Wet Weather Overflow Monitoring Program 2016 to 2024

Synthesis report

Sydney WATER

Wet Weather Overflow Monitoring Program 2016 to 2024 Synthesis Report.

Version 1 - finalised 24 June 2024 - incorporating comments from final peer-review panel workshop (number 22)

Sydney Water Project Team: Michele Cassidy, Colin Besley, Nathan Harrison

SW 73 05/24.

For more info email multimedia@sydneywater.com.au

Executive summary

Sydney Water customers value healthy waterways with clean beaches and estuaries. Our long-term goal is to cost-effectively minimise the impact of our activities and contribute to a more liveable Sydney. For the past 20 years, Sydney Water has been working to meet wet-weather overflow (WWO) frequency targets/limits. Frequency targets, set in 2000 by the NSW Environment Protection Authority (EPA), aimed to reduce the number of WWOs over a 10-year period. The intention of this target was to recover swimming days and/or protect ecosystem health. Until 2020, capital investment mainly involved building large storage tanks, tunnels and bigger pipes and pumps. To meet the frequency targets across the four major coastal catchments (of Environment Protection Licence numbers: North Head 378, Bondi 1688, Malabar 372, and Cronulla 1728), an investment of more than AUD $18 billion (2022) would be required. It could take over 350 years to achieve those licence targets, based on the current source control spend rate. This investment would result in increased costs to our customers and localised inconvenience during construction.

In more recent years, Sydney Water and the EPA have worked to develop a suitable risk-based framework to inform a revised regulatory measure for the management of WWOs in the four major coastal wastewater systems. Under the Pollution Reduction Program (PRP) 307 of the North Head, Malabar, Bondi, and Cronulla EPLs, Sydney Water is required to implement a risk prioritisation methodology to identify the overflows (or emergency relief structures (ERSs)) of highest risk to waterway ecosystem and public health values, and undertake abatement works. From 2020, to meet our licence targets, Sydney Water has transitioned from storage to source control management of infiltration and inflow into the wastewater system. The 2020 – 2024 interim target aimed to reduce volume and/or frequency, however, these EPLs will change from 2024, to achieving a 6% reduction (or 1.6 GL) in WWO volumes. Sydney Water has estimated that all source control works would be completed by 2060 and this approach has been estimated to be 17x cheaper and more time efficient than using storage to abate WWOs.

The EPA has applied a continuous improvement initiative to the risk prioritisation methodology (Pollution Study (PS) 307). PS 307, due by 30 June 2027, states that Sydney Water is obligated '*to resolve technical issues and uncertainties (that is, limitations and assumptions) with the current prioritisation methodology through identification and use of new and improved tools and information to achieve a refined methodology*'. Sydney Water has committed to continuously improve our process of defining the highest risk WWOs, which includes undertaking scientific investigations to better understand the influence of WWOs on the receiving environment. In 2016, Sydney Water commenced the wet-weather overflow monitoring program (WWOM) to address the assumptions and limitations of current ecological and human health inputs in the risk prioritisation methodology. This was the commencement of a nine-year applied research program that has been summarised in this report.

What is a WWO?

The wastewater and stormwater systems are two independent separated systems in the Sydney region. Sydney's wastewater system is designed with emergency relief structures (ERSs). ERSs allow excess wastewater to overflow into stormwater drains or waterways during wet-weather, protecting homes and businesses from flooding. Without these designed overflow points, wastewater could backup through toilets and floor waste drains or gully traps in backyards. Hence, ERSs are a necessary component of the Sydney wastewater system designed to protect public health by stopping sewage backup into homes and business (Bickford et al., 1999).

Most of the wastewater systems in the Sydney region are laid in impermeable rock or clay trenches that often serve as channels where stormwater can follow sewer pipes and enter through faults (Bickford et al., 1999). During light rainfall, the stormwater volume in channels and the depth of the receiving water environment may rise, but there is normally no change in the performance of the wastewater system. Wastewater system pipes are typically designed to carry three times the average expected dry weather flow. Under moderate to heavy rainfall, stormwater can enter (via inflow/infiltration) into the wastewater system increasing flows within the pipes by 5 to 10 times more than in dry weather. This change in dilution of sewer influent under differing magnitude rainfall events was documented for the Sydney wastewater system under a sub-study of this WWOM project (Section 4.4, Besley et al., 2023).

A risk assessment of chemicals in wet-weather discharges from wastewater treatment plants and WWO spills from 11 major ERSs of the Sydney region, evaluated 114 chemicals of potential concern in the late 1990s (Bickford et al. (1999). That assessment documented influent collected from sewer carrier pipes to be toxic under WWO conditions, while no conclusions could be drawn on toxicity in downstream receiving waters (Bickford et al., 1999). Ammonia and nitrate were identified as contaminants of concern in sewer influent of WWOs. This risk assessment also identified that the predominant loading of metals (copper, zinc and silver) to receiving waters was from stormwater transport (Bickford et al., 1999). A subsequent investigation of highly urbanised sub-catchments of the Sydney estuary (Davis and Birch, 2009) showed that the contaminants copper, lead and zinc were predominantly (79-87%) derived from diffuse sources (residential properties and roads). While Birch (2024) recently identified road-derived metals as the chief contributor of metals to stormwater from a review and critical assessment of over three decades of research supplemented by global studies.

Peer-review guided applied research

The Bickford et al. (1999) study is one of the few published studies of separate sanitary sewer/stormwater systems and was based on the Sydney region. A literature (Google Scholar) search illustrated that there was considerable published research on combined sewer overflows (CSOs). The scarcity of published scientific literature for sanitary sewer systems prompted the WWOM project team to pursue a research program guided by leading scientists in their field. An independent expert peer-review panel of five members was appointed to provide assurance that the outcomes of the program would be robust, withstand scrutiny and adequately inform our Executive, Board, the EPA and stakeholders. Each subject matter expert has a specific environmental or human health specialty (see Section 1.3 for more detail). The role of the expert

panel was to review our implementation and application of monitoring tools, analysis of data and approach for meeting the proposed improvement levels.

The WWOM program applied techniques and methodologies new to Sydney Water. To implement these new techniques Sydney Water harnessed leading scientific knowledge and expertise by collaborating with universities and the Commonwealth Scientific and Industrial Research Organisation (CSIRO) to better understand how WWOs influence the potential for adverse ecological response and risk to human health. The use of these new techniques in assessing WWOs from a sanitary sewerage system had not been previously reported in the international scientific literature. Study outcomes have been collaboratively published in 16 journal articles in high-impact international scientific journals and two articles have been submitted. Outcomes from these 18 articles form the basis of this synthesis report.

Conceptual model diagrams and overview of outcomes

The WWOM program comprised four areas of research into the potential influences of WWOs (Figure i). The corresponding overarching research questions were:

- Section 2: Does spill rate and/or volume contribute to sewage-derived gross pollutants being spilt to receiving waters?
- Section 3: Is human or animal faecal contamination the dominant source (in the receiving waters)?
- Section 4: Are there any contaminants of potential concern in WWOs?
- Section 5: Are adverse ecological effects of WWOs apparent?

The synthesis report Sections 2, 3, 4 and 5 discuss outcomes and provide recommendations from these four research areas, respectively.

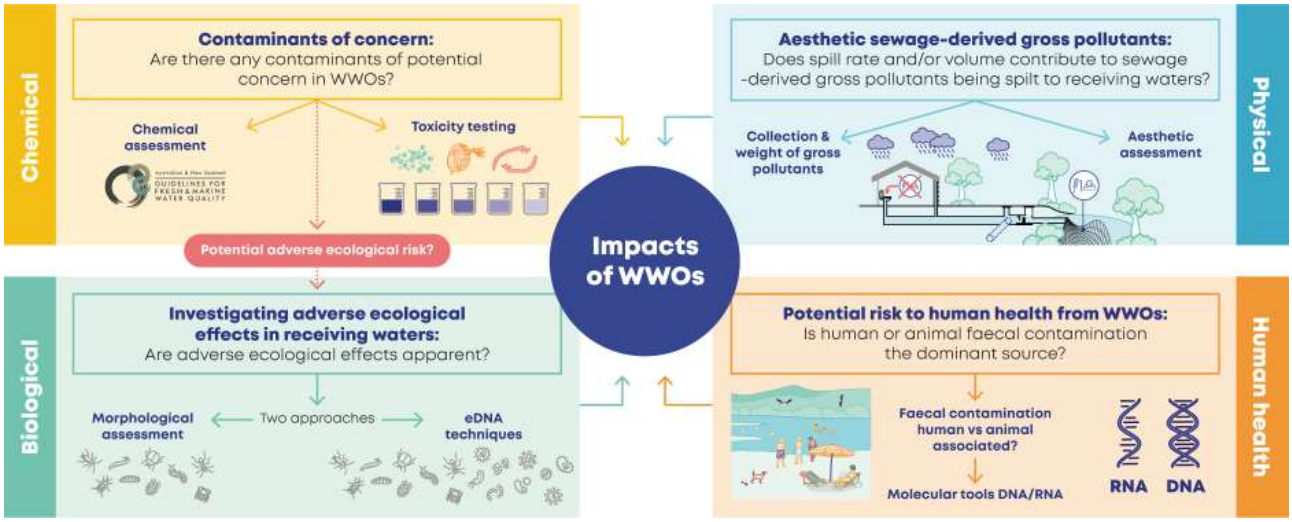


Figure i: Summary of WWOM program showing components investigated and overarching research questions

In conjunction with the peer-review panel, a conceptual model diagram has been raised for each of these four research areas. These conceptual models are described below.

Gross pollutants

Support for periodic community education campaigns (such as, Toilet Blockers Anonymous) for the proactive management of gross pollutants appears necessary, as cited literature in Section 2.1 predicts an upward trend in usage of wet wipe products.

Results from the study of incorrect disposal of gross pollutants (including wet wipes) into the wastewater system suggested that if future screening with modified trash nets (Section 2.2.4) was implemented, it would reduce the presence of sewage-derived gross pollutants and would minimise the risk of public contact with sewage-derived gross pollutants.

A potential reactive management approach is outlined in Section 2.2.4, that recommends selection of ERSs with relatively higher overflow volumes, such as from siphonic overflows that have very high spills rates (> 1000 L/s, Figure ii), would yield the highest capture of sewage-derived gross pollutants. While selection of gravity-fed ERSs that have relatively high spill rates (> 150 L/s) and associated relatively higher overflow volumes (Figure ii) would yield the next highest capture.

A permanent in-pipe screening approach, as discovered in reviewing literature while preparing a scientific publication for this study, may be an option for the highest volume siphonic overflows, such as those to be located within Sydney Airport lands discharging to the Mill Stream (a fringing waterway), to ameliorate the spread of sewage-derived gross pollutants through the receiving environment to reduce current manual labour clean-up effort after WWOs.

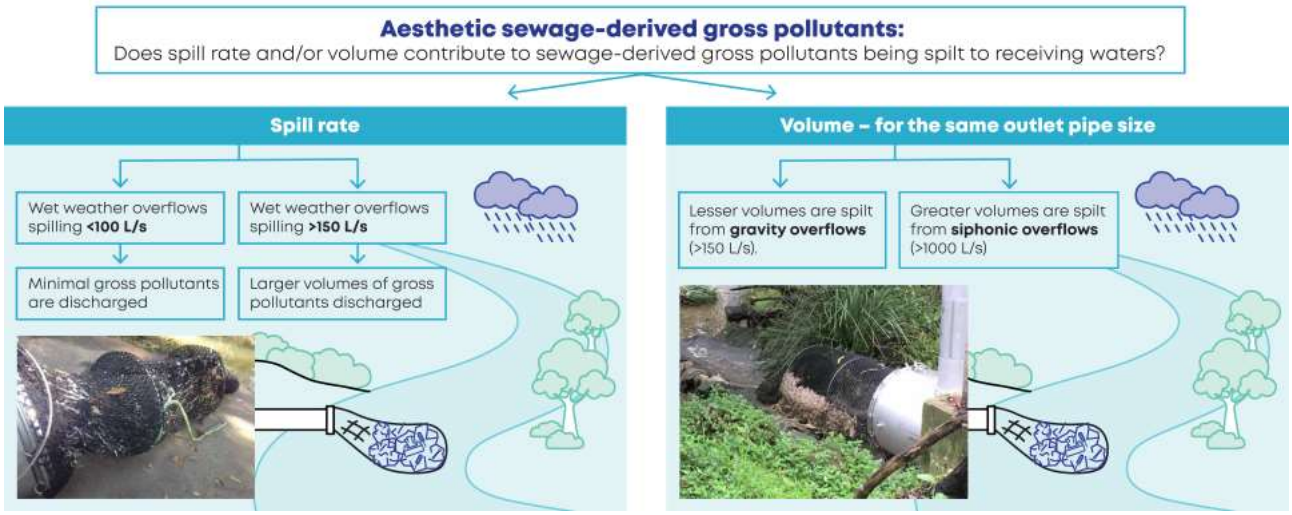


Figure ii: Conceptual model illustrating differing yields of sewage-derived gross pollutants influenced by ERS spill rate

Potential risk to human health from WWOs

In 1993, Sydney Water adopted monitoring with the faecal indicator bacteria, enterococci. A weakness of enterococci is the inability to distinguish human- from animal-faecal contamination. The advent of microbial source tracking with molecular marker genes does not have this limitation and has allowed new insights into tracking effects of WWOs in receiving waters. Capability uplift of Laboratory Services under the WWOM has resulted in three human faecal-associated marker genes (HFMGs) available in-house and capability is being expanded with animal-associated marker genes.

At commencement of the WWOM program in 2016, limited research had been conducted on application of monitoring with microbial source tracking in receiving waterways, with no literature available for sanitary (separate to the stormwater system) WWOs. Hence under the WWOM project, an extensive research program was undertaken with CSIRO that spanned 11 sub-studies (Section 3.2).

In the Sydney region, microbial source tracking with four HFMGs successfully detected human faecal-associated contamination in urban estuarine waters after WWOs at each of the 13 studied estuarine locations and at the freshwater study location (Sections 3.2.4, and 3.2.8). This microbial source tracking also identified leakage of sewage under dry weather (Section 3.2.4). Of the animal faecal-associated molecular marker genes assessed, widespread bird faecal contamination was identified under both dry- and wet-weather, but not at levels that reflect risk to human health.

An outline of when HFMGs may reflect risk to human health from WWO spills is presented in the flow chart (Figure 8-1) drawn from key observations from the eleven human health sub-studies (Section 3.2). These sub-studies informed a QMRA model that established site-specific Risk Based Thresholds (RBTs) for four HFMGs (HF183, Lachno3, crAssphage and PMMoV) (Section 3.3). RBTs are aligned to a benchmark level of GI illness risk, for fresh (day 0) and aged (day 1 to day 10) sewage contamination. Future direct measurement of human faecal-associated marker genes (HFMGs) in the water column during/after a WWO event would provide data to rank departures from RBTs. These departures would then provide an input into the WWOA risk prioritisation methodology to allow for a categorisation of each ERS based on highest illness risk to lowest illness risk (Sections 3.4 and 8.2.2).

The inherent flexibility of RBTs allows HFMG concentration results collected from any day of the 11-day window to be compared (Figure iii and Figure iv). This 11-day sampling window to measure RBTs also minimises the need to expose human resources to under bad weather (unsafe) conditions on the day or day after a WWO. It also avoids the need to collect samples within 24 – 48 hrs (holding time), depending on analyte, for valid results from laboratory processing, which is a problem with autosampler collection. The sampling window afforded by the RBTs helps reduce pressure on limited human resources and be available for reactive event monitoring, such as for keeping the drinking water system safe.

Future direct measurement of HFMGs in the water column during/after a WWO event would provide data to calculate departures above RBTs. These departures would then be ranked to provide a direct input into the WWOA risk prioritisation methodology and allow for a categorisation of each ERS assessed based on highest illness risk (greatest departure from a RBT) to lowest illness risk (further outlined in Sections 3.4 and 8.2.2).

