0.057 7	11.4± 0.646	239± 4.58	119.4± 41.9	11.8± 1.96	7.91± 0.0567	11± 0.595	268± 20.9
K	13.37 ± 2.61	10.1± 2.42	8.04± 0.058 1	11± 0.897	236± 8.54					
L	17.44 ± 0.47	10.2± 2.78	7.94± 0.074 2	11.4± 0.891	502± 126					

Appendix Table A-13 Measured physiochemical parameters for the Coquet and Eden (mean +/- standard deviation).

	Coquet				Eden			
	COD (n=3)	NH_4 (n=3)	PO^3_4 (n=3)	TN (n=3)	COD (n=2-3)	NH_4-N (n=3)	PO_4-P (n=3)	TN (n=3)
	Mg/l				Mg/l			
A	18.4± 4.75	0.0125± 0.00647	0.037± 0.0161	2.16± 0.506	22.8± 7.62	0.0409± 0.0221	0.021± 0.0101	1.27± 0.179
B	23± 4.85	0.0121± 0.00665	0.0293± 0.0171	2.9± 0.024	20.1± 7.29	0.0521± 0.0135	0.01± 0.00755	2.59± 0.0764

C	22.8± 3.95	0.0102± 0.00818	0.000667± 0.000333	1.82± 0.377	17.9± 8	0.138± 0.0978	0.025± 0.0155	2.2± 0.143
D	19.4± 3.25	0.0153± 0.00896	0.01± 0.0052	1.62± 0.388	20.7± 6.53	0.098± 0.0645	0.0503± 0.0338	2.35± 0.0953
E	18.8± 4.35	0.0159± 0.00858	0.007± 0.00208	1.93± 0.683	20.2± 6.68	0.0648± 0.0249	0.0297± 0.0186	2.89± 0.43
F	18.5± 2.15	0.017± 0.00951	0.002± 0.001	1.1± 0.122	14.9± 7.26	0.06± 0.038	0.004± 0.00173	2.43± 0.409
G	19.9± 2.16	0.0116± 0.00431	0.0193± 0.0154	2.43± 0.493	13.4± 7.83	0.162± 0.138	0.057± 0.0417	2.48± 0.284
H	19.5± 0.404	0.0121± 0.00521	0.0467± 0.0235	1.39± 0.383	12.7± 7.93	0.228± 0.211	0.0307± 0.0172	2.14± 0.462
I	19.4± 3.91	0.0103± 0.0041	0.016± 0.00839	2.1± 0.514	13.4± 9.42	0.0478± 0.0301	0.0278± 0.0129	2.41± 0.395
J	21.6± 1.21	0.0164± 0.00348	0.042± 0.0161	1.59± 0.03	14.4± 5.98	0.0762± 0.0591	0.026± 0.0182	2.57± 0.607
K	21.1± 2.36	0.0169± 0.00748	0.071± 0.0323	1.35± 0.143				
L	21.2± 2.68	0.0145± 0.00559	0.0117± 0.00606	1.5± 0.163				

Appendix Table A-14 Most common ASV families assessed through quantitative microbial profiling (QMP).

Phylum	**Class**	**Order**	**Family**	**Genus**
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium (1)
Bacteroidota	Bacteroidia	Cytophagales	Spirosomaceae	Pseudarcicella
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	NA (1)
Bacteroidota	Bacteroidia	Flavobacteriales	Crocinitomicaceae	Fluviicola
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	NA (2)
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter (1)
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium (2)
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Rhodoferax

Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium (3)
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Rhodoferax
Actinobacteriota	Actinobacteria	Frankiales	Sporichthyaceae	hgcl_clade
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter(2)
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	NA (3)
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium (4)
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium (5)
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium (6)
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter (3)
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Limnohabitans
Actinobacteriota	Actinobacteria	Frankiales	Sporichthyaceae	hgcl_clade
Proteobacteria	Gammaproteobacteria	Burkholderiales	Burkholderiaceae	Polynucleobacter
Bacteroidota	Bacteroidia	Chitinophagales	Chitinophagaceae	Sediminibacterium
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium (7)
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	NA (4)
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium (8)
Verrucomicrobiota	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	uncultured

Appendix Table A-15 Summary absolute abundance (copies/mL), relative abundance (copies/16S rRNA) and diversity of genes (number of genes) in each antibiotic class for each site.

		Coquet											Eden									
		C/B	C/C	C/D	C/E	C/F	C/G	C/H	C/I	C/J	C/K	C/L	E/A	E/B	E/C	E/D	E/E	E/F	E/G	E/H	E/I	E/J
Aminoglycoside	**copies/mL**	3320± 5400	4030± 5880	79600± NA	2650± 3990	1240± 562	1210± 727	6460± 7610	872± 850	3080± 3250	2190± 578	7750± 5830	134± 87.8	22900± 14700	5800± 7290	1240± 89.1	18100± 13000	2830± 1530	18100± 20400	2100± 2020	1050± 800	7340± 4080
	copies/16SrRNA	0.0282 ± 0.0446	0.0272± 0.0406	0.339± NA	0.00788 ± 0.00972	0.0104± 0.0105	0.00625 ± 0.00412	0.019± 0.0202	0.00419 ± 0.00517	0.00665 ± 0.0067	0.00592 ± 0.00247	0.0322± 0.0233	0.0138± 0.0186	0.125± 0.0455	0.0145± 0.0152	0.00658 ± 0.00329	0.0576± 0.0538	0.0105± 0.00762	0.0694 ± 0.0818	0.0093± 0.0109	0.00756 ± 0.007	0.0209± 0.00829
	number of genes	5.33± 4.04	3.67± 3.79	11± NA	4± 0	3.5± 0.707	5.33± 3.21	3± 1.41	2.33± 1.53	5± 0	3.33± 0.577	4.67± 0.577	2± 1.41	10± 1.41	4.33± 2.08	5.5± 0.707	10.3± 0.577	4± 2	8± 2.65	7± 2.83	3.67± 0.577	5.67± 2.08
Beta Lactam	**copies/mL**	12100± 4800	10600± 14600	116000 ± NA	577± 927	326± 53.9	695± 259	2970± 4020	193± 159	8880± 1910	442± 290	11500± 5670	4990± 6670	52500± 47300	515± 705	100± 4.09	64400± 23500	531± 384	19400± 30500	2290± 2350	250± 333	3760± 2400
	copies/16SrRNA	0.117± 0.0165	0.0724± 0.1	0.493± NA	0.00162 ± 0.00236	0.00248 ± 0.00207	0.00469 ± 0.0025	0.00847 ± 0.0112	0.0011± 0.00098 7	0.02± 0.00234	0.00156 ± 0.0018	0.121± 0.169	0.0626± 0.0351	0.238± 2.45e-05	0.00126 ± 0.0015	0.00053 ± 0.00025	0.181± 0.0482	0.0022± 0.00225	0.0759 ± 0.121	0.00611 ± 0.00435	0.0019± 0.00271	0.0165± 0.0191
	number of genes	11± 2.83	8.5± 6.36	15± NA	3.67± 4.62	3.5± 0.707	9± 2.83	4± 4.24	3± 2	7± 0	4.33± 2.31	12.3± 2.52	4.5± 4.95	16± 0	2.67± 2.08	2± 1.41	15.3± 1.15	6± 2.65	11± 4.36	9.5± 4.95	4± 3.46	9± 4.36
Integrons	**copies/mL**	35400± 8970	32700± 39400	NA	17000± 18300	48700± 54300	15400± 6500	93400± 96300	20000± 7460	54500± 9440	26700± 9050	52500± 38600	21600± 5870	82400± 79200	33600± 24000	11900± 13700	149000± 71500	19100± 2010	60600± 41500	32500± 17900	12100± 7010	51500± 47100
	copies/16SrRNA	0.381± 0.0189	0.227± 0.265	NA	0.0654± 0.0389	0.49± 0.642	0.0812± 0.0385	0.283± 0.243	0.0952± 0.0369	0.123± 0.00915	0.0678± 0.02	0.236± 0.19	1.84± 2.39	0.357± 0.0373	0.0917± 0.0493	0.0466± 0.0447	0.411± 0.123	0.0712± 0.0306	0.228± 0.173	0.1± 0.00639	0.0832± 0.0596	0.177± 0.156
	number of genes	1± 0	1± 0	NA	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0	1± 0
MDR	**copies/mL**	1430± 1390	1550± 2120	26400± NA	581± 767	935± 901	775± 184	2060± 2150	549± 197	1930± 12.2	890± 192	2670± 2050	809± 583	14100± 12600	1320± 1130	434± 371	17400± 6000	1190± 614	11000± 15600	3660± 3130	570± 518	2860± 1160
	copies/16SrRNA	0.0137 ± 0.0101	0.0107± 0.0144	0.112± NA	0.002± 0.00175	0.00904 ± 0.0113	0.0045± 0.00254	0.00625 ± 0.00546	0.0027± 0.00127	0.00439 ± 0.00040 8	0.00267 ± 0.00173	0.0147± 0.0142	0.0328± 0.0308	0.0645± 0.00114	0.0035± 0.00222	0.00185 ± 0.00095 1	0.0478± 0.00317	0.00485 ± 0.00396	0.0428 ± 0.0621	0.0103± 0.00472	0.00409 ± 0.00436	0.00978 ± 0.00507
	number of genes	3.33± 1.53	3± 2	5± NA	2.67± 2.08	3± 0	4.33± 1.53	3± 0	3.67± 1.15	4.5± 0.707	3.67± 0.577	4.67± 0.577	2.5± 0.707	6± 0	3.67± 1.53	3.5± 0.707	6± 0	4.33± 1.53	5± 1	5± 0	4± 1.73	5.33± 0.577
MG	**copies/mL**	2470± 2080	2190± 2870	51800± NA	1520± 1630	3410± 2900	5990± 3090	4400± 2360	1610± 684	5150± 177	2780± 1170	2990± 2090	369± 211	23300± 18900	8960± 5340	2990± 1570	22400± 7570	4540± 2510	11900± 11100	5530± 3840	1590± 447	4980± 1840