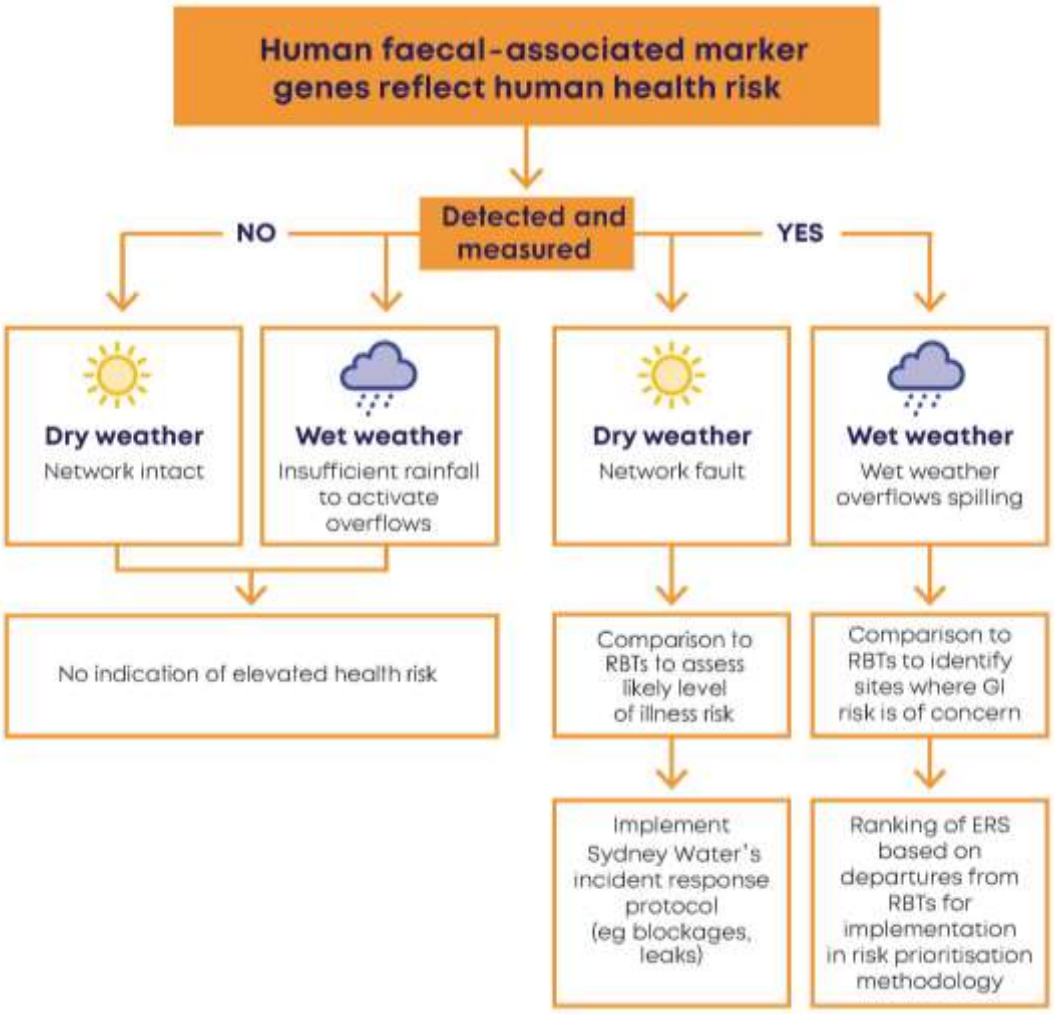


Figure iii: Flow chart of key observations from the human health pilot studies outlining when HFMGs may reflect risk to human health from WWOs

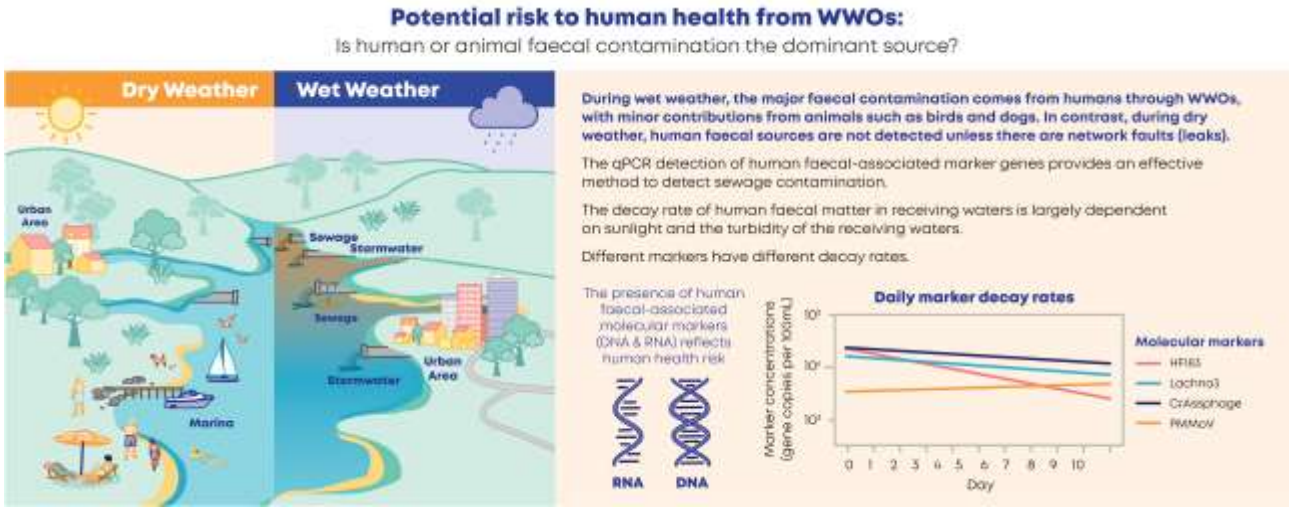


Figure iv: Conceptual model raised from the human health pilot studies

Contaminants of concern

The study of water quality at an atypical node of ERSs spilling into the same reach of Vineyard Creek, determined ammonia as a contaminant of 'potential' concern while ammonia tracked at four other study sites was below a level of concern (Section 4.4). Toxicity testing of influent collected under three weather conditions identified ammonia to be the primary cause of toxicity in both influent and receiving water samples from Vineyard. Toxicity was observed to reduce under increasing infiltration and inflow of stormwater into the sewage system (Section 4.5.1). Ammonia concentration results from Vineyard evaluated against the ANZG (2018) guideline value (for protection of 95% of species) best represented toxicity testing outcomes (Section 4.5.1) and the risk of adverse ecological effect was observed to reduce under increased dilution from inflow and infiltration of rainwater ingress into the sewer system across the weather events studied (Figure v, Section 4).

While toxicity from copper (Cu) and zinc (Zn) was also observed, loadings of these metals in stormwater is much greater (about 80%) than within influent, which suggests metals are of secondary concern in WWOs. Treatment of metals in influent without treatment of stormwater sources is highly unlikely to have an environmental benefit and may take decades as illustrated by the banning of lead in petrol and the slow reduction of lead in estuarine sediments in the Sydney region (Birch, 2024).

The WWOM project strategy to gauge a range of ERSs to represent a cross section of low, medium, and high volume and frequency WWOs (based upon hydraulic modelling) has provided a real-world insight into functioning of the Sydney sewerage system that was not available from the Bickford et al. (1999) review of the Sydney region. Under the WWOM, we determined that WWO spill durations of less than 6 h comprised just over 60% of WWO spill events. While WWO spill durations of 24 h or less comprised over 90% of WWO spill events (Section 4.5.2).

Follow up toxicity testing that better mimicked typical WWO spill durations (Section 4.5.2), based upon 24 h and 6 h pulse exposures of wet-weather influent, established required dilutions greater than 1 in 2 to reduce toxicity (Section 4.5.3). This indicates that urban streams and estuarine waters receiving WWOs need to be of an adequate capacity to achieve a greater than 2x dilution of a WWO spill to remove the risk of an adverse ecological effect (Figure v) based upon the sensitive water flea *Ceriodaphnia dubia* toxicity test outcomes. This is the same test species as used for Sydney Water's toxicity testing of inland wastewater treatment plant effluent to meet EPL obligations.

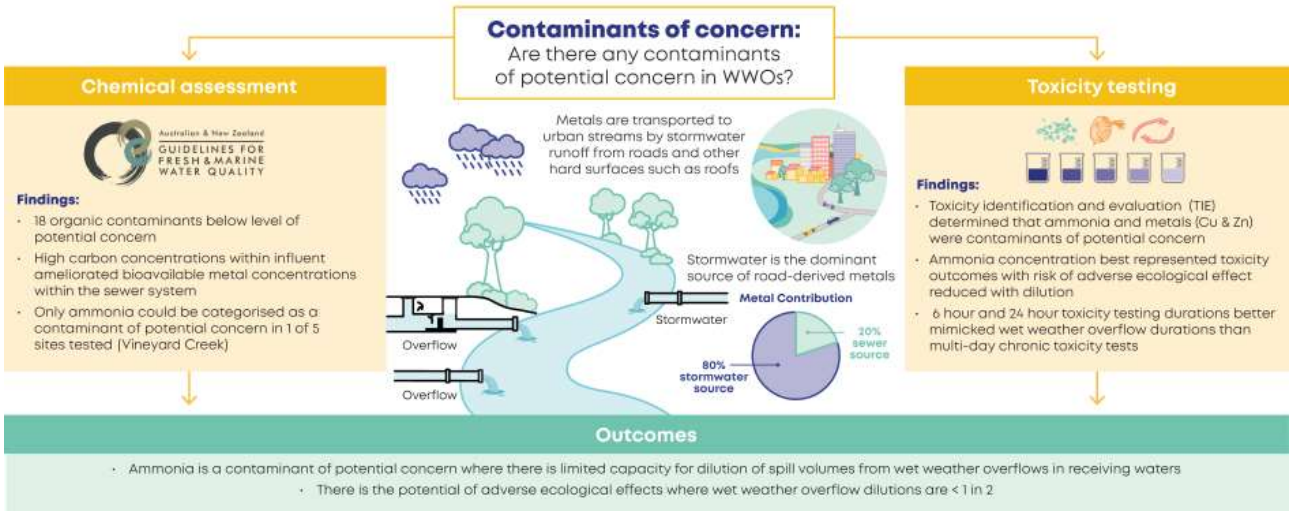


Figure v: Conceptual model illustrating the contaminants of concern

Investigating adverse ecological risk in receiving waters

Outcomes of studies evaluating contaminants of concern (Section 4) and companion toxicity investigations (Section 4.5) have established ammonia as a contaminant of 'potential' concern in receiving waters where insufficient dilution of WWO spills occurred. The insufficient dilution of WWO spills resulted in adverse ecological effects observed under three receiving water situations, from the macroinvertebrate study of sites situated upstream and downstream of an ERS within a stream (Sections 4.7 and 5.3). These three receiving water situations are illustrated in the conceptual model in Figure vi. These low dilution situations were:

1. too many ERSs spilling to same point of an urban stream (for example, Vineyard Creek)
2. oversized ERS spilling to a very small urban stream (for example, Kittys Creek)
3. too many spatially separated ERSs spilling to a small urban stream reach (for example, Girraween Creek)

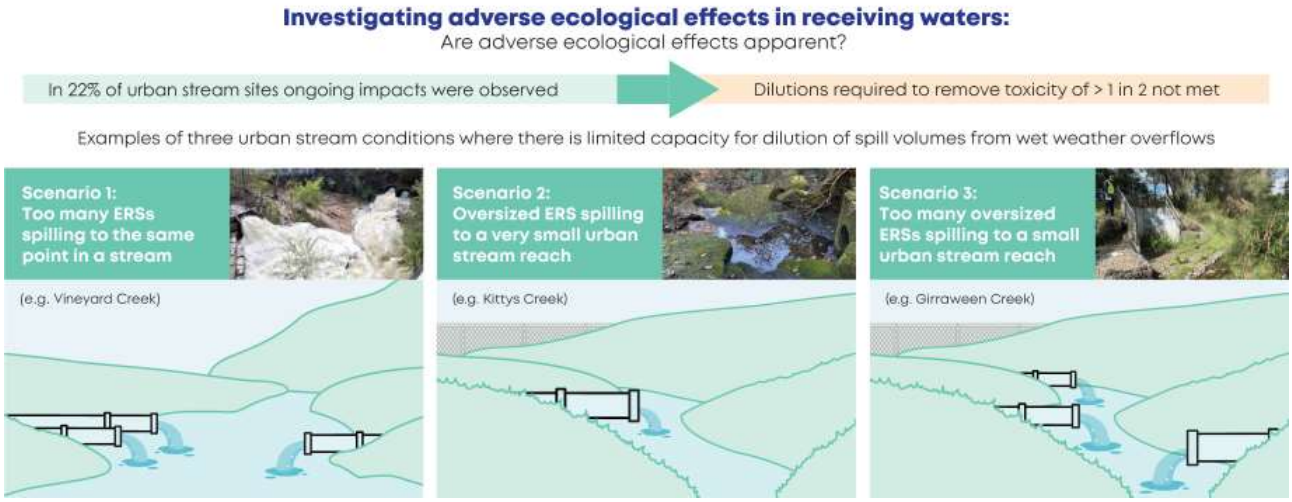


Figure vi: Conceptual model illustrating adverse ecological effects in receiving waters from WWOs

Modelling of traditional microscopic taxonomy (Section 5.4) and DNA-obtained taxonomy (Sections 5.6 and 5.7) against assembled metadata was unable to yield a predictive modelling capability to separate WWO adverse effects. Instead, the natural variation of measured metadata variables (altitude, riparian cover, riparian stream width, and sediment types) across the Sydney region confounded modelling outcomes. In contrast, the close spatial scale paired-site assessments of morphometric macroinvertebrates were successful in identifying adverse ecological effects in low dilution situations. Nevertheless, the morphometric approach would be cost-prohibitive across potentially 660 (22% of) ERSs (Figure vi) in low dilution settings that would need to be assessed. Hence, inclusion of measurements of ecological assemblages would be an unsuitable cost to Sydney Water customers, and as such is not recommended in future assessments. Instead, direct measurement of ammonia using in-situ sensor arrays with subsequent comparison against the corresponding default guideline value of ANZG (2018) is advocated as a cost-effective approach to identify potential 'local' adverse ecological effects from WWO spills that appear to be present in urban streams.

Application of the outcomes to a revised risk prioritisation methodology

The WWOA risk prioritisation methodology is a comparative risk assessment approach for assessing the waterway values of public health and ecosystem health. The 2024 – 2030 comparative risk assessment is based on the "likelihood" and "consequence" of an event occurring, which have been defined as:

- Likelihood: the chance of poor water quality that does not support public health or ecosystem health objectives
- Consequence: the extent or scale of potential water quality impacts on public health or ecosystem health objectives

To maximise the area over which the comparative risk could be assessed, the 2024 – 2030 risk prioritisation methodology relies on data inputs from base datasets that provided the most complete coverage across the study area for the parameter of interest.

For public health, the risk prioritisation applied the 95th percentile enterococci concentration data from Beachwatch. Outcomes from the WWOM proposes that the 95th percentile enterococci concentration be replaced with an assessment of HFMG concentration results of the four HFMGs (HF183, Lachno3, crAssphage, and PMMoV) evaluated against the risk-based thresholds (Section 3.3) for estuarine waters.

Currently, modelled 80th percentile chlorophyll-a (Chla) concentration data from Sydney Water's RMA water quality models raised from data collected at limited estuarine sites as the input used as a proxy to rank the likelihood of poor water quality. WWOM has established that using ammonia as a stressor input (Sections 4.4 and 4.5.1) would be a suitable improvement to the prioritisation methodology and would replace modelled Chla. Ammonia concentration results can be evaluated against the ANZG (2018) guideline value (for protection of 95% of species) to represent risk of adverse ecological effect from WWOs.

A proposed approach to the collection and application of HFMGs and ammonia is outlined in Sections 8.2.2 and Section 8.2.3.

Other program benefits

A more detailed description of benefits, derived from the WWOM, to other Sydney Water programs is outlined in Section 5.9 for the Sydney Water Aquatic Monitoring program (SWAM) and Section 6 dry-weather sewer overflow (DWSO) investigations.

Monitoring using ammonia concentrations that are compared against the ANZG (2018) 95% species protection default guideline value to assess risk of adverse ecological effects is already being employed within DWSO investigations. Findings under the contaminants of concern studies (Sections 4.4 and 4.5.1) support continued use of this chemical marker in DWSO investigations (Section 6.1).

The use of HFMGs in post cleanup assessments where elevated enterococci concentrations are measured have already been implemented in the DWSO investigations. Further development of in-house laboratory capability with animal associated faecal marker genes (Section 3.6) has commenced to support DWSO investigations to assess if animal faecal matter may be responsible for elevated enterococci concentrations, such as in the case of waterfowl observed at cleanup sites. Use of both human- and animal-associated faecal marker genes would add certainty to determine if a human faecal source remains, and further repairs are warranted. If animal faecal contamination was indicated this would then allow human resourcing to be prioritised elsewhere (Section 6.1).

Another potential tool to differentiate recent / active sewage contamination (DWSO) is provided by the suite of eight organic chemical (acetaminophen, ibuprofen, metformin, sucralose, theobromine and three benzotriazoles) markers (tracers) (Section 6.2).

Use of HFMG sediment concentrations for future assessments of DWSO leakage from the Sydney Water sewerage system may be another tool to employ in receiving environment investigations (Section 6.3).

While the above single target HFMG and animal faecal-associated marker gene approach is well advanced, the outcomes of the WWOM DNA pilot studies clearly indicate that further evaluation of Biomonitoring 2.0 ‘rebuild’ and ‘renovation’ approaches with multiple molecular targets are required. A series of recommendations to undertake further evaluation of the application of DNA based taxonomy with ‘paired-site’ assessments under the SWAM program is outlined in Section 5.9 to establish if at least equivalent or better assessment of adverse ecological effects are afforded by measurement of community-DNA or environmental DNA.

Glossary

Term	Abbreviation	Description/explanation
Absolute host sensitivity		Host sensitivity values of human wastewater- and animal scat-associated marker genes were calculated based on three different criteria (i) PCR/RT-PCR detection results (ii) qPCR/RT-qPCR quantifiable data for /mL of human wastewater or /gram of wet weight animal scat, and (iii) qPCR/RT-qPCR data /ng of nucleic acid.
Analysis of variance	ANOVA	Statistical test to evaluate the difference between the means of more than two groups
Aquatic Invertebrates of Australia database	AIA	Database of Australian invertebrates, containing specimen metadata and DNA barcodes. This is housed within the Barcode of Life Data (BOLD) system
Assay		The laboratory workflow from DNA extraction to sequence outputs. Often refers to the target gene and taxonomic group (for example, universal, COI, mitochondrial 16S)
Australian River Assessment System	AUSRIVAS	AURIVAS is a prediction system used to assess the biological health of Australian rivers, developed at a family-level
Barcode of Life Data system	BOLD	International database containing specimen metadata and DNA barcodes
Baseline dry-weather flows		A period where no rain has fallen for at least 72 h prior to sampling
Beachwatch		A water quality monitoring program of NSW ocean beaches, estuaries and inland waterways to assess for swimmability
Benthic		At the bottom of the water column in sediment surface or sub-surface
Best of Both Worlds	BoBW	An article construct of the BOLD and AIA DNA reference libraries
Biomonitoring 2.0		A new paradigm in ecosystem monitoring which employs DNA-based identification of taxa, coupled with high-throughput DNA sequencing on next-generation sequencing platforms
Biostimulation		The modification of the environment to stimulate bacterial growth or nutrient cycling.
Bray-Curtis association measure		The Bray-Curtis association (resemblance) measure has become commonly employed in ecological analyses to raise a similarity matrix to then allow running multivariate statistical analyses upon.