	copies/16SrRNA	0.0242 ± 0.0163	0.0148± 0.0197	0.22± NA	0.00549 ± 0.00304	0.032± 0.0385	0.0292± 0.00943	0.0146± 0.00358	0.00803 ± 0.00456	0.0118± 0.00157	0.00776 ± 0.00412	0.018± 0.0194	0.0366± 0.0489	0.113± 0.0156	0.0251± 0.00775	0.0139± 0.00141	0.0617± 0.00923	0.0152± 0.00578	0.0451 ± 0.0451	0.0164± 0.00393	0.0105± 0.00284	0.0158± 0.00582
	number of genes	6± 2.65	5.33± 2.08	9± NA	6.33± 0.577	7± 0	7.67± 0.577	6.5± 2.12	6.67± 0.577	8± 1.41	7.33± 1.15	6.67± 0.577	3± 0	10± 0	8± 1	7.5± 2.12	9.67± 0.577	8.33± 1.15	9± 1	8± 0	8± 1	8± 1
MLSB	**copies/mL**	1840± 1690	1870± 2930	52300± NA	1180± 745	1800± 2240	2850± 1790	2320± 846	532± 272	2580± 275	1360± 417	5110± 4570	732± 626	18800± 14200	1990± 1520	557± 212	24600± 10900	1100± 546	13500± 18700	1660± 956	702± 554	3400± 1260
	copies/16SrRNA	0.0179 ± 0.0139	0.0129± 0.02	0.222± NA	0.00514 ± 0.00035 1	0.0187± 0.0254	0.0134± 0.00433	0.00789 ± 0.00045 6	0.00256 ± 0.00161	0.00591 ± 0.00121	0.00348 ± 0.00101	0.0194± 0.0128	0.0252± 0.0206	0.0944± 0.0206	0.0054± 0.00266	0.00268 ± 0.00016	0.0665± 0.0113	0.00389 ± 0.00215	0.0525 ± 0.0744	0.00508 ± 0.00048 8	0.00485 ± 0.00382	0.0107± 0.00322
	number of genes	5± 3.61	3.67± 3.06	10± NA	2.33± 2.31	3.5± 2.12	5.33± 3.51	2± 1.41	2.67± 1.53	5± 0	4± 1	6.33± 0.577	3.5± 2.12	9.5± 0.707	5± 2	5± 2.83	10± 0	6.33± 1.15	7.67± 0.577	7.5± 0.707	3.33± 1.15	7± 1
Other	**copies/mL**	1070± 1290	1020± 1140	8680± NA	491± 196	475± 584	1310± 1490	1180± 738	344± 149	1270± 139	626± 346	580± 567	75.1± 14.6	2310± 1570	878± 526	291± 190	3900± 1560	767± 504	1570± 1790	748± 352	383± 276	1540± 1000
	copies/16SrRNA	0.0131 ± 0.0173	0.00687 ± 0.00797	0.0369 ± NA	0.00234 ± 0.00048 9	0.00492 ± 0.00665	0.00543 ± 0.0044	0.00383 ± 0.00135	0.00167 ± 0.00088 7	0.0029± 0.00060 3	0.00142 ± 0.00024 2	0.00303 ± 0.00304	0.00616 ± 0.00788	0.0123± 0.00396	0.00245 ± 0.00094 5	0.00174 ± 0.00166	0.0106± 0.00137	0.00239 ± 0.00050 6	0.0060 2± 0.0071 9	0.00237 ± 6.71e-05	0.00264 ± 0.00193	0.00447 ± 0.00294
	number of genes	2.33± 1.15	2.5± 0.707	5± NA	1.67± 0.577	1.5± 0.707	3± 1	3± 0	2± 1	4± 0	1.67± 0.577	3± 2	1.5± 0.707	5± 0	2.33± 1.53	2± 0	4.33± 0.577	2± 0	3.33± 1.15	4± 1.41	2± 0	
Phenicol	**copies/mL**	148± 49	168± 155	4340± NA	133± 188	38.2± NA	166± 79.3	630± 807	81.2± 28.4	154± 5.37	123± 84.5	277± 168	125± NA	1870± 1370	540± 259	300± NA	2440± 543	184± 117	1440± 1620	411± 86.4	102± 143	395± 333
	copies/16SrRNA	0.0014 5± 0.0001 02	0.00112 ± 0.00111	0.0184 ± NA	0.00041 3± 0.00044	0.00013 5± NA	0.00102 ± 0.00074 3	0.00182 ± 0.00221	0.00052 2± 8.45e-05	0.00035 1± 4.71e-05	0.00044 7± 0.00050 6	0.0018± 0.00202	0.00113 ± NA	0.00959 ± 0.00242	0.00152 ± 0.00015 3	0.00206 ± NA	0.00688 ± 0.00093 7	0.00065 6± 0.00046 4	0.0055 3± 0.0065 3	0.00139 ± 0.00041 8	0.00078 6± 0.00116	0.00128 ± 0.00111
	number of genes	2± 0	1.5± 0.707	4± NA	2± 1	1± NA	2.67± 1.53	2± 1.41	1.5± 0.707	2± 0	1.67± 1.15	2.33± 0.577	1± NA	4± 0	2.5± 0.707	2± NA	3.33± 0.577	2.33± 0.577	3.33± 0.577	3± 0	1.67± 1.15	3± 0
Quinolone	**copies/mL**	2100± 2080	3620± 3980	43900± NA	1610± 2550	570± 105	1220± 809	4710± 5850	818± 690	3040± 1220	1130± 407	4530± 3240	817± 854	17400± 13000	2500± 3510	312± 223	24700± 5190	1690± 1030	12400± 14700	4470± 3360	765± 774	4080± 1570
	copies/16SrRNA	0.0201 ± 0.0167	0.0243± 0.0279	0.187± NA	0.00467 ± 0.00636	0.00436 ± 0.00369	0.00773 ± 0.0063	0.0137± 0.0158	0.00443 ± 0.00428	0.00679 ± 0.00208	0.00365 ± 0.00301	0.0214± 0.0182	0.0211± 0.0117	0.0883± 0.0204	0.00605 ± 0.00755	0.00138 ± 0.00045 5	0.07± 0.00972	0.00695 ± 0.00634	0.0476 ± 0.0589	0.013± 0.00411	0.00554 ± 0.00643	0.0142± 0.00762

	number of genes	3± 1.73	2.5± 0.707	4± NA	1.33± 0.577	1± 0	2± 1	1± 0	1.67± 0.577	2.5± 0.707	1.67± 0.577	3± 1	2.5± 0.707	3.5± 0.707	1.33± 0.577	2± 0	3.67± 0.577	2± 1	3.33± 0.577	3± 0	1.67± 0.577	3± 1
Sulfonamide	**copies/mL**	143± 126	90.1± 42.7	4310± NA	224± 84.2	185± 171	444± 312	310± 13.4	179± 69.1	401± 34.6	311± 117	325± 246	46± NA	1850± 2040	508± 180	393± 92.5	2760± 1260	399± 448	1650± 1970	407± 497	136± 135	381± 289
	copies/16SrRNA	0.00163± 0.00168	0.000549± 0.000196	0.0183 ± NA	0.00108 ± 0.000253	0.00177 ± 0.00218	0.00214 ± 0.000911	0.00111 ± 3e-04	0.000921± 0.00052	0.00091 ± 1.21e-05	0.00081 ± 0.000315	0.00135 ± 0.000971	0.000415± NA	0.00709 ± 0.00287	0.00152 ± 0.00039 3	0.00195 ± 0.00041 2	0.00748 ± 0.00147	0.00113 ± 0.000956	0.00634± 0.00788	0.00102 ± 0.00106	0.000948± 0.000913	0.00104 ± 0.000255
	number of genes	2± 1.41	1.5± 0.707	3± NA	2.33± 0.577	1± 0	2.67± 0.577	1.5± 0.707	2.33± 0.577	3± 0	2± 1	3± 0	2± NA	4± 0	2.33± 0.577	2± 0	4± 0	2± 0	3± 0	2± 1.41	1.67± 0.577	2.33± 1.15
Taxonomic	**copies/mL**	NA	129± NA	262± NA	15.2± NA	NA	20.7± NA	37.1± NA	NA	NA	48.8± 34.5	42.1± NA	30.7± NA	97.1± 42.9	228± NA	28.8± NA	230± 121	NA	234± 168	82.35± 38.65	20.8± NA	52.8± NA
	copies/16SrRNA	NA	0.000885± NA	0.0011 2± NA	0.000125± NA	NA	0.000152± NA	0.000163± NA	NA	NA	0.000145± 4.27e-05	6.57e-05± NA	0.00422 ± NA	0.000595± 0.000342	0.000901± NA	0.000197± NA	0.000605± 0.000166	NA	0.00089± 0.000701	0.000246± 0.000035	0.000111± NA	8.68e-05± NA
	number of genes	NA	2± NA	1± NA	1± NA	NA	1± NA	1± NA	NA	NA	1± 0	1± NA	1± NA	2± 0	1± NA	1± NA	2.67± 0.577	NA	2.5± 2.12	1± NA	1± NA	1± NA
Tetracycline	**copies/mL**	664± 653	646± 774	15600± NA	528± 577	1590± 1870	2470± 1690	1210± 616	450± 233	1930± 373	714± 538	1400± 1180	273± 161	6210± 5880	1200± 655	651± 64.1	6350± 1670	770± 389	3270± 3000	1530± 1080	312± 119	1470± 469
	copies/16SrRNA	0.00636± 0.00523	0.00435 ± 0.0053	0.0664 ± NA	0.00189 ± 0.0011	0.0162± 0.0216	0.0114± 0.00563	0.00402 ± 0.000869	0.00221 ± 0.00145	0.00436 ± 0.000412	0.00189 ± 0.00129	0.00649 ± 0.0056	0.0127± 0.013	0.0273± 0.00208	0.00339 ± 0.00089	0.00333 ± 0.00115	0.0178± 0.00127	0.00267 ± 0.00126	0.0124 ± 0.0122	0.00454 ± 0.00113	0.00218 ± 0.00112	0.00471 ± 0.00169
	number of genes	3.67± 2.52	2.33± 1.15	7± NA	3.67± 1.53	3.5± 0.707	6.67± 1.53	2± 0	2.67± 0.577	6± 1.41	4± 3	5± 2	2.5± 2.12	7.5± 0.707	4.67± 0.577	6.5± 2.12	8± 1	4.67± 0.577	6.33± 0.577	5.5± 0.707	4.67± 1.15	5.33± 3.06
Trimethoprim	**copies/mL**	156± 124	290± NA	1740± NA	NA	NA	13.4± NA	NA	NA	NA	NA	129± 125	28.1± NA	879± 613	NA	NA	1440± 479	NA	2320± 3180	262± NA	NA	97.2± 22.9
	copies/16SrRNA	0.00143± 0.000848	0.00198 ± NA	0.0074 1± NA	NA	NA	9.86e-05± NA	NA	NA	NA	NA	0.000519± 0.000389	0.000253± NA	0.0046± 0.00137	NA	NA	0.00422 ± 0.00163	NA	0.00911± 0.0125	0.000609± NA	NA	0.000334± 0.000209
	number of genes	2± 0	2± NA	2± NA	NA	NA	1± NA	NA	NA	NA	NA	1.67± 0.577	1± NA	2± 0	NA	NA	1.67± 0.577	NA	1.5± 0.707	2± NA	NA	2± 0
Van	**copies/mL**	1570± 1000	1630± 2250	22400± NA	476± 627	82.8± NA	383± 131	1170± 1560	88.1± 79.5	1160± 413	164± 38.2	2210± 1760	787± NA	11000± 8470	1210± NA	41.9± NA	13300± 3240	265± 276	8290± 12900	1590± 1780	187± 96.5	912± 405

	copies/16SrRNA	0.0148 ± 0.00602	0.0112± 0.0154	0.0952 ± NA	0.00135 ± 0.00152	0.00029 3± NA	0.00258 ± 0.0013	0.00334 ± 0.00433	0.00046 ± 0.000489	0.00259 ± 0.000679	0.00051 1± 0.000412	0.00898 ± 0.00638	0.00709 ± NA	0.0547± 0.0109	0.00273 ± NA	0.00015 2± NA	0.0374± 0.00409	0.00118 ± 0.00147	0.0324 ± 0.0511	0.00411 ± 0.00355	0.00133 ± 0.000983	0.00366 ± 0.00273
	number of genes	4.5± 0.707	3± 2.83	5± NA	3± 2.83	2± NA	4.5± 0.707	2± 1.41	1.33± 0.577	4± 0	1.33± 0.577	4.33± 0.577	5± NA	5± 0	3± NA	1± NA	5± 0	2± 0	4± 1	3± 1.41	2.5± 0.707	3.67± 1.53

Appendix Table A-16 average relative abundance (copies/16S rRNA), absolute abundance (copies/mL), diversity (number of genes) of ARGs/MGEs per month per catchment and Kruskal Wallis significance tests (p<0.05). Green cells indicate statistically significant values.