Term	Abbreviation	Description/explanation
Canonical correlation of principle coordinates	CAP	The purpose of CAP is to find axes through the multivariate cloud of points that have the strongest correlation with some other set of variables (canonical correlation).
Combined sewer overflow	CSO	In other parts of the world, the sewerage and stormwater systems can be intentionally integrated into a combined system. The term combined sewer overflows (CSOs) is used to describe discharge points in an integrated combined system where stormwater runoff and wastewater from both domestic and industrial sources are combined and become an influent spilt to downstream receiving waters
Conceptual Process Model	CPM	A simplified representation of a systematic procedure
Deoxyribonucleic acid	DNA	DNA is the molecule that carries genetic information for the development and functioning of an organism
DNA barcodes		Fragments of DNA used to identify taxa, unique to a species
Distanced-based linear models	DISTLM	DISTLM is a routine for analysing and modelling the relationship between a multivariate data cloud, as described by a resemblance matrix, and one or more predictor variables
Dry-weather sewer overflow	DWSO	Overflow during dry weather of undiluted influent (wastewater) from the wastewater system
Dry-weather sampling		Sampling conducted during baseline dry weather flow typically with a minimum of 72 hours with no rainfall
Duplex assay		Combined use of two markers in one assay
Emergency relief structure	ERS	Overflow structure, such as a weir and pipe, into a waterway
Environment Protection Authority	EPA	The EPA is the primary environmental regulator for New South Wales and is also responsible for setting and regulating our EPLs.
Environment Protection Licence	EPL	An EPL are the central means to control the localised, cumulative and acute impacts of pollution in NSW.
Eukarya		These are organisms with cells that contain a nucleus as well as membrane-bound organelles (most commonly associated with plants, animals and fungi)

Term	Abbreviation	Description/explanation
GenBank®	-	GenBank is the National Library of Medicines collection of publicly available DNA genetic sequence database
Human faecal marker genes	HFMGs	In this report, specifically referring to crAssphage CPQ_056, *Bacteroides* HF 183, Pepper Mild Mottle Virus PMMoV and Lachnospiraceae Lachno 3
Human (public) health		The health and wellbeing of the community
Impervious surface		An impervious surface are generally artificial structures like pavements, roads, and buildings. Impervious surfaces impede filtration of water into the soil (NSW Movement and Place)
Infiltration/Inflow	I/I	Avenues where stormwater can enter the wastewater system during wet weather events.
Independent Pricing and Regulatory Tribunal	IPART	IPART is the independent pricing regulator and operating licence administrator for water utilities in NSW
In silico	-	Experimentation performed by computer
Macroinvertebrate		Animals without a backbone that can be seen without a microscope
Microbial Assessment Category	MAC	Beachwatch defines Microbial Assessment Category or MAC as "There are 4 Microbial Assessment Categories (A to D) and these are determined from the 95th percentile of an enterococci dataset of at least 100 data points. Each MAC is associated with a risk of illness determined from epidemiological studies. The risks of illness shown below are not those associated with a single data point but are the overall risk of illness associated with an enterococci dataset with that 95th percentile (Wyer et al. 1999).
Mitochondrial cytochrome c oxidase	COI	COI is a protein complex found in bacteria, archaea, and the mitochondria of eukaryotes.
Microbial source tracking	MST	A technique for evaluating water quality for faecal pollution.
Model for Urban Sewers	MOUSE	Sydney Water's hydraulic wastewater system model
	NSW	New South Wales
National Health and Medical Research Council	NHMRC	The NHMRC is the leading expert body in health and medical research, who provide clinical, public health and environmental health guidelines to

Term	Abbreviation	Description/explanation
		support the translation of research into health practice and policy.
Pearson correlation coefficient	PCC	PCC is a measure of strength of a linear association between two variables
Permutational multivariate ANOVA	PERMANOVA	PERMANOVA is a routine for testing the simultaneous response of one or more variables to one or more factors in an analysis of variance (ANOVA) experimental design on the basis of any resemblance measure (such as Bray-Curtis), using permutation methods
Polymerase Chain Reaction	PCR	Method widely used in molecular biology to make many copies of a specific DNA segment.
PCR bias	PCR bias	Amplification of a sequence may vary from one sequence to the next. This can cause incorrect information on the abundance and diversity of genes, either over or underrepresenting sequences after amplification.
Polar Organic Chemical Integrative Samplers	POCIS	An absorbent membrane for water soluble compounds
Pollution reduction program	PRP	A PRP is a requirement in EPLs to carry out works or to install plant for the purpose of preventing, controlling, abating or mitigating pollution
Pollution study	PS	PS are set by the EPA to require a licensee to undertake investigations into any aspect of the environmental impact (air, water, land, sensitive receivers in the case of noise and the appropriate management of waste) of the activity or work authorised or controlled by an EPL. Pollution studies can be initiated by the EPA or a licensee.
Quantitative Microbial Risk Assessment	QMRA	A QMRA is a mathematical modelling approach used to estimate the risk of infection and illness when a population is exposed to microorganisms in the environment.
Quantitative Polymerase Chain Reaction	qPCR	qPCR is a PCR-based technique that couples amplification of a target DNA sequence with quantification of the concentration of that DNA species in the reaction
Operational taxonomic unit	OTU	An OTU is the basic unit used in numerical taxonomy, and can refer to individual, species, genus or class.
One standard deviation	1SD	1SD defines a region that includes 68% of all the data points.

Term	Abbreviation	Description/explanation
Risk assessment		A process to identify hazards and evaluate the risks associated with those hazards.
Risk-based threshold	RBT	The HFMG marker concentration that corresponds to the benchmark gastrointestinal (GI) illness rate is referred to here after as the risk-based threshold.
Ribonucleic acid	RNA	RNA is a long, single-stranded chain of cells that processes protein and carries genetic information. 16S is used to sequence bacteria and 18S is commonly used to sequence animals.
Road density		Road density is calculated by road length (km) divided by the catchment area (km^2)
Sewer Catchment Asset Management Plan	SCAMP	A SCAMP is a sub-section of the wastewater hydraulic system
Sewer Overflow Licensing Project	SOLP	Environmental impact statements for each of the 27 sewerage systems as part of the submission by Sydney Water to the NSW EPA for EPLs (in 2000).
Stream Invertebrate Grade Number Average Level	SIGNAL	SIGNAL is a biotic index. A biotic index has a measure of pollution sensitivity from grades assigned to taxa at either family or genus taxonomic level, and when based on count data a measure of abundance
Stream Invertebrate Grade Number Average Level – Sydney Genus	SIGNAL-SG	As per SIGNAL, where 'S' indicates Sydney region version and 'G' indicates taxonomy is at the genus taxonomic level
Stormwater urban runoff marker		A chemical found only in stormwater, in sufficient concentrations to be used to identify stormwater from sewer overflows
Student Newman Keuls	SNK	A post hoc test for differences in means
Sydney Water Aquatic Monitoring Program	SWAM	Previously the STSIMP. The SWAM is an annual regulatory EPA license requirement to conduct waterway health monitoring to assess the performance of the wastewater network and treatment facilities.
Sewage		See 'Wastewater'
Sewage Treatment System	STS	See Wastewater system
Sewage Treatment System Impact Monitoring Program	STSIMP	See SWAM

Term	Abbreviation	Description/explanation
Uncontrolled/undirected overflows		Discharges from the sewerage system at points not designed to spill during wet-weather events
Wastewater		The contents of a wastewater system pipe, including discharges from homes and businesses. The term 'wastewater' has generally replaced the term 'sewage'
Wastewater markers		Chemicals specifically found in wastewater. Able to be used to identify if wastewater has entered the stormwater.
Wastewater system		The system of pipes and other infrastructure that transports wastewater. The term 'wastewater system' has generally replaced the term 'sewerage system'
Water Services Association of Australia	WSAA	WSAA is the peak industry body representing the urban water industry.
Weight of evidence assessment	WEA	Uses a combination of adequate and reliable documentation from different sources to add weight to a conclusion
Wet weather overflow	WWO	The release of excess, rainwater ingress diluted influent (wastewater) from the wastewater system (sewerage system)
Wet Weather Overflow Monitoring Program	WWOM	WWOM is a scientific research program investigating the influence of WWOs on the receiving environment
Wet Weather Overflow Abatement	WWOA	WWOA delivers the regulatory obligations as detailed in the relevant clauses of Environment Protection Licences for our sewage systems

Table of contents

Figures

Tables

1 Background

Over the last 20 years, Sydney Water has been working with the NSW Environment Protection Authority (EPA) to meet the overflow frequency targets/limits from the landmark Sewer Overflow Licensing Project (SOLP). Our long-term goal is to minimise the impact of our activities and contribute to a more liveable Sydney, helping to deliver healthier waterways, with cleaner beaches and estuaries. We plan to do this by effectively managing our wastewater system, thereby meeting government and community expectations to have access to fresh and marine water for recreation and ensuring sustainable aquatic ecosystems.

Frequency targets, set in 2000 by the NSW Environment Protection Authority (EPA), aimed to reduce the number of WWOs over a 10-year period. The intention of this target was to recover swimming days and/or protect ecosystem health. Since 1998, we have invested approximately AUD $1.5 billion (2012) to reduce the frequency of WWOs contributing to the cleaner beaches and waterways currently enjoyed by the community. Until 2020, capital investment mainly involved building large storage tanks, tunnels and bigger pipes and pumps. It was estimated in 2015, that to meet the frequency targets across the four major coastal catchments (of North Head (Environment Protection Licence (EPL) number: 378), Bondi (EPL 1688), Malabar (EPL 372), and Cronulla (EPL 1728)), we would need to build around 48 storage tanks at 48 ML each, amplify 2,600 km of pipes and upgrade 31 pumping stations. Under the frequency targets, this abatement solution was calculated to cost more than AUD $8 billion (2018), which would significantly increase customer's water bills. Recent estimates place the cost storage solutions at more than AUD $18 billion (2022) and take over 350 years.

In 2015, Sydney Water identified WWO regulation as its number one risk and this was included as a performance metric in the Managing Director's Contribution Development Plan. As a more cost-effective solution, Sydney Water proposed to the EPA that frequency targets be replaced with a risk-based approach for abating waterway ecosystem health and human health impacts. Sydney Water and the EPA have worked for several years to develop a suitable risk-based framework to inform the new regulatory measure for the management of WWOs in the four major coastal wastewater systems. In July 2019, the EPA formally revised the EPLs for these four sewage treatment systems by including a delivery obligation (Pollution Reduction Program 307) to reduce volume and/or frequency of overflows at sites identified as highest relative risk using a risk prioritisation methodology during the 2020 – 2024 IPART period. More recent refinements of the regulatory measure obligate Sydney Water to reduce volumes at the highest priority sites for the 2024 – 2030 Independent Pricing and Regulatory Tribunal (IPART) period. In addition, the EPA has applied a continuous improvement requirement to the prioritisation methodology (detailed within Pollution Study 307). The EPA considers the Wet Weather Overflow Abatement program (WWOA) as a rolling commitment to abate WWOs and has identified that these licence conditions will be an enduring requirement.

Abatement of WWOs was initially aimed at managing human health risks in recreational waters. Investment was generally applied based on outputs from the wastewater hydraulic Model for Urban Sewers (MOUSE) model. Consequently (with the exception of the Northside Storage Tunnel), it has been difficult to characterise adequately the environmental and human health benefits from

abatement. Depending on the volume and duration of a rain event and the local wastewater system performance, waterways can receive both WWO spills and urban stormwater runoff (containing pollutants). The ability to distinguish between the influence of stormwater and wastewater has historically been difficult. Merely confirming the presence of wastewater in the environment is not a direct measure of the risk or impact. Suitable waterway monitoring could confirm the significance of WWOs to human health and waterway ecosystem health. Targeted and robust monitoring would allow Sydney Water to invest in areas where there is greater certainty around the impacts of its wastewater system and provide clarity for the potential to derive a benefit from abatement works. During the process of transitioning from frequency targets to volume reduction, Sydney Water has committed to continuously improve our process of defining the highest risk WWOs, which includes undertaking scientific investigations to better understand the influence of WWOs on the receiving environment.

1.1 What are wet-weather overflows?

Sydney Water's wastewater system consists of a network of pipes and pumping stations which transport wastewater from households, businesses and industries to wastewater treatment plants and water recycling plants. Sydney Water provides wastewater services to about 5.2 million people.

These services cover over 12,700 square kilometres (km) across 24 separate wastewater systems. Wastewater is transported to about 30 treatment plants through 27,000 km of wastewater pipes and 695 wastewater pumping stations. Customers own about another 22,000 km of wastewater pipes on private properties, which connect into the Sydney Water wastewater system.

The wastewater system and stormwater system are two independent systems. During light rainfall, the stormwater volumes in channels and the depth of the receiving water environment may rise, but there is normally no change in the performance of the wastewater system. Wastewater system pipes are typically designed to carry three times the average expected dry weather flow. Under moderate to heavy rainfall, stormwater can enter (via inflow/infiltration) into the wastewater system increasing flows within the pipes by 5 to 10 times more than in dry weather. This change in dilution of sewer influent under differing magnitude rainfall events was documented for the Sydney wastewater system by Besley et al, (2023).

Inflow and infiltration can be caused by faults in both private sewers and Sydney Water's wastewater system. Inflow can occur when assets are in poor condition or properties are incorrectly connected into the wastewater system. The Water Services Association of Australia (WSAA, 2023) notes common faults causing inflow as:

- downpipes from roof or pool drains incorrectly connected into the wastewater system
- cross connections between stormwater and wastewater pipes
- faulty or damaged maintenance holes
- faulty valve or backflow prevention devices in emergency relief structures (ERS – a designed overflow point)
- faulty or low gully traps
- wastewater pump station wet wells inundated during localised flooding

Infiltration is a gradual increase of stormwater or groundwater flow into the wastewater system during and after wet-weather events. Most of the wastewater systems in the Sydney region are laid in impermeable rock or clay trenches that often serve as channels where stormwater can follow sewer pipes and enter through faults (Bickford et al., 1999). Examples of faults include cracks in pipes; displaced pipe joints and damaged maintenance holes and ERSs (Figure 1-1).

Sydney's wastewater system is designed with ERSs. ERSs allow excess wastewater to overflow into stormwater drains or waterways during wet-weather, protecting homes and businesses from flooding (see Figure 1-2). Without these designed overflow points, wastewater could backup through toilets and floor waste drains and flood homes and businesses or gully traps in backyards (see Figure 1-3). Hence, ERSs are a necessary component of the Sydney wastewater system and are designed to protect public health by stopping sewage backup into homes and business (Bickford et al., 1999).