		Coquet					**Eden**					Coquet vs Eden *P-value*
		March	September	November	Sampling Month *P-value*	Mean ± sd	March	September	October	Sampling Month *P-value*	Mean ± sd	
ARG	Relative abundance (copies/16SrRNA) (mean ± sd)	0.132± 0.178	0.084± 0.116	0.0539± 0.0551	0.69	0.090± 0.13	0.213± 0.243	0.212± 0.336	0.172± 0.264	0.34	0.198 ±0.270	0.075
	Absolute abundance (copies/mL) (mean ± sd)	17200± 14600	18100± 22400	12700± 13200	0.62	15989 ±16699	56500± 81600	67400± 105000	33200± 47000	0.76	51771 ± 78316	0.16
	Diversity (number of genes) (mean ± sd)	40.7± 16.2	24.3± 15.6	32.6± 14.5	0.06	32.5 ±16.3	45.3± 20.4	46.8± 22.7	44.9± 19.6	0.91	45.6 +/ - 20	0.0207
MGE	Relative abundance (copies/16SrRNA) (mean ± sd)	0.221± 0.183	0.28± 0.325	0.132± 0.108	0.61	0.211± 0.224	0.631± 1.12	0.22± 0.202	0.147± 0.152	0.05	0.337 ±0.687	0.499
	Absolute abundance (copies/mL) (mean ± sd)	44600± 49200	46300± 32500	28800± 23200	0.51	39892± 36060	69700± 56500	72000± 84100	33100± 37400	0.17	57749± 61519	0.4549

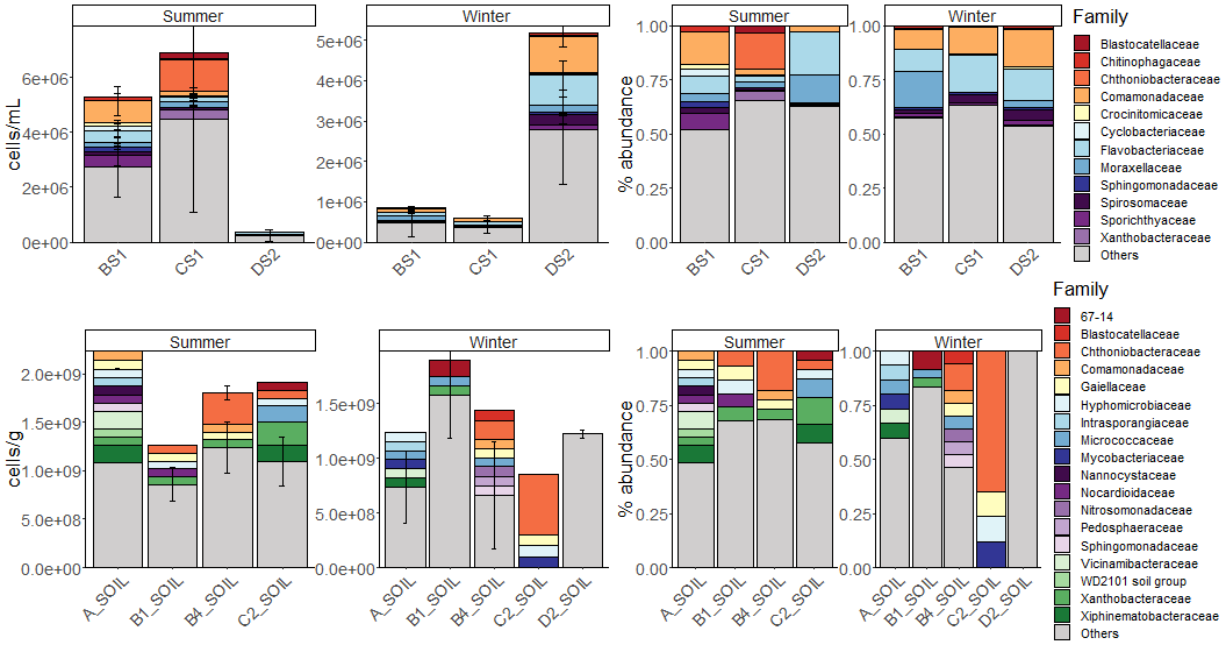
Appendix Table A-17 Network Analysis: Most highly correlating ARGs, MGEs (absolute abundance (copies/mL), taxa (by order level) (quantitative microbial abundance (QMP)/mL) and MST markers (Ruminant *Bacteroidetes* and Human *Bacteroidetes*; RuBac and HuBac) by degree of connectivity (calculated by Gephi) for both catchments, the Coquet and Eden.

	ARGs/MGEs	degrees	Order/MST	Degrees
Both Catchments	blaTEM	3	Bacteroidales	3
	cfxA	2	Acidaminococcales	1
	blaVIM	2	Aeromonadales	1
	tetO_2	2	Azospirillales	1
	qnrS2	1	Babeliales	1
	tetQ	1	Cyanobacteriales	1
	tetW	1	Cytophagales	1
	blaCMY_2	1	Fibrobacterales	1
	tetPB_1	1	Methanomassiliicoccales	1
	sul4	1	Nitrosopumilales	1
			Silvanigrellales	1
Coquet catchment	tetM	16	Aeromonadales	3
	aacC4	5	Bacteroidales	3
	erm36	4	Candidatus_Azambacteria	3
	mcr1	4	Pirellulales	3
	blaOXA48	3	Babeliales	2
	cfxA	3	Cellvibrionales	2
	cmxA	3	Chthoniobacterales	2
	arr3	3	Frankiales	2
	oleC	2	KD4-96	2
	blaNDM	2	Lactobacillales	2
	vanYD_1	2	Phycisphaerales	2
	ermB_3	2	Planctomycetales	2
	mdtH	2	Sphingomonadales	2
	qacF/H	1	Acetobacterales	1
	tet39	1	Acidaminococcales	1
	lnuC	1	Bathyarchaeia	1
	aadA_1	1	Blastocatellales	1
	tetW	1	Campylobacterales	1
	aac(6)-iic	1	Candidatus_Nomurabacteria	1
	sul3_1	1	Candidatus_Zambryskibacteria	1
	msrE	1	Chitinophagales	1
	blaTEM	1	Chlamydiales	1
	qnrB	1	Cyanobacteriales	1

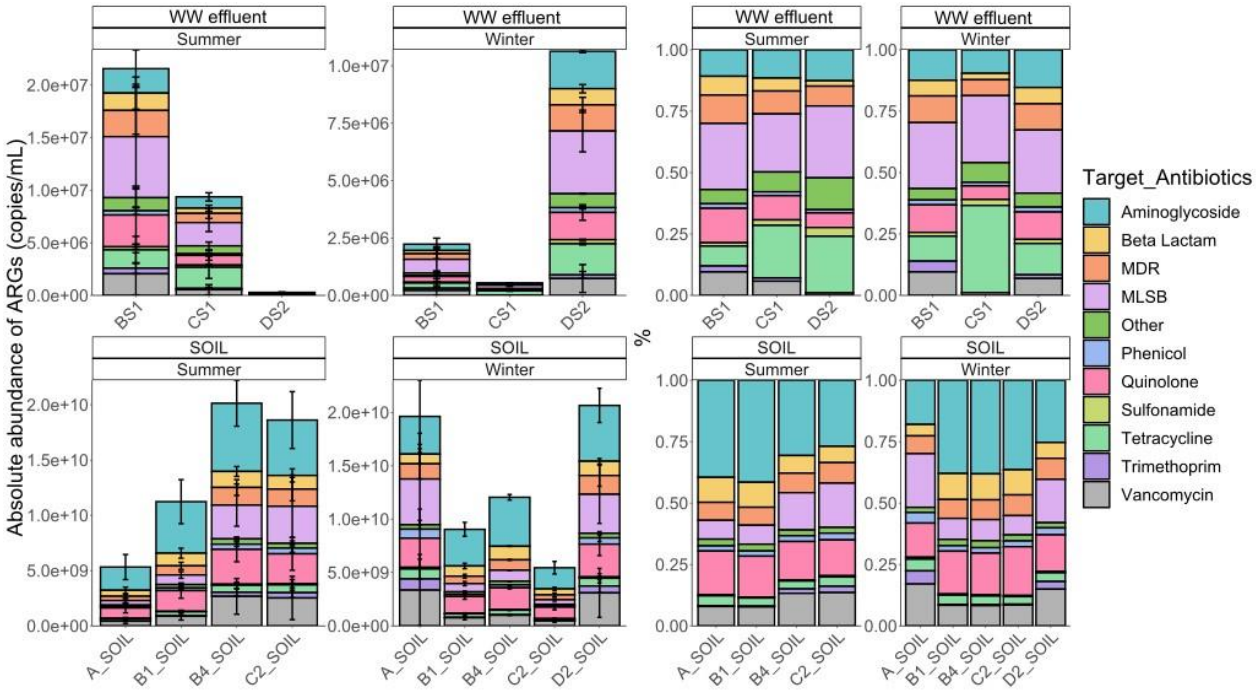
			Desulfitobacteriales	1
			Diplorickettsiales	1
			Flavobacteriales	1
			Geobacterales	1
			Gracilibacteria	1
			Lachnospirales	1
			Methanobacteriales	1
			Micrococcales	1
			Nannocystales	1
			Nitrososphaerales	1
			Paracaedibacterales	1
			Propionibacteriales	1
			Pseudomonadales	1
			Rhizobiales	1
			Rokubacteriales	1
			Silvanigrellales	1
			Sphingobacteriales	1
			Spirochaetales	1
			Steroidobacterales	1
			Verrucomicrobiales	1
			Xanthomonadales	1
Eden Catchment	strB	12	HuBac_mL	1
	tetQ	10	RuBac_mL	1
	tetW	8	Bacteroidales	7
	aadA_1	4	Acidaminococcales	6
	tetPB_1	4	Campylobacterales	6
	blaSHV11	3	Peptostreptococcales-Tissierellales	6
	cfxA	3	Clostridiales	5
	lnuC	3	Pseudomonadales	5
	oleC	3	Aeromonadales	4
	vat(A)	3	Rhodobacterales	4
	aacC4	2	Burkholderiales	3
	blaCMY_2	2	Erysipelotrichales	3
	sul2_2	2	Flavobacteriales	3
	sul4	2	Lachnospirales	3
	aadA10	1	Azospirillales	2
	ant4-ib	1	Christensenellales	2
	blaKPC	1	Micrococcales	2
	carB	1	Oscillospirales	2

	dfrA25	1	R7C24	2
	mcr1	1	Spirochaetales	2
	msrE	1	Bacillales	1
	qnrS2	1	Candidatus_Nomurabacteria	1
	tetM	1	Cellvibrionales	1
	tnpA_5	5	Chitinophagales	1
	tnpA_1	4	Cytophagales	1
	ISAba3	3	Deinococcales	1
	ISPps	2	Gammaproteobacteria_Incertae_Sedis	1
	IncP_oriT	1	Lactobacillales	1
			Methanobacteriales	1
			Methanomassiliicoccales	1
			Nitrosopumilales	1
			Nitrososphaerales	1
			Omnitrophales	1
			Verrucomicrobiales	1
			Vicinamibacterales	1
			Xanthomonadales	1

Appendix B



Appendix Figure B-1 Bar plots showing the 25 most abundant ASVs grouped by families, with remaining pooled as 'other'. Figures show the wastewater effluent samples (top row), and the soil samples (bottom row)



Appendix Figure B-2 Bar plots showing the absolute ARG abudnance for wastewater effluent (top row) and soil (bottom row).

Appendix Table B-1 Location of sample sites

Cluster	Site		Northing	Easting
Control (clean site)	A		55.363683,	-2.184614
	A soil	100m	55.363979,	-2.186750
		200m	55.364001,	-2.188026
Cluster B	B1		55.312831,	-1.958716
	B2		55.313780,	-1.953614
	BS1		55.320595,	-1.955053
	B3		55.318928,	-1.958282
	B4 soil	100m	55.306640,	-1.936820
		200m	55.311646,	-1.961147
	B4		55.305222,	-1.934570
Cluster C	C1		55.309135,	-1.891346
	CS1		55.304616,	-1.890016
	C2 soil	100m	55.302656,	-1.889995
		200m	55.303891,	-1.888699
	C2		55.302344,	-1.886153
Cluster D	D1		55.296581,	-1.709643
	DS2		55.299476,	-1.700201
	D2 soil	100m	55.298408,	-1.699770
		200m	55.298155,	-1.705259
	D2		55.298785,	-1.696386

Appendix Table B-2 Summary of samples collected

Sample Code	Sample Type	March 7/03/2022 10/03/2022	March/April 21/0/2022 14/4/2022	July 11/07/2022 14/07/2022	July 25/07/2022 28/07/2022	August 08/08/2022 11/08/2022
A	River water	✓	✓	✓	✓	✓
A soil	Soil	✓	✓	✓	✓	✓
B1	River water	✓	✓	✓	✓	✓
B1 soil	Soil	✓	✓	✓	✓	✓
B2	River water	✓	✓	✓	✓	✓
B3	River water		✓			
B4	River water	✓	✓	✓	✓	✓
B4 soil	Soil	✓	✓	✓	✓	✓
BS1	Sewage effluent	✓	✓	✓	✓	✓
C1	River water	✓	✓	✓	✓	✓
CS1	Sewage effluent	✓	✓	✓	✓	✓
C2	River water	✓	✓	✓	✓	✓
C2 soil	Soil	✓	✓	✓	✓	✓
D1	River water	✓	✓	✓	✓	✓
DS1	Sewage effluent	✓	✓	✓	✓	
D2	River water	✓	✓	✓	✓	✓
D2 soil	Soil	✓	✓			

Appendix Table B-3 Top 25 Taxa River Water Samples

Phylum	Class	Order	Family	Genus
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Simplicispira
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Limnohabitans
Bacteroidota	Bacteroidia	Cytophagales	Spirosomaceae	Pseudarcicella
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter
Cyanobacteria	Cyanobacteriia	Cyanobacteriales	Phormidiaceae	Tychonema CCAP 1459-11B
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	NA
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
Bacteroidota	Bacteroidia	Flavobacteriales	Crocinitomicaceae	Fluviicola
Bacteroidota	Bacteroidia	Sphingobacteriales	env.OPS 17	NA
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter
Bacteroidota	Bacteroidia	Chitinophagales	Chitinophagaceae	Sediminibacterium
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
Actinobacteriota	Actinobacteria	Frankiales	Sporichthyaceae	NA
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingorhabdus
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
Actinobacteriota	Actinobacteria	Micrococcales	Microbacteriaceae	Rhodoluna
Bacteroidota	Bacteroidia	Flavobacteriales	Crocinitomicaceae	Fluviicola
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Limnohabitans
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	Candidatus Udaeobacter

Actinobacteriota	Actinobacteria	Frankiales	Sporichthyaceae	Candidatus Planktophila
Proteobacteria	Gammaproteobacteria	Methylococcales	Methylomonadaceae	Crenothrix

Appendix Table B-4 Top 25 Taxa Sewage Effluent samples

Phylum	**Class**	**Order**	**Family**	**Genus**
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	NA
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Limnohabitans
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	Candidatus Udaeobacter
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	Bradyrhizobium
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	Candidatus Udaeobacter
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter
Bacteroidota	Bacteroidia	Cytophagales	Spirosomaceae	Pseudarcicella
Actinobacteriota	Actinobacteria	Frankiales	Sporichthyaceae	NA
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingorhabdus
Bacteroidota	Bacteroidia	Cytophagales	Cyclobacteriaceae	Algoriphagus
Bacteroidota	Bacteroidia	Chitinophagales	Chitinophagaceae	Sediminibacterium
Actinobacteriota	Actinobacteria	Frankiales	Sporichthyaceae	Candidatus Planktophila
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Simplicispira
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Limnohabitans
Bacteroidota	Bacteroidia	Flavobacteriales	Crocinitomicaceae	Fluviicola

Acidobacteriota	Blastocatellia	Blastocatellales	Blastocatellaceae	JGI 0001001-H03
Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	NA
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	Candidatus Udaeobacter
Actinobacteriota	Actinobacteria	Frankiales	Sporichthyaceae	hgcI clade
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Aquabacterium