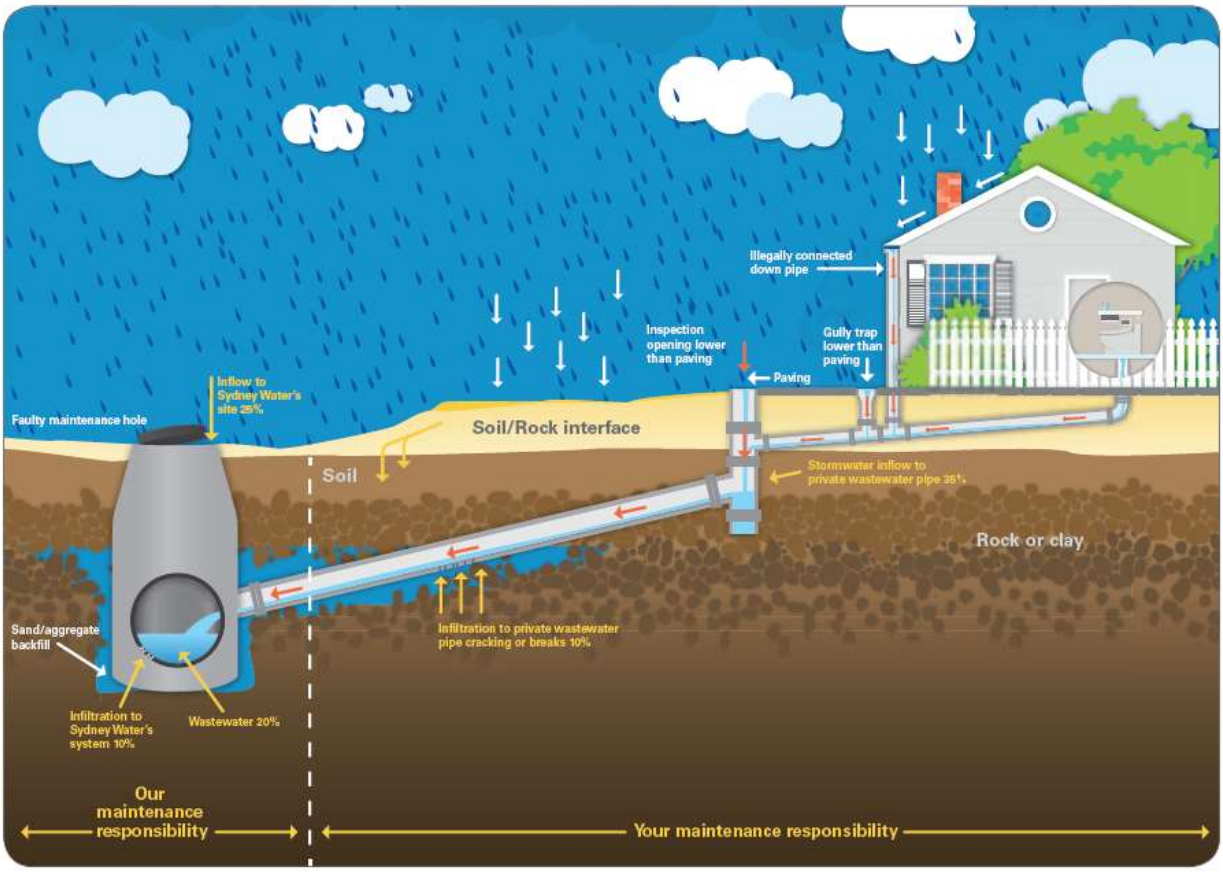


Figure 1-1: How water enters the wastewater system during wet weather

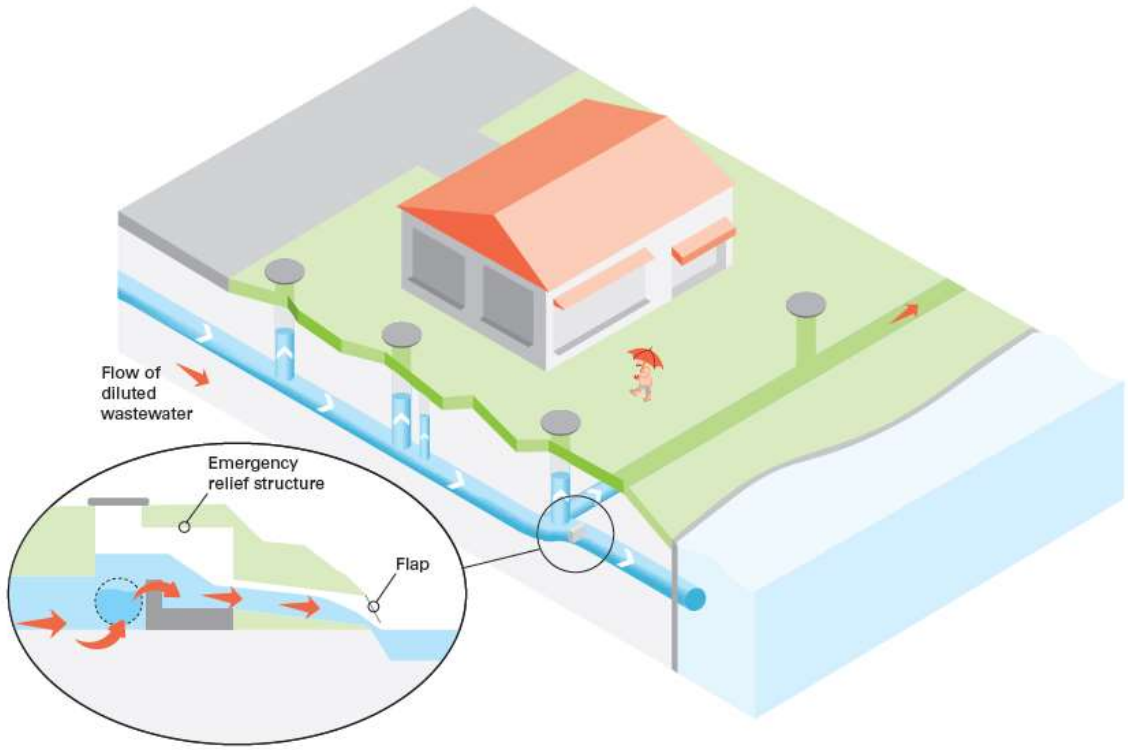


Figure 1-2: Mechanics of a WWO

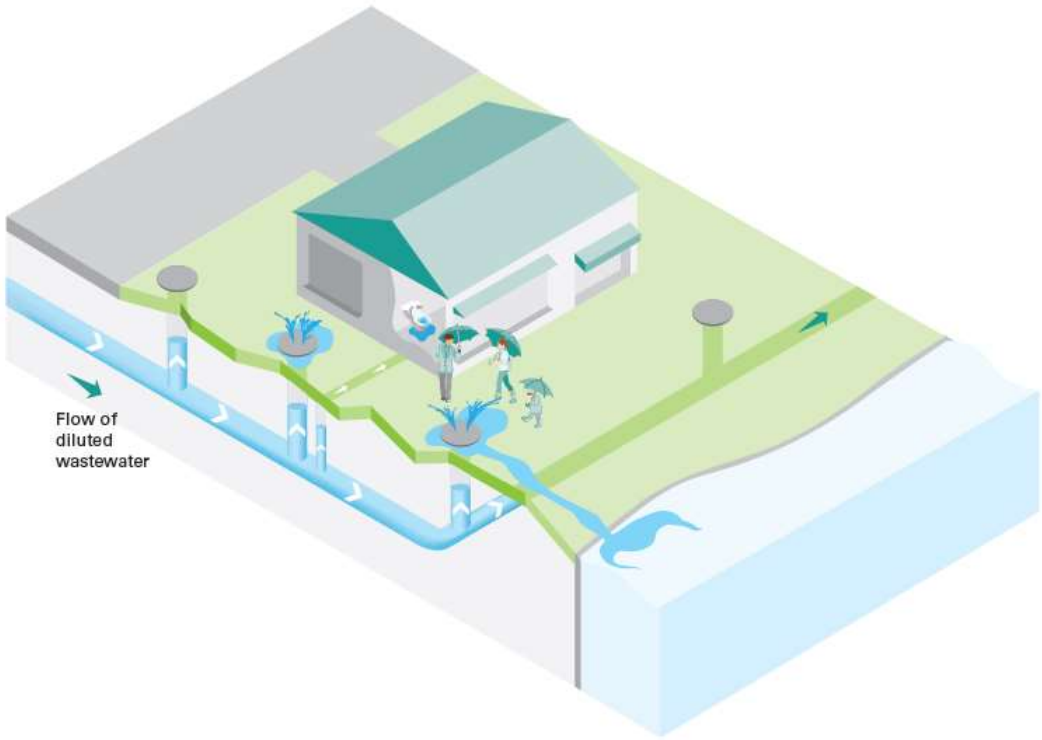


Figure 1-3: Potential human health impact to homes when emergency relief structures are not installed to redirect excess diluted wastewater within the system

1.2 Developing the Wet Weather Overflow Monitoring Program

1.2.1 Historic and current knowledge and capabilities

Sydney Water has over thirty years' worth of existing knowledge and capabilities in studying the influence of our operations on receiving waters. We have:

- Collected data for standard water quality parameters (temperature, dissolved oxygen, conductivity) and nutrient (total nitrogen (TN), total phosphorus (TP), chlorophyll-a (Chla)) parameters, the latter focusing on biostimulation across many of our waterways. Collect every three weeks at 18 estuarine water quality sites
- over 20 years' data from freshwater macroinvertebrate sampling, with a widely recognised reference specimen voucher collection. Collected six monthly at 48 freshwater river and stream sites
- intertidal data for inputs into the Environmental Indicators Monitoring Program reports and the subsequent report series (from 2008 to 2023) of the Sewage Treatment System Impact Monitoring Program. Measured six monthly at 26 intertidal rockplatform sites
- data collected while assisting Beachwatch by collecting water samples for the Illawarra area to analyse for the faecal indicator bacteria (FIB) enterococci. Colected every 6 days between October to April, then geenrally monthly at key beaches (see Beachwatch for further details).

Benthic estuarine data have not been historically collect by Sydney Water. That decision was informed by a water column study documenting the predominant disturbance to be in the surface layer of the estuarine water column (Besley et al., 1993). This preliminary study informed decisions to construct the Northside Storage Tunnel (this storage tunnel is described further in Section 3.1) to store excess WWO volumes. This early study also informed studies of subsequent intertidal biota (epibenthos on rock platforms) conducted under the above-mentioned two report series.

Standard measurements of water quality parameters remain an adequate method of quantifying stressors on the receiving waters. The response from organisms was not well understood. We recognised that whilst there were abundant data from traditional methods of impacts assessment such as freshwater macroinvertebrates (identified and enumerated), there are limitations to this traditional data collection. Whilst macoinvertebrates can provide good knowledge of impacts from ongoing (press) disturbances from a single point source such as a wastewater treatment plant with a paired upstream and downstream site design, this indicator may not be able to tease out the potential adverse ecological effects of WWOs from numerous disparate point sources interspersed among multiple stormwater inflows within a catchment. Enterococci measurements are currently used to assess potential impacts on human health and have been used since 1993 to detect sewage contamination. However, these cannot dissect a human source from animal sources (that is, birds, dogs, cattle, etc.).

A literature (Google Scholar) search illustrates that there is a lot of published research on combined sewer overflows (CSOs), but such information is limited on separate sanitary sewer/stormwater systems, such as exists in Sydney, hence the need to invest in applied research. In Sydney, the most recent assessment was conducted by Bickford et al. (1999). That

study focussed on discharges from treatment plants and spills from the largest eleven overflows, which represented the majority of WWOs in terms of frequency and volume at the time.

1.2.2 The Wet Weather Overflow Monitoring Program study intent

The focus of the WWOM program was to increase the scope of the WWO points investigated beyond the subset documented by Bickford et al. (1999). The high to low volume/frequency overflows were obtained from the 1000 licenced modelled WWO points (L7.1 of Sydney Water's EPLs). This was expected to enable a finer understanding of the adverse effects of WWOs (human pathogens in sewage) from those of urban stormwater runoff (and the contaminants transported from the urban landscape in stormwater, for example, road-derived metals) in receiving waterways. Sydney Water aimed to improve its past approach by investigating more recent technologies that have emerged in the last decade and to test the application of these methods; and in turn understand the cost differences between methods. The knowledge gaps in understanding the influence of WWOs involved:

- Establishing the method(s) and technologies to assess if WWOs adversely effect ecosystem health and if it can be separated from contaminants transported in urban stormwater, for example, road-derived metals
- Assessing if microbial source tracking of human-faecal associated marker genes can potentially replace (subject to cost) the enterococci indicator in the existing risk prioritisation assessment for capital solution planning
- Investigating contaminants that may pose an adverse ecological risk
- Developing a greater understanding of the operation of WWOs and how the wastewater system functions at a local level
- Gaining insight into the types and amount of sewage-derived gross pollutants spilled from Sydney's wastewater network

The outcomes of the Wet Weather Overflow Monitoring (WWOM) Program will allow us to target capital expenditure in the right areas with confidence that benefits can be achieved and utilise the learnings as an input into a refined risk prioritisation methodology, thereby delivering the ongoing continuous improvement commitment.

A summary of the four study components (themes) of the WWOM program is outlined in Figure 1-4 below.

References

Bickford, G., Toll, J., Hansen, J., Baker, E., Keessen, R. 1999. Aquatic ecological and human health risk assessment of chemicals in wet weather discharges in the Sydney region, New South Wales, Australia. Mar. Pollut. Bull. 39, 1-12.

Besley, C., Collie, T. and Scammell, M. 1993. Recovery after wet weather water quality sampling Middle Harbour. AWT Science & Environment, Water Board. 93/153.

Besley, C.H., Batley, G.E., Cassidy, M. 2023. Tracking contaminants of concern in wet-weather sanitary sewer overflows. Environ. Sci. Pollut. Res. 30, 96763-96781. https://doi.org/10.1007/s11356-023-29152-x

WSAA 2023. Managing Wet Weather Flows. Guidelines for the Australian & New Zealand Water Industry. Water Services Association of Australia. ISBN 978-1-922808-09-7

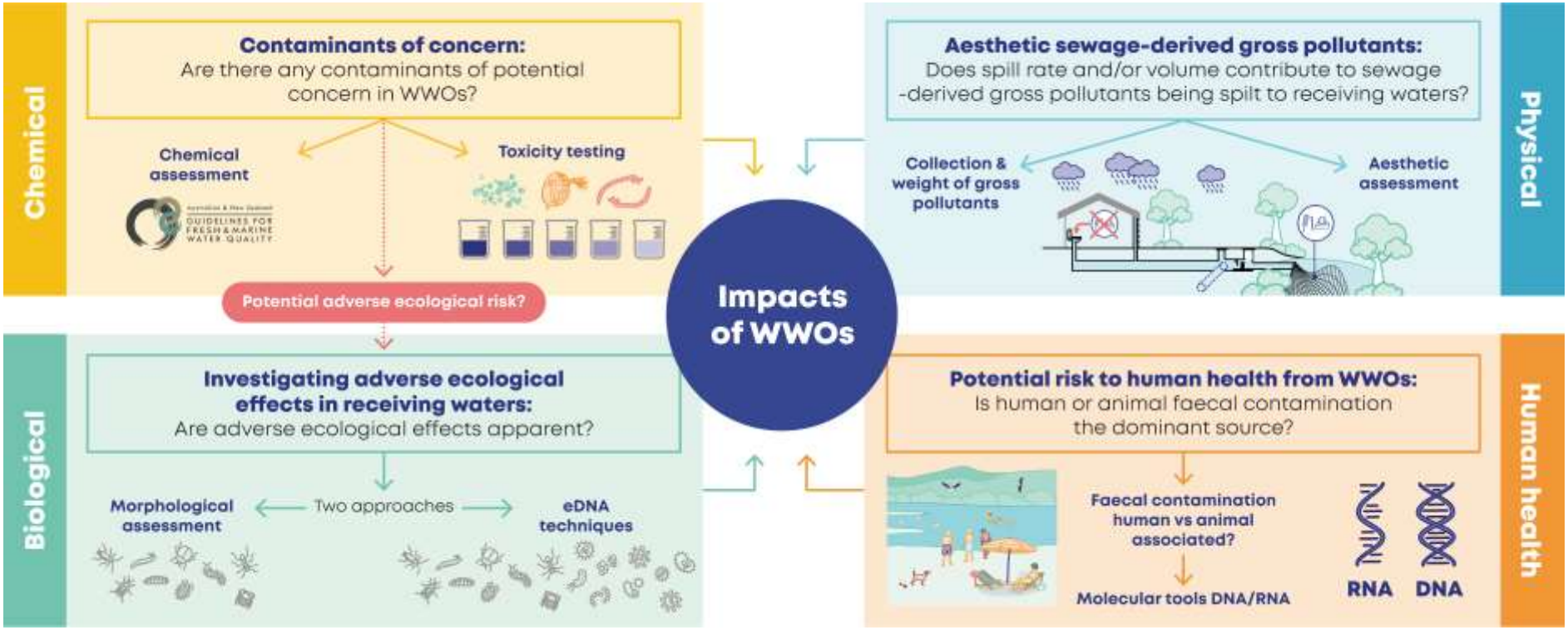


Figure 1-4: Summary of WWOM program showing components investigated and overarching research questions

1.3 Expert peer panel

The WWOM program applied techniques and methodologies new to Sydney Water. Use of these new techniques in assessing WWOs from a sanitary sewerage system has not been reported in international scientific literature. As such, an independent expert peer review panel of five members was appointed to provide assurance that the outcomes of the program would be robust, withstand scrutiny and adequately inform our executive, Board, the EPA and stakeholders. Each subject matter expert has a specific environmental or human health specialty. The role of the expert peer panel was to review our implementation and application of monitoring tools, analysis of data and approach for meeting the proposed improvement levels.

The expert peer panel consists of:

Doctor Graeme Batley AM (2016 - 2024), Commonwealth Scientific and Industrial Research Organisation (honorary fellow), has extensive expertise in the assessment of risks associated with contaminants in aquatic ecosystems. Dr Batley has been working in his field for over 50 years. His major research interest is in analytical and environmental chemistry, and he was involved in the development of the water and sediment quality guidelines and the authorship and revision of revision of the Australian and New Zealand water quality guidelines. Dr Batley has worked with various government agencies (including OEH, NSW Maritime, and the Commonwealth Department of Climate Change, Energy, the Environment and Water), the mining industry and other utilities such as Melbourne Water Corporation. In addition, Graeme is involved with many government and private industry committees and panels. Graeme has 471 research publications. Dr Batley also presented to the Sydney Water Board in 2015 on the proposed risk-based method and supported further scientific studies.

Associate Professor Gavin Birch (2016 - 2024), Sydney University (retired), has extensive experience in estuarine and marine contaminants, with specific interest in the effects of contaminant constituents in stormwater on receiving waterways and sediment toxicity. A/Prof Birch collaborates extensively with universities and government institutions overseas, including the United States National Oceanographic and Space Administration (NOAA), the United States Geological Survey (USGS), Simon Fraser, Southern Cross and Hong Kong Universities. Gavin has 259 publications, as well as 115 conference abstracts. Gavin is a reviewer for several international journals and granting bodies; and contributes to many advisory agencies, such as the NSW EPA, Sydney Metropolitan Catchment Authority and Sydney Harbour Board Strategic Plan.

Professor Melanie Bishop (2016 - 2024), Macquarie University, is an expert in coastal ecosystems. She has over 15 years' experience in researching coastal and estuarine ecosystems, with an interest in the impacts of nutrient contaminants on estuarine ecosystems. Prof Bishop co-leads the green engineering working group of the World Harbour Project and the Living Seawalls Program. Melanie has over 160 peer-reviewed publications on estuarine and coastal ecosystems. Melanie has also provided advice on the Threats and Risks Framework for the NSW Marine Estate.

Professor Angus Webb (2016 - 2024), University of Melbourne, is a specialist in hydrology and water resources. Prof Webb is a quantitative ecologist, his research focuses on the study of

landscape-scale impacts of human-induced disturbances on freshwater systems, and what can be done to restore these systems. Originally training in marine ecology, Prof Webb has been focusing on freshwater systems since 2001 and has led monitoring, evaluation and research efforts for the outcomes of environmental flows delivered under the Murray-Darling Basin Plan since 2014. Angus has published 213 peer-reviewed research publications.

Associate Professor Susan Petterson (2016 - 2022), Water & Health/Griffith University, is a senior risk assessment expert with over 20 years in water-related microbial risk both nationally and internationally. A/Prof Petterson serves as an editor for the IWA Journal of Water and Health and was a lead author of the WHO guidance document for undertaking quantitative microbial risk assessment (QMRA) (2016). Susan is a member of the National Health and Medical Research Council Water Quality Advisory Committee, the Independent Metropolitan Water Advisory Panel for the Sydney and Hunter and the Technical Assistance Committee for the environmental surveillance of SARS-CoV-2 for PATH. Susan is currently a non-executive Director of the Sydney Water Board (2022 - 2025).

Professor Anne Roiko (2022 - 2024), Griffith University, is a specialist in human health risk assessment, risk communication and applying research to policy. She has over 30 years' experience as an environmental health scientist, predominantly in the academic field. Professor Roiko is a member of several advisory boards including the International Advisory Board of the SARA Project, an international research project about microbial water quality in aquatic ecosystems, the Recreational Water Quality Advisory Committee for the National Health and Medical Research Council and the Public Health Scientific Expert Panel for Healthy Land & Water (as chair). Anne has published 116 peer-reviewed research publications.

1.4 Collaborations and publication

Distinguishing the adverse effects of WWOs (human pathogens and contaminants in sewage) from urban stormwater runoff (and the animal faecal matter and contaminants they contain) is a complex and difficult exercise. A review (Section 1.2.1) of the monitoring tools currently employed by Sydney Water for existing regulatory programs revealed that we would not be able to achieve our desired outcomes using these tools. Sydney Water recognised the need to improve capability with emerging methods and technologies; and to also understand the cost differences between methods to allow for future adoption of cost effective lines of evidence. We harnessed leading scientific knowledge and expertise by collaborating with universities and the Commonwealth Scientific and Industrial Research Organisation (CSIRO) to better understand how WWOs influence the potential for adverse ecological effects and human health risk (Figure 1-5).

In this project, Sydney Water has collaboratively undertaken novel research. Rather than have our collaborators generate grey reports, the project has pursued the publication of the research studies in international scientific journals, which is a standard academic tenet. Journal publication involves a blind peer review of our collaborative work and ensures robust outcomes of the WWOM to assure our regulator and Sydney Water customers that this investment has been cost-effective. This work will be highly valued by other water utilities.

Figure 1-5: Schematic of collaborations with academic institutions and Sydney Water service providers

The following section provides a background on our collaborators and their inputs:

Doctor Warish Ahmed (2017 - 2024), Principal Research Scientist, Commonwealth Scientific and Industrial Research Organisation (CSIRO), is a human health industry expert with extensive expertise in microbial source tracking methods to identify point and non-point sources of microbial contaminants in receiving waters. Dr Ahmed's main research areas include next generation sequencing to identify microbial diversity of aquatic ecosystems; using microbial source tracking (MST) to track the source of microbial contaminants in receiving waters; quantifying pathogens in the receiving environment; determining the decay of contaminants; and quantitative microbial risk assessment. During the recent Covid (SARS-CoV-2) pandemic, Dr Ahmed led the Queensland monitoring of and research into tracking COVID-19 in wastewater. Dr Ahmed has published more than 226 peer reviewed journal articles. Dr Ahmed guided Sydney Water in how to apply new DNA methods in MST to analyse our WWOs and receiving waters. Warish also trained our laboratory staff in the qPCR analysis protocols.