Appendix Table B-5 Top 25 Taxa soil samples

Phylum	**Class**	**Order**	**Family**	**Genus**
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	Candidatus Udaeobacter
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	Candidatus Udaeobacter
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Xiphinematobacteraceae	Candidatus Xiphinematobacter
Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	Pseudolabrys
Actinobacteriota	Actinobacteria	Micrococcales	Micrococcaceae	Pseudarthrobacter
Actinobacteriota	Thermoleophilia	Solirubrobacterales	67-14	NA
Actinobacteriota	Thermoleophilia	Gaiellales	Gaiellaceae	Gaiella
Actinobacteriota	Actinobacteria	Corynebacteriales	Mycobacteriaceae	Mycobacterium
Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Pedomicrobium
Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	Piscinibacter
Planctomycetota	Phycisphaerae	Tepidisphaerales	WD2101 soil group	NA
Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Hyphomicrobium
Acidobacteriota	Vicinamibacteria	Vicinamibacterales	Vicinamibacteraceae	Vicinamibacter
Myxococcota	Polyangia	Nannocystales	Nannocystaceae	Enhygromyxa
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Xiphinematobacteraceae	Candidatus Xiphinematobacter
Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	NA
Actinobacteriota	Actinobacteria	Micrococcales	Intrasporangiaceae	NA

Actinobacteriota	Actinobacteria	Propionibacteriales	Nocardioidaceae	Nocardioides
Acidobacteriota	Vicinamibacteria	Vicinamibacterales	Vicinamibacteraceae	NA
Actinobacteriota	Thermoleophilia	Gaiellales	Gaiellaceae	Gaiella
Verrucomicrobiota	Verrucomicrobiae	Pedosphaerales	Pedosphaeraceae	NA
Proteobacteria	Gammaproteobacteria	Burkholderiales	Nitrosomonadaceae	Ellin6067
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	Chthoniobacter
Acidobacteriota	Blastocatellia	Blastocatellales	Blastocatellaceae	JGI 0001001-H03

Appendix Table B-6 Results of the FEAST analysis for each sink including paired t-test to idenitfy differences in the percentage contribution seasonally (summer v winter) and between the estimates based on the microbiome and resistome. The significance threshold was set at p<0.05

	Site		Summer/Winter	Microbiome		Resistome		T Test Microbiome vs Resistome
				Percentage contribution	T test (p value)	Summer vs. Winter	T test (p value)	
Sink B4	B1	RW	Summer	37.8 ± 19	0.221	72.1 ± 5.96	0.5118	0.07967
		RW	Winter	4.88 ± 3.32		43.6 ± 29.8		
	B1_SOIL	SOIL	Summer	0.0298 ± 0.0298	0.3	7.42 ± 5.52	0.423	0.2133
		SOIL	Winter	0.73 ± 0.359		1.9 ± 0.325		
	B2	RW	Summer	35.5 ± 19.3	0.8798	13.7 ± 4.38	0.4264	0.8364
		RW	Winter	29.5 ± 29		51.7 ± 30.4		
	B3	RW	Winter	0.0391 ± NA		0.693 ± NA		
	B4_SOIL	SOIL	Summer	0.0354 ± 0.0118	0.1761	5.84 ± 5.78	0.5149	0.3223
		SOIL	Winter	0.551 ± 0.15		1.31 ± 0.181		
	BS1	SE	Summer	5.89 ± 3.05	0.4442	0.946 ± 0.781	0.9311	0.1562
		SE	Winter	25.3 ± 16.2		1.08 ± 1.08		
	Unknown		Summer	20.7 ± 9.76	0.4363	2.16e-67 ± 2.16e-67	0.4226	0.03061
			Winter	39 ± 15.7		2.19e-235 ± 0		
Sink C2	C1	RW	Summer	31.6 ± 28.6	0.5954	31 ± 5.46	0.2157	0.8925
		RW	Winter	13.7 ± 4.11		20.7 ± 3.67		
	C2_SOIL	SOIL	Summer	0.134 ± 0.107	0.3376	0.0975 ± 0.0567	0.5142	0.3786
		SOIL	Winter	0 ± 0		2.23 ± 2.23		
	CS1	SE	Summer	44.1 ± 23	0.9559	68.9 ± 5.4	0.6763	0.2116
		SE	Winter	46.7 ± 34.8		71.9 ± 3.59		
	Unknown		Summer	24.2 ± 17.9	0.7125	7.97e-30 ± 7.97e-30	0.4909	0.1193

			Winter	39.6 ± 30.7		5.18 ± 5.03		
Sink D2	D1	RW	Summer	28.9 ± 1.67	0.929	68.7 ± 11.8	0.3576	0.001652
		RW	Winter	30.9 ± 17.7		82.7 ± 3.74		
	D2_SOIL	SOIL	Winter	0.793 ± 0.793		1.44 ± 0.111		0.5651
	DS2	SE	Summer	5.57 ± 5.57	0.6814	41.6 ± 9.54	0.1933	0.05758
		SE	Winter	2.5 ± 2.21		15.8 ± 3.63		
	Unknown		Summer	67.3 ± 5.06	0.9337	3.54 ± 3.5	0.4175	8.28e-05
			Winter	65.8 ± 14.7		6.29e-51 ± 6.28e-51		

Appendix Table B-7 Measured on site parameters

Cluster	Site		Calculated flow rate	Temp	pH	Conductivity	DO
A	A	Summer	1.68± 0.6	19.5± 2.05	8.66± 0.00577	221± 9	9.76± 0.397
		Winter	7.38± 5.11	8.45± 0.495	8.94± 0.134	136± 24.7	12.9± 0.0566
B	B1	Summer	4.78± 1.19	18.9± 1.91	8.54± 0.0889	269± 7.51	9.11± 0.455
		Winter	42.5± 28.3	7.65± 3.04	8.56± 0.276	184± 33.3	12.7± 1.11
	B2	Summer	1.67± 0.79	17± 1.65	8.32± 0.11	432± 5.51	9.24± 0.411
		Winter	2.45± 1.53	7.15± 1.34	8.48± 0.106	416± 40.3	13.1± 0.424
	B3	Winter		8.2± NA	8.43± NA	441± NA	12.6± NA
	B4	Summer	3.08± 1.32	20.7± 3.3	8.32± 0.0755	296± 2.31	9.61± 0.291
		Winter	5.62± 1.28	6.75± 0.354	8.38± 0.148	238± 41.2	13± 1.1
	BS1	Summer		18.5± 1.15	7.64± 0.117	607± 37.6	6.87± 0.184
		Winter		7.7± 0.283	7.99± 0.325	746± 142	8.97± NA
C	C1	Summer	1.4± 0.164	16.9± 0.737	8.34± 0.169	306± 7.02	9.06± 0.225
		Winter	5.89± 0.631	10.6± 0.212	8.5± 0.247	238± 20.5	11± 0.24
	C2	Summer	0± 0	17.5± 1.5	8.04± 0.278	336± 7.37	9.24± 0.344
		Winter	0± 0	9.8± 0.283	8.27± 0.332	254± 28.3	11.3± 0.184
	CS1	Summer		17.8± 2.06	28.8± 37.3	597± 5.51	5.03± 0.446
		Winter		10.3± 0.141	7.42± 0.262	700± 189	7.12± 1.72
D	D1	Summer	2.38± 0.23	19.1± 1.74	8.63± 0.0265	332± 9.45	10.2± 0.723
		Winter	9.93± 1.03	10.6± 0.0707	8.52± 0.0707	251± 1.41	11.8± 0.0849
	D2	Summer	NA	18.6± 0.778	8.3± 0.0141	332± 10.6	9.71± 0.0849
		Winter	0± 0	11.9± 0.778	8.38± 0.141	260± 2.12	11.7± 0.0778
	DS2	Summer		18.2± 0.636	7.32± 0.0212	820± 4.95	5.27± 0.509
		Winter		11.4± 0.778	7.48± 0.389	711± 5.66	7.5± 0.962