Doctor David Roser (2016 - 2017), University of New South Wales, has over 40 years' research/consultancy/work/post graduate experience in microbial ecology and environmental microbiology, especially in respect to waterborne pathogen microbiology, and environmental/risk management, with 69 publications in his area of expertise. Dr Roser is qualified in both microbiology and environmental planning and management. He has extensive experience in the collection use of water microbiology sample data and the use of this data in Quantitative Microbial Risk Assessment especially in respect to sewage impacts. To determine the risk of WWO on the community use of receiving environments, we needed to verify the factors that are causing a risk to human health in the area and determine if a solution has reduced risk to human health and improved community usage. Dr Roser was engaged to review the early concepts for sampling for human health.

Associate Professor Anthony Chariton (2018 - 2024), Macquarie University, pioneered environmental DNA (eDNA) metabarcoding in Australia. His research focusses on the development, application and integration of 'omic' technologies and traditional monitoring and assessment of aquatic ecosystems. His key interest is in how communities respond to natural and anthropogenic stressors. A/Prof Chariton has 120 publications in this field and is a co-author of the revised Australian and New Zealand Sediment Quality Guidelines. A/Prof Chariton was initially engaged on the WWOM program in an advisory role to develop the eDNA pilot studies learnings into a broader monitoring campaign. Later, A/Prof Chariton led a team of national and international experts, to undertake bioinformatic analysis of eDNA sequenced data, using this and a range of other relevant data collected during the study, to develop the Predictive Ecological Response Model. Anthony has also provided training to Laboratory Services staff in the collection of eDNA using specialised sediment cores.

Doctor Anu Kumar (2020 - 2024), Principal Research Scientist, Commonwealth Scientific and Industrial Research Organisation (CSIRO), is an ecotoxicologist with more than 25 years' experience in tracking and identifying emerging contaminants of concern, isolating pollution sources, assessing impacts of aquatic and terrestrial pollution on ecosystems; developing cost-effective monitoring systems; and developing water management guidelines and practices. Her work spans the agriculture, aquaculture, industrial and urban sectors, both nationally and

internationally and she is a leading expert in ecotoxicological assessment in Australia. Dr Kumar has 95 publications. Dr Kumar led the ecotoxicological assessment pilot study, which is the final component to assess the contaminants that potentially pose an adverse biological risk.

Associate Professor Katherine Dafforn (2016 - 2018), previously of the University of New South Wales (UNSW) now Macquarie University, was engaged to conduct a pilot study investigating the utility of using genomics or molecular analysis to document macroinvertebrate community composition (structure) in fresh and estuarine waters. A/Prof Dafforn's expertise is in sensitive monitoring tools for large spatial areas and using a range of methods to measure aquatic health. Katherine has 127 publications. Our collaboration with A/Prof Dafforn involved a pilot study to assess how DNA analysis (genomics) of sediments compares against our current macroinvertebrate methodologies (macrofauna sampling methods and morphological taxonomy); and determining if community function data collected using genomic assessment provided input consistent with the needs of the predictive ecological response model. A/Prof Dafforn also trained our Laboratory Services staff in the collection of eDNA samples. Section 7.3 outlines the findings of this work.

Doctor Michael Shackleton (2017 - 2022), La Trobe University, the Murray-Darling Freshwater Research Centre (MDFRC), specialises in macroinvertebrate systematics and molecular ecology for freshwater environments. Michael has published 37 journal papers. Sydney Water engaged Dr Shackleton to conduct CO1 genetic analysis of freshwater macroinvertebrate samples collected during the pilot study stage (Section 5.5.1). We used the outputs of this assessment to compare with the UNSW molecular analysis of community composition using 16S/18S and traditional analysis (Section 7.3).

The Centre Aquatic Pollution Impact and Management (CAPIM) (2016 - 2018), University of Melbourne. CAPIM was a scientific research organisation established in 2010 to identify and address the impact of pollution in water environments, with a multidisciplinary team with expertise in chemistry, ecotoxicology, biomarker research, fish histology, metabolomics, animal morphology and animal behaviour. CAPIM led development in pollution detection technologies to detect acute pollution events, especially in stormwater drains; detect pesticides and endocrine disrupting chemicals, and monitor water sediment quality with cost-effective, integrated water monitoring tools. CAPIM has worked with Melbourne Water, Yarra Valley Water and WaterNSW using passive sampling techniques. Sydney Water engaged CAPIM to conduct a passive sampler trial to compare results to traditional autosampler techniques. Our Laboratory Services staff were trained in the techniques and methods used by CAPIM in the passive sampler trial.

Doctor Mayumi Allinson (2016 - 2022), University of Melbourne, specialises in investigating organic contaminants in water, sediment and biological samples. Dr Allinson was initially engaged by CAPIM to analyse the chemicals collected from the passive sampler pilot study. Dr Allinson was key to the success of the method development and analysis of passive sampler results. Mayumi also trained our laboratory staff in method development. Refer to Section 4.1 for outcomes of the passive sampler pilot study.

Glenn McDermott (2016 - 2019), Enviromon, has over 40 years' in the field of flow measurement design, analysis and characterisation services. Glenn has been working with Sydney Water for many years on various sewer gauging projects and has unparalleled knowledge and experience

working on Sydney Waters sewer flow monitoring programs. Enviromon developed theoretical rating tables for the overflow volumes from each of the complex measurement structures.

Sydney Water Laboratory Services (2016 – 2024) undertook equipment deployment, sampling and analysis for the WWOM program:

- Field Sampling and Testing – deployment and retrieval of passive samplers in estuarine waterways, collection of waters for toxicological analysis and sediment samples from estuarine waters for eDNA analysis, collection from autosamplers of water column samples to chemcial metadata analyses to support toxicological interpretations
- Aquatic Ecology – collection and identification of macroinvertebrates; deployment and collection of passive samplers in freshwater streams, collection of sediment samples for eDNA analysis
- Hydrometric Services - deployment, calibration, reporting of level and pulsar gauges; autosamplers; and rain gauges
- Trace Analysis – analysis of passive samplers and chemical metadata laboratory analyses to support toxicological (metals, pesticides, PFAS) interpretations and preparation of sample containers from collection of toxicological water samples
- Water chemistry – analysis of ammonia, CBOD, and BOD to support toxicological interpretations
- Technical Services/Microbiology – eDNA extracion for subsequent qPCR labotatory analysis of MST from water column and sediment samples, additional extraction of eDNA from sediment samples for subsequent metabarcoding studies

2 Gross pollutant pilot study

2.1 Proactive management of a growing wet wipe problem

Customer complaints arise from visible gross pollutants deposited on receiving waterway margins and in associated riparian vegetation after WWOs. This is an increasing problem. The need for Sydney Water to continue proactive messaging as a management action is apparent. For example, Ó Briain et al. (2020) advocated the need to increase public awareness of microplastic fibre pollution in the marine environment arising from inapt disposal of sanitary products down the toilet and they suggested a message to promote disposal via land-based waste management was required. This seems justified, as Ashley et al. (2002) commented that over the last 20 years there has been a substantial increase in the amount and types of disposable items flushed, and the volume of products containing plastics are predicted to continue to increase. This statement is supported by the Grand View Research (2019) estimate of the global personal-care wipes market value at AUD $15.8 billion in 2018. They suggested that this is expected to grow at a compound annual growth rate of 5.6% from 2019 to 2025 to reach USD $23.1 billion by 2025. Grand View Research (2019) identified some of the major characteristics of this growth are cost, convenience, hygiene, performance, ease of use, time saving, disposability, and consumer-centric aesthetics. Furthermore, there has been the introduction over the past decade of wipes with diverse applications, such as intimate, wet, feminine, and scented.

In the Sydney region, 500 t of flushed wet wipes are inappropriately disposed into the sewerage system with removal costs upwards of AUD $8 million every year and this figure is increasing (Sydney Water, 2020). The messaging that Ó Briain et al. (2020) suggested is encompassed in Sydney Water's 'Clean up not down' campaign that asks customers to use a rubbish bin to dispose of wipes, or any other bathroom products, rather than flushing them down the toilet (Sydney Water, 2020). Messaging campaigns such as this are clearly needed by water utilities in the face of an upward trend in usage of wipe products. The current proactive messaging campaigns of Sydney Water are Toilet Blockers Anonymous (Only flush the 3Ps: Pee, Poo and Toilet Paper! initiative) and the It's Best To Bin It Factsheet (sydneywater.com.au).

2.2 Reactionary management presents another approach

A potential reactionary management action is to collect inappropriately disposed gross pollutant items at the designed emergency relief structures (ERS - designed overflow points of the sewerage network). A gross pollutant pilot study was conducted to assess the capability, logistics and cost-effectiveness of a commercial trash net to capture gross pollutants that spill from overflow points (Figure 2-1).

This study had four objectives:

- to examine the types of gross pollutants that spill during an overflow event
- to assess the amount of gross pollutants captured
- to investigate if the captured gross pollutant composition is similar across the four ERSs that had different modelled frequencies of spilling
- to examine for the presence of visible gross pollutants on adjacent land and/or caught in fringing bankside vegetation of another 95 categorised as high risk overflow points.

Besley and Cassidy (2022) documented this pilot study. Text, graphic and citations in this section is drawn from the peer-reviewed journal publication: Besley, C.H. and Cassidy, M., 2022. The composition of gross pollutants contained in wet weather overflows for different locations, spill frequencies and discharge volumes. J. Environ. Manage. 303, 114256. https://doi.org/10.1016/j.jenvman.2021.114256.

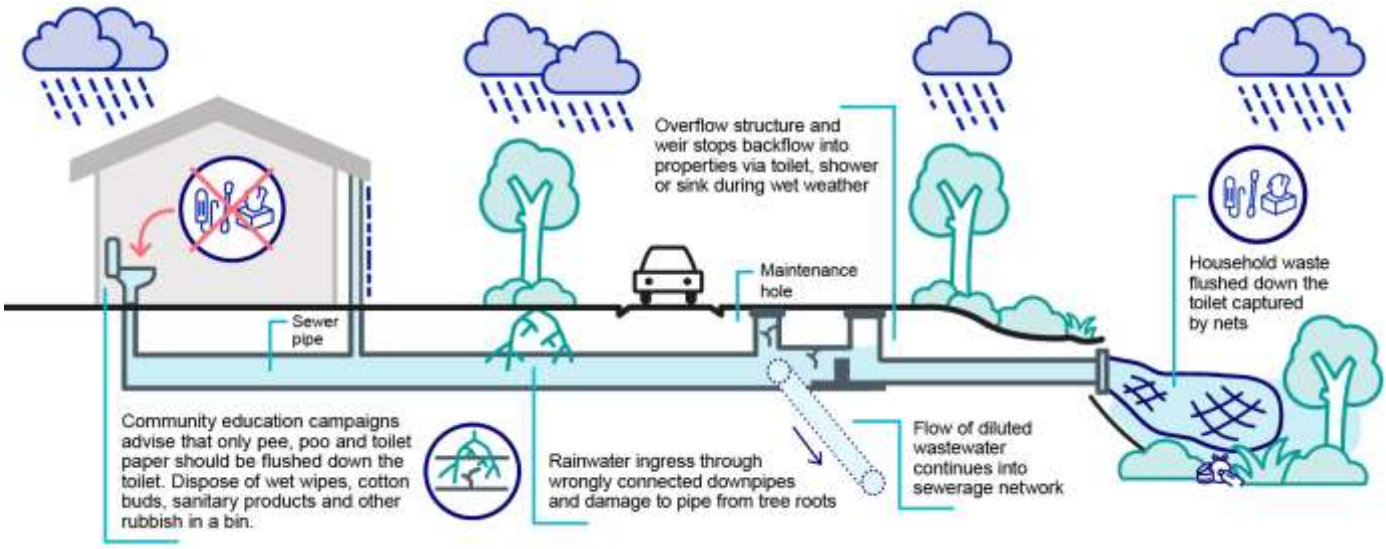


Figure 2-1: Graphical abstract illustrating the transport of sewage-derived gross pollutants via the wastewater system

2.2.1 Site selection

Nets were installed on the end of overflow outlet structures at four locations in the Sydney metropolitan area, at: Mill Creek, Alfords Point; Hunts Creek, North Rocks; The Ponds Creek, Rydalmere; and Salt Pan Creek, Padstow (Figure 2-2). These overflow structures were chosen to represent two low, one medium and one high volume and frequency WWOs (based on hydraulic modelling), respectively. An additional 95 overflow points were selected for study based upon modelled sewer overflow volume.

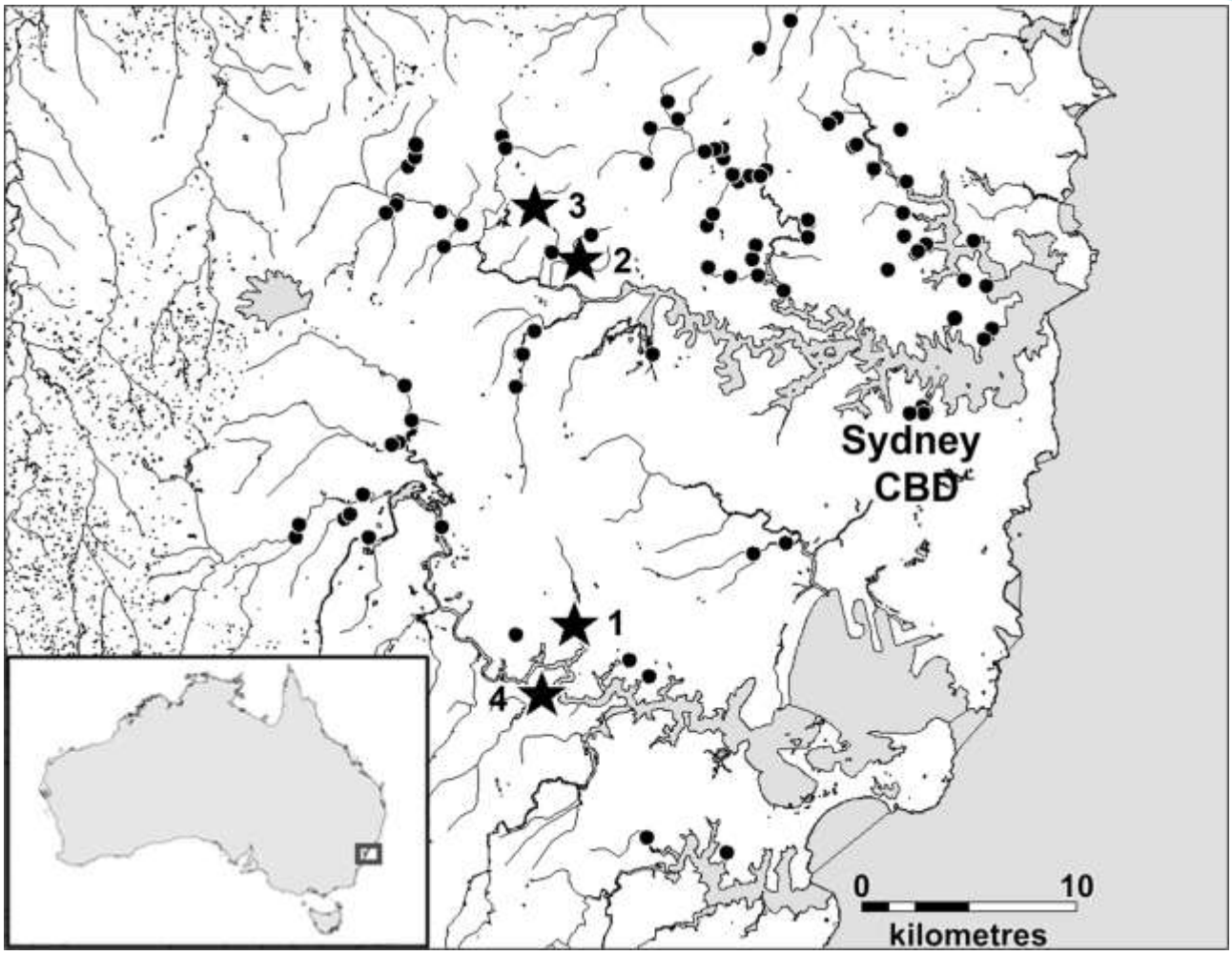


Figure 2-2: Gross polluants study locations

Four overflow locations with nets attached to collect sewage-derived gross pollutants (black stars) and another 95 overflow locations where aesthetic visual inspections were untaken for both sewage-derived and stormwater-style gross pollutants (black dots).
Netted overflow locations: 1 = Salt Pan Creek; 2 = The Ponds Creek; 3 = Hunts Creek; 4 = Mill Creek

2.2.2 Main findings of pilot study

The gross pollutant pilot study was conducted to assess the capability of a modified commercial stormwater trash-net to capture gross pollutants that spill from ERSs of the sewerage network. Trash nets were installed on four ERSs with differing gauged spill frequencies and overflow volumes (Figure 2-3). Five categories of gross pollutants (wet wipes and paper products, sanitary items, condoms, miscellaneous identifiable items, and matted immeasurable items) were collected in similar proportions across the two highest volume overflows, although the weight of material was highest from the largest volume overflow. Only wet wipes and paper products were collected from the lowest volume overflow. Wet wipes and paper products were the main gross pollutant collected

comprising 54% (159 kg) of the total weight over the study period. This weight was potentially an under-representation as wet wipes were observed matted together and intertwined in the net that inhibited separation, with this material allocated into the 'immeasurable' category comprising 30% (89 kg) of the collected material. Regression modelling indicated that the weight of sewage-derived gross pollutants captured was most influenced by overflow spill volume. The five categories of gross pollutants were collected in similar proportions across the two highest volume overflows of Salt Pan Creek and The Ponds Creek, although the weight of material was highest from the largest volume overflow of Salt Pan Creek. In contrast, only wet wipes and paper products were collected from Hunts Creek which spilt the lowest volume of wastewater during the study period. The gross pollutant pilot study demonstrated a modified commercial stormwater trash net was able to capture sewage-derived gross pollutants spilt from the sewer network, and this capture also prevented those pollutants from entering receiving waters.