Appendix Table B-8 Physiochemical characteristics of each site

Cluster	Site	Winter_Summer	COD	NH4-N	TN	PO4-P
A	A	Summer	62.1± 16.3	0.0133± 0.00577	1.14± 0.626	0.00667± 0.00577
		Winter	57.2± 2.45	0.005± 0.00707	0.935± 0.106	0.01± 0.0141
B	B1	Summer	64.3± 9.62	0.01± 0	1.22± 0.537	0.00333± 0.00577
		Winter	71.3± 6.54	0.005± 0.00707	1.37± 0.41	0.03± 0.0283
	B2	Summer	83.3± 20.8	0.0267± 0.00577	2.26± 0.38	0.0367± 0.00577
		Winter	65.7± 12.3	0.005± 0.00707	3.14± 0.983	0.025± 0.00707
	B3	Winter	59.6± NA	0.01± NA	2.81± NA	0.09± NA
	B4	Summer	61.8± 17.6	0.0167± 0.00577	1.42± 0.812	0.00333± 0.00577
		Winter	62.1± 1.45	0.03± 0.0283	1.62± 0.0354	0.005± 0.00707
	BS1	Summer	86.8± 7.14	0.273± 0.0551	40.6± 3.93	2.5± 0.286
		Winter	127± 56.7	0.21± 0.0424	24.4± 1.91	2.01± 0.226
C	C1	Summer	61.6± 15.6	0.02± 0.01	1.12± 0.249	0.00667± 0.00577
		Winter	76.5± 8.77	0.005± 0.00707	1.28± 0.495	0.01± 0.0141
	C2	Summer	61.3± 7.95	0.0867± 0.0577	2.36± 0.882	0.223± 0.15
		Winter	68.2± 1.84	0.055± 0.0495	2.14± 1.12	0.11± 0.0849
	CS1	Summer	143± 2.47	1.42± 0.452	28.8± 5.74	2.66± 0.548
		Winter	162± 24.4	1.32± 0.53	22.4± 0.212	2.09± 0.0354
D	D1	Summer	59.2± NA	0.0133± 0.00577	1.04± 0.343	0.01± 0.01
		Winter	74.6± 21.6	0.01± 0	1.37± 0.537	0.005± 0.00707
	D2	Summer	57± 17.5	0.02± 0.01	0.967± 0.154	0.03± 0.02
		Winter	71.1± 2.26	0.01± 0	1.38± 0.431	0.03± 0.0283
	DS2	Summer	178± NA	0.775± 0.488	34.4± 6.47	2.74± 0.403
		Winter	170± 10.6	0.5± 0.099	26.6± 4.6	2.27± 0.269

Appendix Table B-9 E. coli, total coliforms, ESBL E. coli and ESBL coliforms (colony forming units (CFU)/mL)

			E. coli (CFU/mL)	Coliforms (CFU/mL)	ESBL *E. coli* (CFU/mL)	ESBL coliforms (CFU/mL)
A	A	Summer	1.26± 0.678	6.14± 1	0± 0	0.04± NA
		Winter	0.63± 0.849	3.3± 3.08	0± 0	0.43± 0.453
B	B1	Summer	7.75± 2.54	15.5± 1.5	12.7± 11	7.31± 11.6
		Winter	1.3± 0.424	3.15± 1.34	0.025± 0.00707	0.26± 0.0849
	B2	Summer	11.7± 5.45	17.6± 7.32	24.2± 13.5	16.6± 13.5
		Winter	3.6± 1.03	6.73± 2.4	0.055± 0.00707	0.78± 0.099
	B3	Winter	27.1± NA	39.9± NA	0.19± NA	1.56± NA
	B4	Summer	3.23± 0.466	13.3± 4.53	40.6± 42.1	0.42± 0.0424
		Winter	2.08± 0.0707	3.82± 0.495	0.06± 0.0424	0.385± 0.106
	BS1	Summer	347± 281	1150± 742	7.53± 7.27	8.51± 4.43
		Winter	437± 269	1140± 636	2.24± 2.73	38.2± 49.7
C	C1	Summer	3.63± 1.61	13.6± 4.45	76.6± 55.2	0.807± 0.862
		Winter	1.74± 0.615	3.78± 0.87	0.035± 0.0212	0.235± 0.0495
	C2	Summer	52.7± 7.79	164± 50.3	2.51± 1.3	1.09± 0.0929
		Winter	6.16± 8.72	16.5± 23.3	0.15± NA	0.64± NA
	CS1	Summer	787± 407	3260± 1780	2.03± 0.376	23.7± 14.6
		Winter	833± 603	2000± 863	13.7± 9.43	41.3± 21
D	D1	Summer	3.99± 3.36	11.4± 1.49	10.7± NA	0.12± NA
		Winter	3.45± 1.63	8.12± 1.2	0.05± 0.0566	0.705± 0.417
	D2	Summer	5.73± 1.82	24.2± 11	21.6± 23.3	0.713± 0.232
		Winter	3.58± 1.25	9.55± 1.34	0.045± 0.0354	0.705± 0.417
	DS2	Summer	340± 9.43	1390± 311	7.31± 3.54	20.4± 6.18
		Winter	460± 283	1640± 99	10.9± 1.77	33.2± 11.4

Appendix Table B-10 N:P ratios in the Coquet sites (mean ± standard deviation)

Cluster	Site	N:P Ratios
A	A	32.1 ± 24.8
B	B1	28 ± 14.9
	B2	28.6 ± 11.6
	B3	10.2
	B4	43.7 ± 14.3
C	C1	35.2 ± 9.68
	C2	5.83 ± 2.86
D	D1	34.6 ± 19.4
	D2	19.2 ± 14.4

Appendix C

Appendix Table C-1 Sources of the data input into the model

Data type	Resolution	Source
Digital Elevation Model	12.5m	ALOS PALSAR
Soil	1km	Various (see Appendix Table C-2)
Land-use	100m	CORINE land cover (2018)
Climate	Rainfall (3 locations, see Appendix Table C-4) Temperature (1 location, see Appendix Table C-4)	DEFRA Hydrology Data Explorer Met Office Historic Station Data (Met Office 2023)

Appendix Table C-2 Soil types identified in the catchment.

	Dystric Gleysols	Humic Gleysols	Dystric Cambisols	Gleyic Luvisol	Placic Podzol	Dystric Histosol	Reference
Type of Soil	Clay Loam	Clay loam/Sandy Clay Loam/Clay Loam	Clay Loam	Clay loam/Clay	Clay	Sandy Loam	USDA soil texture triangle
Geological Series	Brickfield Series	Wilcocks Series	Belmont Series	Dunkeswick Series	Winter Hill Series	Malvern Series	(ESDB 2006)
Soil Hydraulic Group	C	C	C	C	D	C	Estimated
porosity	55	45	52	44	52	88	(Batjes 1997)
Depth from soil surface layer (L1/L2/L3)	200/500/1000	200/500/1000	200/500	200/600/1000	200/700/1200	200/600/1150	(LandIS 2023)
Available water capacity at surface (L1/L2/L3)	118	125	132	95	114	480	(Batjes 1997)
Saturated Hydraulic Conductivity	5/5/5	5/5.5/5	5/5	5/5/0.55	0.55/0.55/0.55	50/50/50	(Maidment 1993)
Organic carbon content (%) (L1/L2/L3)	1.5/0.34/0.34	3.04/0.6/06	1.94/0.4	0.64/0.2/0.2	6.67/1.3/1.3	35/37.4/37.4	(Batjes 1997)
Clay Content (L1/L2/L3)	35/35/35	35/25/35	33/33	35/35/60	60/60/60	10/10/10	(LandIS 2023)
Silt content (L1/L2/L3)	35/35/35	35/15/35	33/33	35/35/20	20/20/20	20/20/25	(LandIS 2023)
Sand content (L1/L2/L3)	30/30/30	30/60/30	33/33	30/30/20	20/20/20	70/70/65	(LandIS 2023)
Rock fragment content (L1/L2/L3)	0	0/5	0	0	0	0	(Batjes 1997; LandIS 2023)
Moist soil albedo ratio (L1/L2/L3)	0.05/0.05/0.5	0.05/0.08/0.05	0.05/0.05	0.05/0.05/0.02	0.02/0.02/0.02	0.09/0.09/0.09	(Van Wijk and Scholte Ubing 1963)
USLE equation soil eroadability (K) factor m^3 metric ton cm	0.26/0.26/0.26	0.26/0.32/0.26	0.26/0.26	0.26/0.26/0.26	0.19/0.19/0.26	0.27/0.27/0.27	(David 1988)