The companion aesthetic survey observed gross pollutants originating from the sewer at 36% of the 95 aesthetic study locations while gross pollutants originating from the stormwater system were found at 82% of the aesthetic study locations. Sewage-derived gross pollutants were observed to extend further downstream at ERSs with increased modelled WWO spill volumes. The observed build-up of wet wipes and paper products at ERS outlet on coarse screen bars was the dominant sewage- derived gross pollutant, which was consistent with the main sewage- derived gross pollutant collected in the deployed nets.

2.2.3 Proposed approach to assist implementation of reactionary management

In response to customer complaints of visible aesthetic sewage-derived gross pollutants, maintenance crews are dispatched to clean-up gross pollutants. This current method is reactive and labor-intensive displacing human resources from other maintenance activities. Deployment of nets at strategic ERSs would allow collection of sewage-derived gross pollutants once spills commence and ameliorate the spread of sewage-derived gross pollutants through the receiving environment. By implementing a planned maintenance screening program to gauge, net and dispose of laden nets, water utilities could mitigate visible aesthetic sewage-derived gross pollutants and in-turn reduce customer complaints. The balance between the cost of sending maintenance crews in a reactive manner against using this screening tool on strategic ERSs of concern would need to be evaluated to balance operational costs.

Records of customer complaints may form a primary source of evidence to define strategic ERSs for a screening maintenance program. Inclusion of ERSs that release high-volume siphonic sewer overflows would seem appropriate from the pilot study results. A cost-effective selection of gravity-fed ERSs for future screening may include aesthetic inspections for presence of sewage-derived gross pollutants combined with assessment of potential overflow spill volume from hydraulic modelling of the sewerage system. Adopting the screening approach as documented in the pilot study with the installation of both nets and gauging equipment would incur the cost of purchase, installation, and maintenance for the gauging component. If overflow locations were found to spill infrequently and at low volume, the equipment could be moved to other overflow points on the sewage network. If gauging was not installed, a more frequent visitation schedule would be prudent in warmer seasons to minimise odour complaints.

Figure 2-3: Modified Ecosol ™ Net Tech trash net attached to sewer overflow structures

(A) Hunts Creek, North Rocks; (B) gross sewage- derived pollutants collected in net; and (C) trash net attached to overflow at The Ponds Creek, Rydalmere, closeup of net showing captured gross sewage-derived pollutants

2.2.4 Application of reactionary management and continued periodic community education campaigns for proactive management

Results from Besley and Cassidy (2022) suggest that if future screening with these modified trash nets was implemented as a reactionary management tool, it would reduce the presence of sewage- derived gross pollutants and would minimise the risk of public contact with gross pollutants. Selection of ERSs with relatively higher overflow volumes, such as from siphonic overflows would yield the highest capture of sewage- derived gross pollutants. The approach outlined in Section 2.2.3 is advocated in selection of gravity-fed ERSs that have relatively higher spill volumes.

Besley and Cassidy (2022) cited the more permanent in-pipe screening used by the USEPA (1999). This type of structure may be an option for the highest volume siphonic overflows, such as those located within Sydney Airport lands discharging to the Mill Stream (a fringing waterway), to ameliorate the spread of sewage-derived gross pollutants through the receiving environment. Although the potential for backup into homes and businesses would need to be considered with more permanent screen structures compared with detachable nets used in the pilot study.

Periodic community education campaigns for the proactive management of gross pollutants remain necessary as cited literature predicts an upward trend in usage of wet-wipe products.

Recommendation

Screening with modified trash nets at siphonic and at larger volume and frequency spilling gravity-fed ERSs would afford the most cost-effective approach to use this potential management tool and to reduce physical effort by field teams in site clean ups post spills. At the highest volume siphonic overflows such as Mill Stream, a more permanent in-pipe screening may be a more suitable option for the potentially greater weights of spilled gross pollutants.

3 Human health risk from wet-weather overflows

Microbial source tracking (MST) has developed over the past two decades as a technique for evaluating water quality, specifically for the purpose of determining the sources of faecal pollution in environmental waters (Paruch and Paruch, 2022). This approach is founded on the premise of leveraging host-associated molecular marker genes found in bacteria, protozoa, and viruses in the faeces of a diverse range of animal species, including humans (Harwood et al., 2014). These marker genes are identified through the implementation of sequence analysis or other molecular methods such as subtractive hybridisation (Dick et al., 2005; Gomi et al., 2023). Many quantitative PCR-based (qPCR) assays have been developed to detect and quantify faecal pollution source-specific marker genes in environmental waters. MST stands as a powerful tool in the ongoing effort to monitor, with the potential to better manage environmental (receiving) water quality (Harwood et al., 2014).

The utility of human faecal marker genes (HFMGs) arises from their concentrations being 3–5 orders of magnitude greater than infectious enteric viruses in untreated sewage influent (as documented in Section 3.2.7 Ahmed et al., 2022, sub-study 7). Ahmed et al. (2022) found that initial high concentration within influent assists with more frequent detection of HFMGs in receiving waters where dilution of sewage influent occurs.

The human health pilot study of the WWOM project sample collection was conducted between 2016 and 2022 with the overall aim to apply emerging microbial source tracking monitoring tools to understand the human health microbial contamination risk posed by WWOs from ERSs in receiving waters.

In parallel to the WWOM program, a risk prioritisation assessment tool has been developed to address the NSW EPA licence pollution reduction program (PRP 307) and pollution study (PS 307), requiring to Sydney Water to undertake abatement works at prioritised sites. The current version of the prioritisation methodology is based upon enterococci enumerated using traditional culture-based laboratory methods. Outcomes of the human health pilot study are intended to inform a revised risk prioritisation methodology based on HFMGs that positively affirm human sewage as the source in receiving waters and meets the pollution study (PS307) objective of continuous improvement to the prioritisation methodology. Information was obtained under the human health sub-studies outlined in Sections 3.1 and 3.2 below that were collaboratively conducted with Dr Warish Ahmed of CSIRO. This preceding work allowed the development of a Quantitative Microbial Risk Assessment (QMRA) that defined Sydney-specific risk-based thresholds (RBTs) for each of the four studied HFMGs. These RBTs correspond to the USEPA benchmark of 32 illnesses per 1000 swimmers (USEPA, 2012). The outcomes of the QMRA and Sydney specific risk-based thresholds are documented in Section 3.3. Discussed in Section 3.4 is the application of risk-based thresholds (RBTs) to inform future revision of Sydney Water's risk prioritisation methodology in management of WWO abatement. It should be noted, this work is not being used to inform the public about swimmability, that work is undertaken by Beachwatch.

Explanation of abbreviations in journal publications and this section

Within this report, HFMGs refer to the MST gene assays that are human specific. The published papers also refer to human wastewater-associated marker genes for human faecal pollution tracking.

Within the published papers the terms enteric viruses or pathogenic viruses have been used to describe infectious enteric viruses. In this section, these have been referred to as enteric viruses.

In addition, the phrase wet-weather sanitary sewer overflows (WWO) will be observed or text describing WWOs from a sanitary sewerage system that is separate to the stormwater system. This wording has been adopted to distinguish the Sydney work from other areas of the world where a combined stormwater and sewer systems exist, and those overflows are described in scientific literature as 'combined sewer overflows' (CSO). Studies of combined sewer and stormwater systems do not afford direct comparison of results in the Sydney region with separate systems.

In published papers, we refer to 'fecal' to align with the American English language standards used in the international journal publications. In this report, we use the Australian English "faecal".

References

DPIE, 2020. Protocol for Assessment and Management of Microbial Risks in Recreational Waters. Department of Planning, Industry and Environment, NSW. Accessed 7 June 2024

NHMRC, 2008. National Health and Medical Research Council. Guidelines for Managing Risks in Recreational Water. Australian Government Publication Services ISBN 1864962666.

USEPA 2012. Recreational water quality criteria. United States Environmental Protection Agency: Washington, DC.

3.1 Human health pilot sub-studies informing the quantitative microbial risk assessment

In 2017, in collaboration with Dr Ahmed of CSIRO, a series of human health sub-studies commenced. Sample collection was completed in 2022, while formal analysis and documentation continued in early 2024, culminating in a Sydney-specific QMRA that developed RBTs for each of the four HFMGs studied. The 11 sub-studies explored eight research aspects to inform the QMRA framework (Figure 3-1). This stepped approach in applied research has built confidence in the applicability of MST for determining the sources of faecal pollution (human or animal) in receiving waters. Each of the sub-studies illustrated in Figure 3-1 are documented in Section 3.2.

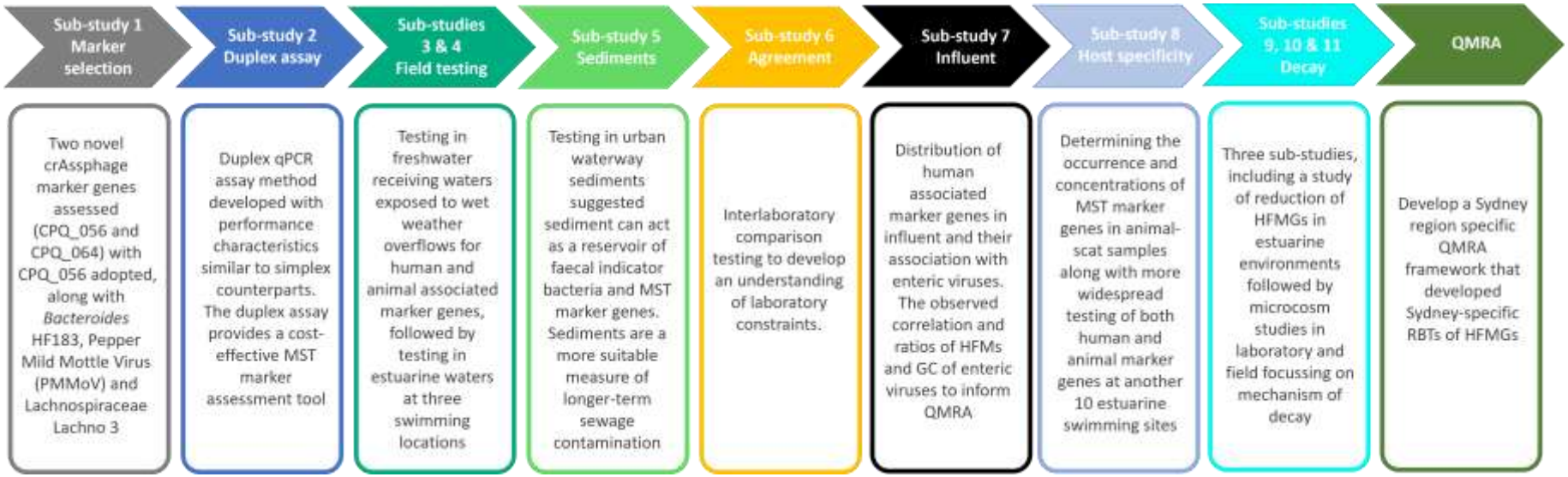


Figure 3-1: Sub-studies conducted to develop a wet weather overflow QMRA framework for the Sydney region

Ahmed et al. (2019a) advocated that, to build confidence in MST approaches, accurate results from field studies are crucial as large financial investments are required to remediate faecal contamination sources in waterways. This opinion was considered prudent given Sydney Water's experience in capital solution investments, for example, the large containment works (Figure 3-2) to date have cost:

- Northern Beach Storage Tank (18 ML) – $80 million
- Northside Storage Tunnel (480 ML / 18 = 27 storage tanks) – $460 million

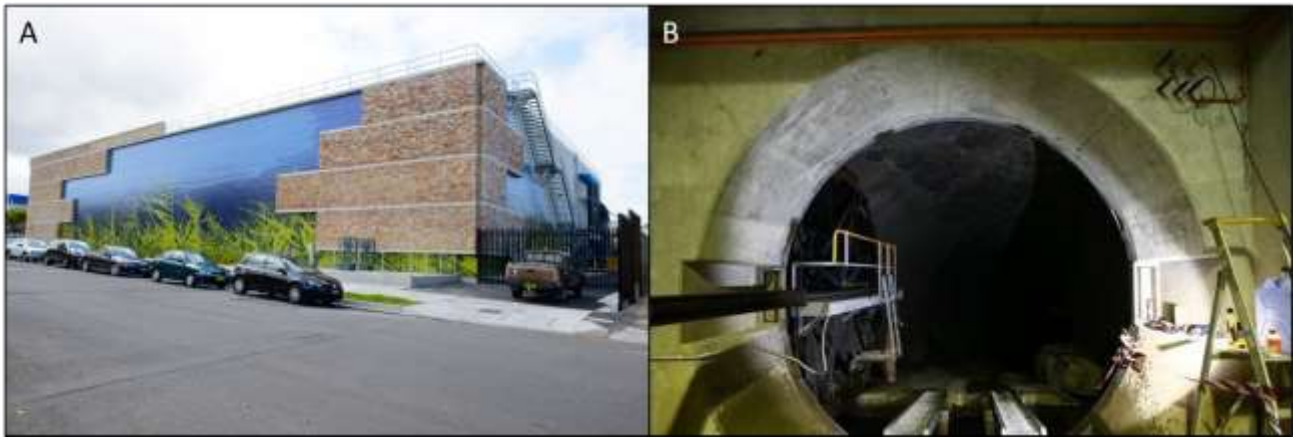


Figure 3-2: Northern Beaches Storage Tank (A) and the Northside Storage Tunnel during construction (B)

Ahmed et al. (2019a) suggested that host specificity and sensitivity are two main performance characteristics for consideration when choosing MST assays. They went on to state, that for field applications, care should be taken to choose appropriate MST marker genes and assays based on available host specificity and sensitivity data and background knowledge of the contaminating sources in the study area. This approach helped identify a cost-effective combination of HFMGs that will be least influenced by false-positive (cross reactivity from animals) and false-negative results (where a human sewage source is not correctly identified) to give confidence in test outcomes in determining if a human sewage source is present or not at a study site. Hence, the series of sub-studies (Figure 3-1) have been conducted within the Sydney region to incorporate the Ahmed et al. (2019a) recommendations for the combined usage of simultaneous marker assays for more accurate and informative sewage contamination detection. Four HFMGs of *Bacteroides* HF183, pepper mild mottle virus (PMMoV), crAssphage CPQ_056 assay, and Lachnospiraceae Lachno3 were incorporated into the sub-studies and then evaluated within the QMRA framework developed. To complement this evaluation of MST marker genes, animal marker genes (of typically encountered urban [dogs, cats, aquatic birds, chickens] and peri-urban animals [cattle, horses]) were also assessed along with several enteric viruses.

References

Ahmed, W., Gyawalib, P., Fengc, S., McLellan, S.L., 2019a. host specificity and sensitivity of established and novel sewage-associated marker genes in human and nonhuman fecal samples. Appl. Environ. Microbiol, 84, https://doi.org/10.1128/AEM.00641-19

Dick, L.K., Simonich, M.T., Field, K.G., 2005. Microplate subtractive hybridization to enrich for Bacteroidales genetic markers for fecal source identification. Appl. Environ. Microbiol. 71, 3179-3183. https://doi.org/10.1128/AEM.71.6.3179-3183.2005

Gomi, R., Haramoto, E., Wada, H., Sugie, Y., Ma, C.-Y., Raya, S., Malla, B., Nishimura, F., Tanaka, H., Ihara, M., 2023. Development of two microbial source tracking markers for detection of wastewater-associated *Escherichia coli* isolates. Sci. Total Environ. 864., 160952. https://doi.org/10.1016/j.scitotenv.2022.160952

Harwood, V.J., Staley, C., Badgely, B.D., Borges, K., Korajkic, A., 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), 1-40. https://doi.org/10.1111/1574-6976.12031

Paruch, L., Paruch, A.M., 2022. An overview of microbial source tracking using host-specific genetic markers to identify origins of fecal contamination in different water environments. Water 14, 1809. https://doi.org/10.3390/w14111809

3.2 Eleven human health pilot sub-studies

As illustrated in Figure 1-4 above, there are 11 published sub-studies that inform the QMRA framework, outlined in Section 3.3.

3.2.1 Pilot sub-study 1: Testing novel assays

At the start of the WWOM project, crAssphage was a novel MST HFMG gene with its status as a surrogate for enteric viruses yet to be established. This initial sub-study assessed the then available crAssphage assays of CPQ_056 and CPQ_064.

Objectives of pilot sub-study 1

The primary objective of this initial sub-study (Ahmed et al., 2018) was to determine the host-sensitivity and -specificity of the then newly designed (novel at that time) crAssphage qPCR assays CPQ_056 and CPQ_064 HFMGs in faecal samples collected from various human and animal host groups in Australia. A secondary objective was to assess the utility of these HFMGs to detect sewage pollution within the freshwater designated-swimming location of Lake Parramatta which receives urban stormwater runoff that can transport WWO spills to the swimming location. A graphical outline of this sub-study is provided in Figure 3-3.

Text, graphic and citations in this section were drawn from the peer-reviewed journal publication: Ahmed, W., Payyappat, S, Cassidy, M., Besley, C., and Power, K., 2018. Novel crAssphage marker genes ascertain sewage pollution in a recreational lake receiving urban stormwater runoff. Water Res. 145, 769-778, https://doi.org/10.1016/j.watres.2018.08.049

Freshwater pilot study location

This initial, second and the third pilot sub-studies were conducted in the urbanised freshwater body of Lake Parramatta (Figure 3-4), which was the only freshwater designated-swimming location within the Sydney region at the time. Companion ERS gauging confirmed that WWOs were active within the catchment. This enabled assessment under these three studies of wet-weather driven spills transported by stormwater to the swimming location within the receiving water environment of this recreational lake.