Appendix Table C-3 land-use categories for SWAT and CORINE land-cover database

SWAT		**CORINE land-cover**		
SWAT category	SWAT label	Label1	Label2	Label3
Residential	URBN	Artificial surfaces	Urban fabric	Discontinuous urban fabric
Residential-Low Density	URLD	Artificial surfaces	Artificial, non-agricultural vegetated areas	Sport and leisure facilities
Agricultural Land-Generic	AGRL	Agricultural areas	Arable land	Non-irrigated arable land
Pasture	PAST	Agricultural areas	Pastures	Pastures
Agricultural Land-Generic	AGRL	Agricultural areas	Heterogeneous agricultural areas	Land principally occupied by agriculture, with significant areas of natural vegetation
Forest-Deciduous	FRSD	Forest and semi natural areas	Forests	Broad-leaved forest
Forest-Evergreen	FRSE	Forest and semi natural areas	Forests	Coniferous forest
Forest-Mixed	FRST	Forest and semi natural areas	Forests	Mixed forest
Range Grasses	RNGE	Forest and semi natural areas	Scrub and/or herbaceous vegetation associations	Natural grasslands
Forest-Mixed	FRST	Forest and semi natural areas	Scrub and/or herbaceous vegetation associations	Moors and heathland
Forest-Mixed	FRST	Forest and semi natural areas	Scrub and/or herbaceous vegetation associations	Transitional woodland-shrub
Range Grasses	RNGE	Forest and semi natural areas	Open spaces with little or no vegetation	Sparsely vegetated areas
Wetlands-Mixed	WETL	Wetlands	Inland wetlands	Inland marshes
Wetlands-Mixed	WETL	Wetlands	Inland wetlands	Peat bogs
Water	WATR	Water bodies	Inland waters	Water bodies
Water	WATR	Water bodies	Marine waters	Estuaries

Appendix Table C-4 Sources for Climatic data for the model

Data	Source	Site Name	Coordinates
Daily precipitation (mm)	Rainfall Guage Sites (DEFRA Hydrology Data Explorer; https://environment.data.gov.uk/hydrology/landing)	Linbriggs	55.350245 , -2.1703339
		Rothbury	55.308447 , -1.8959052
		Warkworth	55.346569 , -1.6298228
Monthly maximum and minimum temperature (°C)	Met Office Historic Station Data (Met Office, 2023)	Durham	54.76786, -1.58455

Appendix Table C-5 Approximation of the kg of manure applied in each subbasin per hectare per day. Calculations assume 3.5kg of manure produced per sheep per day and 27.5 kg of manure produced from one cattle per day

		Number of animals per subbasin		Number of animals/hectare		Kg of manure per hectare per day	
Subbasin	**Area (hectares)**	**Sheep**	**Cattle**	**Sheep**	**Cattle**	**Sheep**	**Cattle**
1	6173	17037	0	2.76	0.00	9.66	0.00
2	13	47	0	3.59	0.00	12.56	0.00
3	2158	5956	0	2.76	0.00	9.66	0.00
4	1188	4246	515	3.57	0.43	12.50	11.92
5	1467	5105	111	3.48	0.08	12.18	2.08
6	1241	3731	649	3.01	0.52	10.52	14.39
7	516	1850	0	3.59	0.00	12.56	0.00
8	2074	7441	0	3.59	0.00	12.56	0.00
9	338	1212	53	3.59	0.16	12.56	4.35
10	358	1185	57	3.31	0.16	11.59	4.35
11	3023	8596	478	2.84	0.16	9.95	4.35
12	567	1707	288	3.01	0.51	10.54	13.99
13	1246	4167	541	3.34	0.43	11.70	11.94
14	882	3150	383	3.57	0.43	12.50	11.94
15	2317	7748	8	3.34	0.00	11.70	0.10
16	1699	5116	864	3.01	0.51	10.54	13.99
17	2044	6406	1219	3.13	0.60	10.97	16.40
18	1359	3306	41	2.43	0.03	8.51	0.83
19	1341	3263	5	2.43	0.00	8.51	0.10
20	491	1513	213	3.08	0.43	10.78	11.94
21	2234	6880	970	3.08	0.43	10.78	11.94
22	1	4	0	3.08	0.43	10.78	11.94
23	538	1266	233	2.36	0.43	8.24	11.94
24	11	33	5	3.08	0.43	10.78	11.94
25	144	443	62	3.08	0.43	10.78	11.94
26	1829	3656	114	2.00	0.06	7.00	1.72
27	1001	2558	447	2.56	0.45	8.95	12.28
28	703	1797	16	2.56	0.02	8.95	0.62
29	9	23	4	2.56	0.45	8.95	12.28
30	56	144	25	2.56	0.45	8.95	12.28
31	2758	8188	1150	2.97	0.42	10.39	11.46
32	655	1676	293	2.56	0.45	8.95	12.28
33	833	2770	362	3.33	0.43	11.64	11.94
34	1747	5083	758	2.91	0.43	10.19	11.94
35	1706	3410	704	2.00	0.41	7.00	11.35
36	1575	3955	650	2.51	0.41	8.79	11.35
37	1242	3120	513	2.51	0.41	8.79	11.35
38	1993	5005	823	2.51	0.41	8.79	11.35
39	2868	8345	278	2.91	0.10	10.19	2.66
40	2628	7096	288	2.70	0.11	9.45	3.01
41	4297	12655	1984	2.94	0.46	10.31	12.69

Appendix Table C-6 National River Flow Archive (NRFA; https://nrfa.ceh.ac.uk/) Sites and Coordinates

NRFA Site	Coordinates
Coquet at Morwick	55.333099 , -1.6326910
Coquet at Rothbury	55.308447 , -1.8959997

Appendix Table C-7 List of Sites with Observed E. coli and ESBL E. coli data. All sites were sampled as part of the sampling campaign for Chapter 4 unless otherwise specified.

Site	Coordinates	Observed Dates
A	55.363683, -2.184614	15/03/2021 (Site A, Chapter 3) 07/03/2022 21/03/2022 11/07/2022 25/07/2022 08/08/2022
B1	55.312831, -1.958716	07/03/2022 21/03/2022 11/07/2022 25/07/2022 08/08/2022
B2	55.313780, -1.953614	07/03/2022 21/03/2022 11/07/2022 25/07/2022 08/08/2022
B4	55.305222, -1.934570	15/03/2021 (Site D, Chapter 3) 07/03/2022 21/03/2022 11/07/2022 25/07/2022 08/08/2022
C1	55.309135, -1.891346	15/03/2021 (Site E, Chapter 3) 10/03/2022

		14/04/2022 14/07/2022 28/07/2022 11/08/2022
C2	55.302344, -1.886153	15/03/2021 (Site F, Chapter 3) 10/03/2022 14/04/2022 14/07/2022 28/07/2022 11/08/2022
D1	55.296581, -1.709643	15/03/2021 (Site I, Chapter 3) 10/03/2022 14/04/2022 14/07/2022 28/07/2022 11/08/2022
D2	55.298785, -1.696386	10/03/2022 14/04/2022 14/07/2022 28/07/2022 11/08/2022