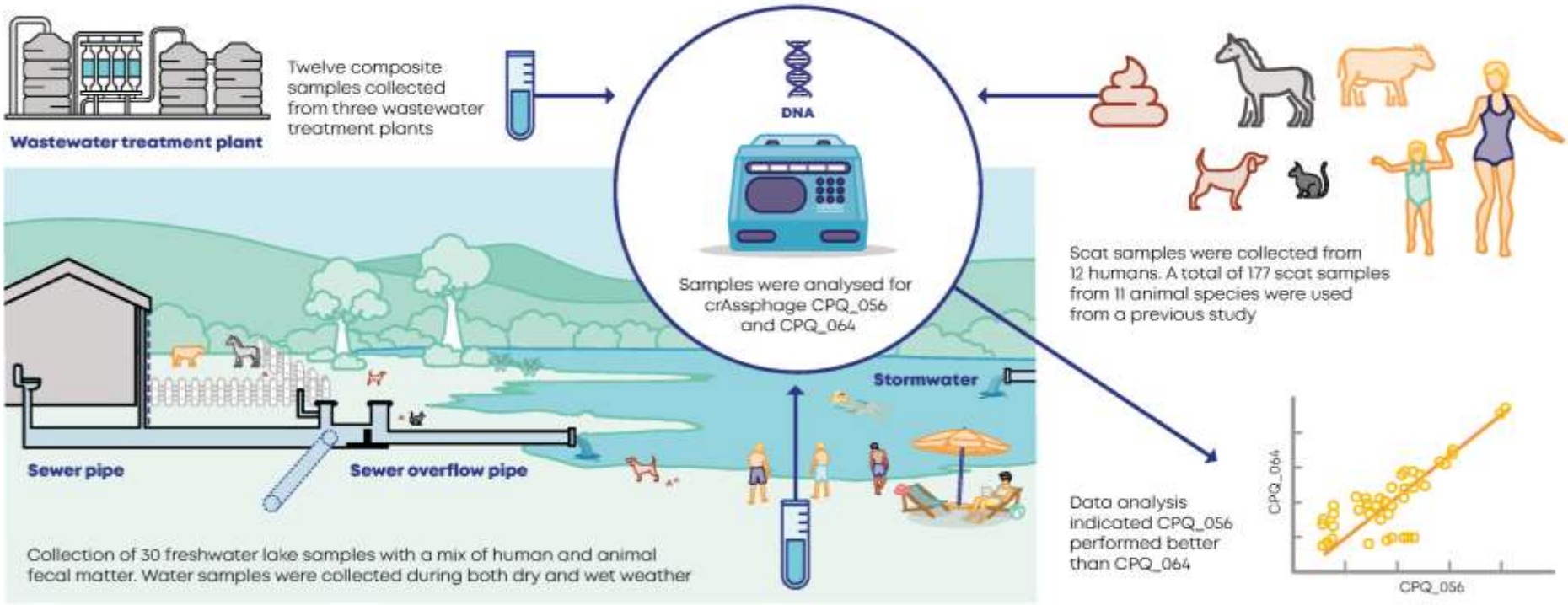


Figure 3-3: Graphical abstract outlining assessment of host sensitivity and specificity of CPQ_056 and CPQ_064 marker gene assays

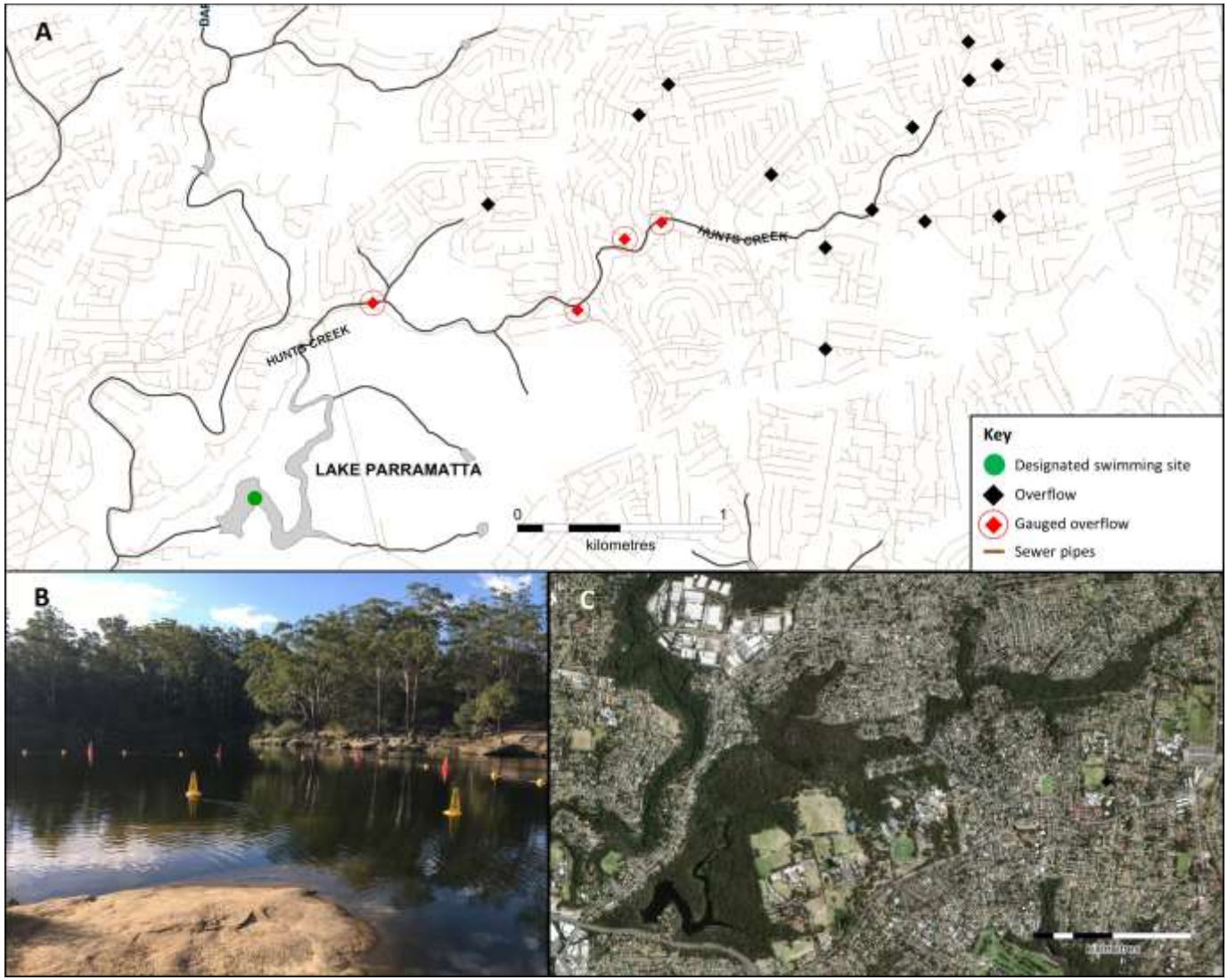


Figure 3-4: The urbanised freshwater swimming location at Lake Parramatta

(A) location of overflows in relation to the designated swimming site; (B) designated swimming area; (C) aerial view of the Lake Parramatta catchment

Main findings of pilot sub-study 1

Outcomes of this study documented mean concentrations of untreated sewage to be marginally higher for CPQ_056 than CPQ_064, although concentrations for both were 2 to 3 orders of magnitude higher than other sewage-associated viruses. Among 177 animal faecal samples tested from 11 species, the host-specificity values for CPQ_056 and CPQ_064 marker genes were 0.95 and 0.93 (1 is the maximal value, zero is lowest), respectively. Limited cross-reactivity was observed from testing of cat and cattle faecal samples. This outcome indicated that a crAssphage HFMG should be simultaneously assessed with another HFMG assay that is not based upon crAssphage.

The concentrations of both markers in the urban lake, receiving stormwater runoff, were quantified in all 20 storm samples for CPQ_056 and in 18 of 20 storm samples for CPQ_064 when WWOs had occurred (as indicated by gauging). Conversely, samples collected in dry-weather returned negative qPCR results for all 10 samples for CPQ_056 and for 8 of 10 samples for CPQ_064. These documented outcomes of Ahmed et al. (2018) shaped the decision to take the better performing crAssphage CPQ_056 HFMG forward into further studies.

These outcomes also provided initial confidence in the MST approach with HFMGs, as it was able to detect human faecal contamination from WWOs in receiving lake water body. This detection was despite the sampling sites being somewhat distant (1 to 3 km upstream) from the ERSs of the sewerage system situated higher up in the urbanised catchment of this lake (Figure 3-4A).

Further research was advocated to understand the decay kinetics of crAssphage markers in environmental waters in relation to FIB and pathogens.

Pilot sub-study 1 key outcome and enacted recommendation

The crAssphage CPQ_056 HFMG assay successfully detected sewage pollution when active WWOs had occurred from distant ERSs situated higher up in the catchment. Importantly, CPQ_056 did not detect sewage under dry-weather conditions when spills from ERSs of the sewerage system were not active.

CrAssphage CPQ_056 HFMG was included as an assessment tool in subsequent sub-studies based upon better performance characteristics to those documented for CPQ_064 assay.

Decay studies were incorporated into the human health WWOM pilot study program based on the recommendation documented in the published paper on this initial pilot sub-study (Ahmed et al., 2018).

3.2.2 Pilot sub-study 2: Developing a duplex assay

Pilot study 2 was undertaken to investigate the potential of combined usage of simultaneous marker assays into a multiplex assay for more accurate and informative sewage contamination detection. A successful outcome would also afford a cost saving over running simplex markers in parallel. This latter aspect was important from a production laboratory point of view with high sample numbers employed in assessing wastewater system performance.

Objectives of pilot sub-study 2

The aim of pilot sub-study 2 (Ahmed et al., 2019a) was to develop a duplex qPCR assay that allows simultaneous quantification of *Bacteroides* HF183 and crAssphage CPQ_056 HFMGs in untreated sewage and environmental water samples without compromising sensitivity and reproducibility. The duplex assay was developed in the CSIRO laboratory. The newly developed duplex qPCR assay was validated in a collaborative laboratory study with the aim to evaluate reproducibility, sensitivity and concordance for field study.

The laboratories that took part in this validation study were CSIRO Land and Water laboratory at the Ecosciences Precinct, Dutton Park, Brisbane and the Sydney Water Laboratory Services laboratory, West Ryde, NSW. A graphical outline of this sub-study is provided in Figure 3-5.

Text, graphic and citations in this section were drawn from the peer-reviewed journal publication: Ahmed, W., Payyappat, S., Cassidy, M., Besley, C., 2019a. A duplex PCR assay for the simultaneous quantification of *Bacteroides* HF183 and crAssphage CPQ_056 marker genes in untreated sewage and stormwater. Environ. Int. 126, 252-259. https://doi.org/10.1016/j.envint.2019.01.035

Stored DNA extracted from sewage influent previously collected from a sewage treatment plant in Brisbane, Australia was used in development and validation of the duplex assay. Stored DNA extracted from previously collected water samples from the urbanised freshwater body of Lake Parramatta (Figure 3-4) for pilot sub-study 1 was used for the validation study conducted between the two laboratories (Ahmed et al., 2019a).

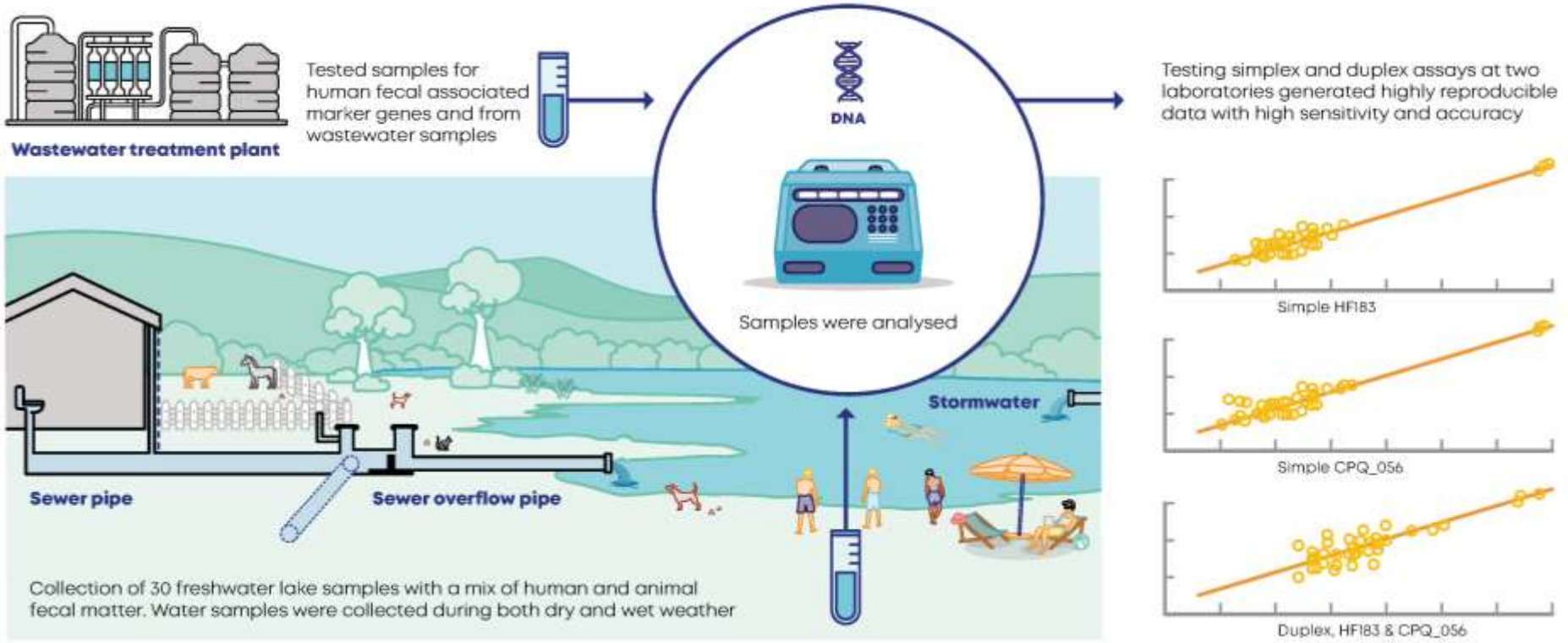


Figure 3-5: Graphical abstract outlining development of duplex assay based on stored DNA from Lake Parramatta waters and wastewater treatment plants

Main findings of pilot sub-study 2

The same primer and probe sequences were used in the duplex qPCR assay as used in published simplex qPCR assays for these two HFMG assays. The performance characteristics of the duplex qPCR assay were similar to its simplex counterparts (Ahmed et al., 2019a).

Outcomes of the collaborative laboratory validation study of field samples documented:

- concordance values between the simplex versus duplex qPCR assays for HF183 and crAssphage CPQ_056 marker genes ranged from 96.7 to 100%
- mean concentrations of HF183 and CPQ_056 in environmental water samples were remarkably similar or in some cases slightly greater for the duplex qPCR assay.

This suggested that the duplex assay would provide reliable data for quantifying HF183 and CPQ_056 simultaneously from a sample (Ahmed et al., 2019a).

Pilot sub-study 2 key outcomes

The duplex qPCR assay was shown to generate highly reproducible data with high sensitivity and accuracy.

To the best of the authors' knowledge, this is the first description of a duplex qPCR assay for simultaneous detection of HF183 and CPQ_056 HFMGs in sewage and environmental water samples.

The duplex assay is a valuable addition to the MST toolbox for sewage pollution monitoring and allows simultaneous sample analysis for two HFMGs to minimise the risk of false negative results in environmental receiving water samples contaminated with low levels of sewage.

This simultaneous sample analysis under the duplex assay is cost effective, being just over half the cost compared to that of two (single) simplex assays.

3.2.3 Pilot sub-study 3: Testing multiple assays in an urban freshwater lake

This sub-study applied a combined usage of multiple marker assays to investigate human and animal sources of faecal contamination in freshwater Lake Parramatta.

Objectives of pilot sub-study 3

The objectives of this study (Ahmed et al., 2019b) were to:

- determine the magnitude of sewage and animal faecal contamination in Lake Parramatta water samples collected during a dry weather period and from two storm events that coincided with WWOs with three HFMG assays [*Bacteroides* HF183, crAssphage CPQ_056 and pepper mild mottle virus (PMMoV)]
- determine the sources of animal faecal contamination from these same sample events from assessing three animal faeces-associated marker gene assays: cattle (CowM2); dog (*Bacteroides* BacCan-UCD); and avian (*Helicobacter* spp. associated GFD)
- determine the correlations between culturable faecal indicator bacteria (*Escherichia coli, Enterococcus* spp) and MST marker genes in these collected lake water samples
- compare the concentrations of HF183 and PMMoV to the established marker gene threshold values that would exceed the health risk benchmark for primary contact recreators

A graphical outline of this sub-study is provided in Figure 3-6.

Text, graphic and citations in this section were primarily drawn from the peer-reviewed journal publication: Ahmed, W., Payyappat, S., Cassidy, M. Besley, C., 2019b. Enhanced insights from human and animal host-associated molecular marker genes in a freshwater lake receiving wet weather overflows, Sci. Rep. 9, 12503. https://doi.org/10.1038/s41598-019-48682-4; and from

Preston, C.A. 1995. The impact of urbanisation on water quality in the Lane Cove River, Sydney New South Wales: a comparison of urban and non-urban catchments. Aust. Geogr. Stud. 33: 19-30. https://doi.org/10.1111/j.1467-8470.1995.tb00682.x

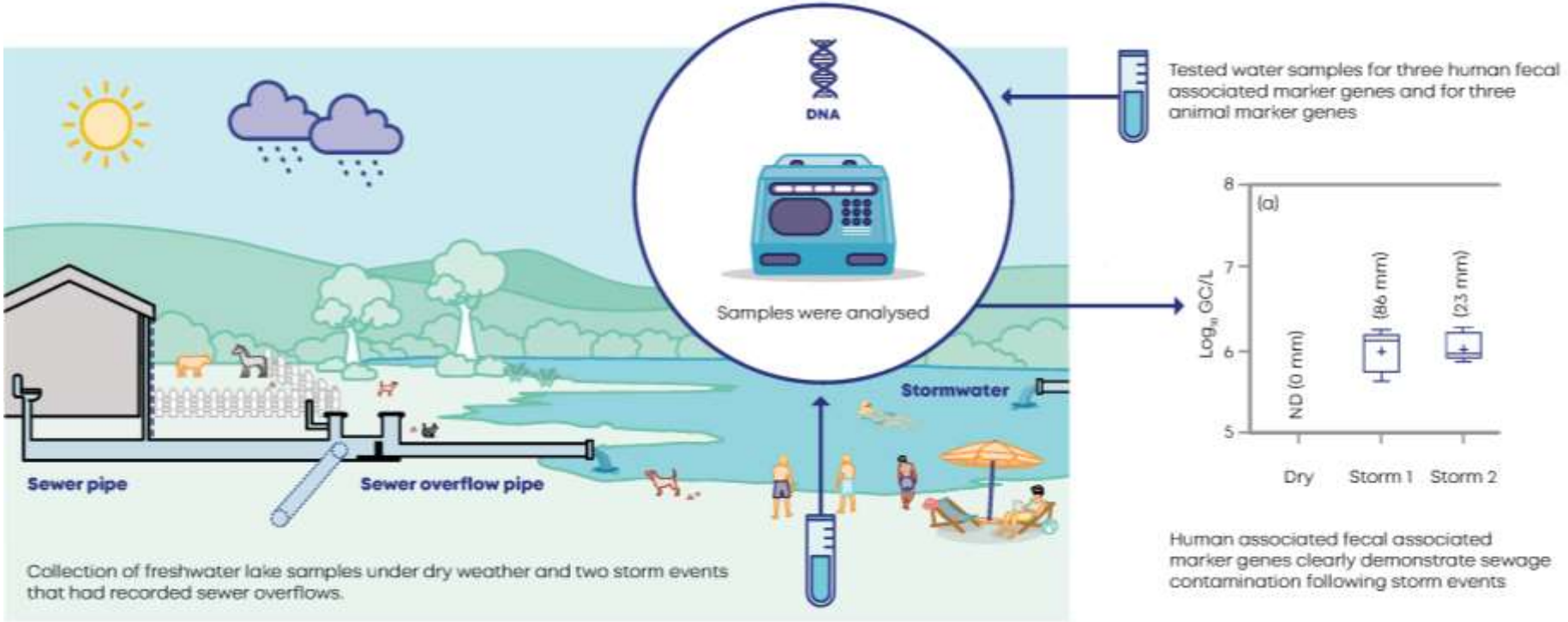


Figure 3-6: Graphical abstract outlining application of MST to distinguish human and animal faecal sources in an urban freshwater lake

Main findings of pilot sub-study 3

The magnitude of general and source-specific faecal pollution was low in water samples collected during dry weather compared to storm events when WWOs were recorded by gauging. The concentrations of HF183, crAssphage and PMMoV in water samples collected during storm events were relatively high, and in most stormwater samples exceeded the risk benchmark threshold values established in the literature for primary recreational contact. Moderate to strong positive correlations were observed among the quantitative occurrence of the three HFMGs (HF183, crAssphage and PMMoV). None of the samples tested were positive for the cowM2 (cow) marker gene, while BacCan-UCD (dog) and GFD (avian) animal associated markers were sporadically detected in water samples collected from both dry-weather and storm events.

The animal marker results aligned to observations of water birds observed under both dry- and wet-weather at the swimming site along with dogs accompanying owners while swimming. The wet-weather source of dog marker detections may be from washed-in animal faecal matter from the catchment. It was previously found by Preston (1995) in a control creek (without ERSs) within the Lane Cove area of Sydney that limited faecal inputs were from cats, dogs and native animals.

Pilot sub-study 3 key outcomes

Faecal indicator bacteria, such as enterococci, cannot separate human and animal faecal sources of contamination. Whereas the MST approaches employed in this sub-study provided insights into both human and two animal sources (dog and avian) of faecal contamination.

The results from multiple HFMG analysis clearly demonstrate that the study lake is impacted by sewage contamination. The source of sewage contamination was confirmed by gauging records of WWOs within the Lake Parramatta catchment.

Outcomes from this sub-study suggest that insight into the decay rates of HFMGs in relation to each other and enteric viruses would be beneficial in development of the QMRA for the Sydney region, as would development of Sydney-specific risk benchmark threshold values for primary contact recreators.

3.2.4 Pilot sub-study 4: Testing multiple assays in urban estuarine receiving waters

In the first three sub-studies, HFMGs were observed in storm samples together with no observation under dry conditions in urban receiving freshwaters, provided sufficient confidence in MST to progress the pilot studies into urban estuarine (saline) receiving waters to better understand sources of faecal contamination at swimming locations.

Objectives of pilot sub-study 4

The main aim of sub-study 4 (Ahmed et al. 2020a) was to assess the human health risk of sewage contamination from WWOs and dry-weather spills/leakage (if any) at estuarine locations. Water samples were collected from five or six sampling points at two different depths (0.5 m below the water surface and 1 m above the bottom surface) during a dry-weather period and compared with samples from two storm events that coincided with WWOs. To establish the magnitude of contamination, all water samples were analysed for established [*Bacteroides* HF183 and pepper mild mottle virus (PMMoV)] or novel [crAssphage CPQ_056 or Lachnospiraceae (Lachno3)] HFMGs.

A secondary aim was to determine the extent of animal faecal contamination in water samples by analysing animal faecal-associated marker genes of avian (GFD), dog (BacCan-UCD), cow (cowM2) and horse (HoF597) species.

A graphical outline of this sub-study is provided in Figure 3-7.

Text, graphic and citations in this section were drawn from the peer-reviewed journal publication: Ahmed, W., Payyappat, S., Cassidy, M., Harrison, N., Besley, C., 2020a. Sewage-associated marker genes illustrate the impact of wet weather overflows and dry weather leakage in urban estuarine waters of Sydney, Australia. Sci. Total Environ. 705, 135390. https://doi.org/10.1016/j.scitotenv.2019.135390 and from the following NSW government publications:

State of NSW and Department of Planning and Environment, 2022. State of the beaches 2021–2022: Statewide summary and how to read this report (nsw.gov.au). Environment and Heritage Department of Planning and Environment, ISBN 978-1-922840-67-7. Accessed November 8, 2023.

DPIE, 2020. Protocol for Assessment and Management of Microbial Risks in Recreational Waters. Department of Planning, Industry and Environment, NSW. Accessed 7 June 2024

NHMRC, 2008. National Health and Medical Research Council. Guidelines for Managing Risks in Recreational Water. Australian Government Publication Services ISBN 1864962666

State of NSW and Department of Planning and Environment 2021. State of the beaches 2020-2021: Sydney region (nsw.gov.au). Environment and Heritage Department of Planning and Environment, ISBN 978-1-922672-14-8. Accessed November 21, 2023.

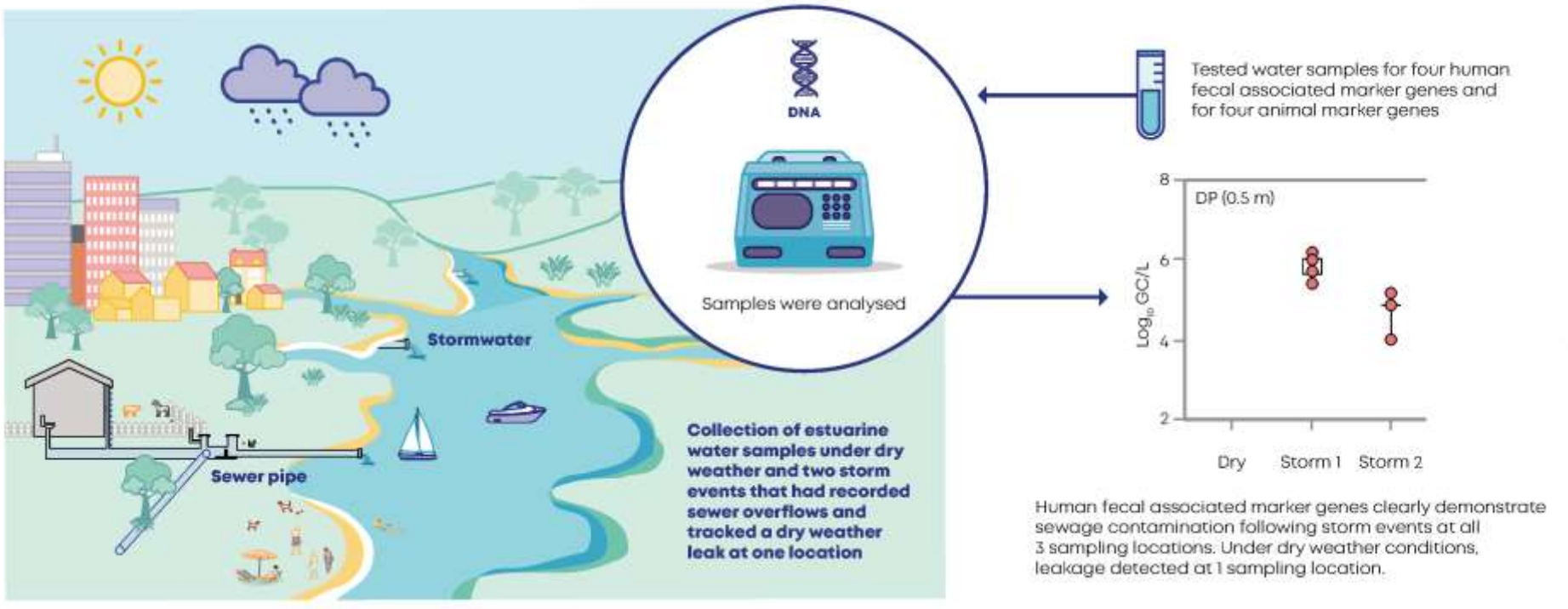


Figure 3-7: Graphical abstract outlining application of MST to distinguish human and animal faecal sources in urbanised estuarine waters at three locations

Estuarine pilot study locations

Three urbanised estuarine locations of Davidson Park, Hen and Chicken Bay, and Gymea Bay were chosen to represent differing hydrodynamic conditions that occur across the broader range of Beachwatch locations within the tidally drowned river valleys of the Sydney region. The study location of Davidson Park (Figure 3-8A and B) was situated within the channel of the upper Middle Harbour of Sydney Harbour (Port Jackson). In contrast, Hen and Chicken Bay (Figure 3-8C and D) was situated at the end of a relatively long embayment within the Parramatta River in the upper Sydney Harbour. While Gymea Bay (Figure 3-8E and F) is situated in a discrete embayment within mid-Port Hacking.

A secondary criterion for selection of these locations was the microbial assessment category from Beachwatch monitoring (State of NSW and Department of Planning and Environment, 2022), which is based upon the culturable faecal indicator bacteria of enterococci (Table 3-1). As all three locations had apparent sporadic bacterial contamination, as indicated by the microbial assessment category, this suggested that these three locations were most likely impacted by WWOs and would probably yield suitable study information.

Table 3-1: Microbial Assessment Categories of initial three estuarine study locations

Beach	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21	2021-22
Davidson Reserve	C	C	C	C	C	C	C
Cabarita Beach (Hen & Chicken)	B	B	B	B	B	B	B
Gymea Bay Baths	C	C	C	B	B	C	C

The Microbial Assessment Categories are determined from a modified 95th percentile from an enterococci dataset of at least 100 data points using Enterotester developed by Dr. Richard Lugg (DPIE, 2020). Category: A ≤40 cfu/100 mL, GI illness risk < 1%, AFR illness risk < 0.3%; B 41–200 cfu/100 mL, GI illness risk < 1–5%, AFR illness risk: < 0.3–1.9%; C 201–500 cfu/100 mL, GI illness risk < 5–10%, AFR illness risk < 1.9–3.9%; D > 500 cfu/100 mL, GI illness risk > 10%, AFR illness risk > 3.9%; GI = gastrointestinal illness risk per 100 exposures; AFR = acute fever and rash illness risk per 1000 exposures. See page 72-73 of NHMRC (2008) for notes associated with use of these threshold values

Gauges were installed into ERSs at locations of the sewerage network indicated in Figure 3-8. Companion telemetry installed with these gauges enable SMS text messages to be sent when ERSs were actively spilling to receiving waters of each catchment, to allow the dispatch of field crews for reactionary sampling of these wet-weather driven spill events.

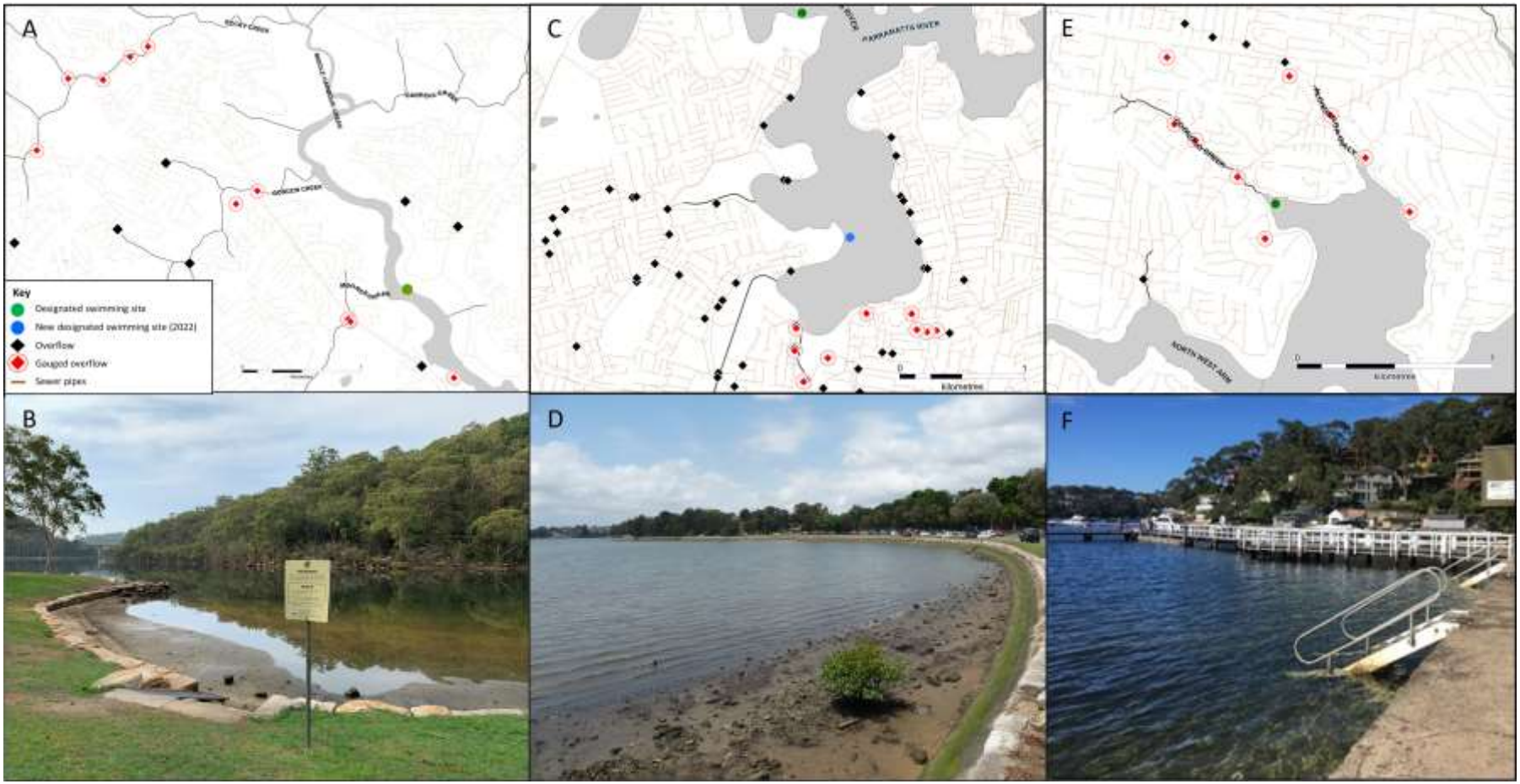


Figure 3-8: Initial three estuarine human health study locations

Davidson Park (A), location of overflows in relation to the designated swimming site, (B) designated swimming area; Hen and Chicken Bay (C) location of overflows in relation to the designated swimming site, (D) designated swimming area; Gymea Bay Baths (E) location of overflows in relation to the designated swimming site, (F) designated swimming area

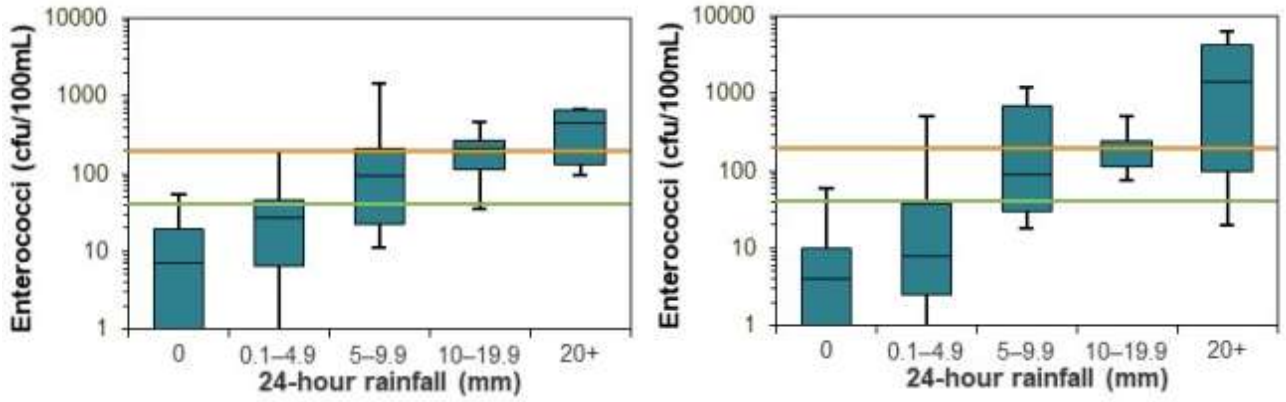
Main findings of pilot sub-study 4

The results from multiple HFMG analysis clearly demonstrated that the studied estuarine waters were contaminated by WWOs transported by urban stormwater runoff. This analysis of HFMGs showed greater (that is, 3–5 orders of magnitude) concentrations in water samples collected during the storm events compared to dry weather.

Water samples were also analysed for four animal faeces-associated marker genes to determine the extent of animal faecal contamination. Among the four animal faeces-associated marker genes, cow (cowM2) and horse (HoF597) could not be detected. In contrast, the avian (GFD) marker gene was most frequently detected in about 80% of sampled collected under both dry- and wet-weather conditions. While the dog (BacCan-UCD) marker gene was detected in about 27% of wet-weather samples from surface waters (within 0.5 m). Dairy cattle are known to be a reservoir for a wide range of zoonotic pathogens such as *Escherichia coli* O157:H7, *Cryptosporidium parvum*, *Campylobacter jejuni,* and *Salmonella* spp. (An et al., 2018; Stein and Katz, 2017; Vermeulen et al., 2017). Based on results of this sub-study 4, Ahmed et al, (2020a) concluded that risk from zoonotic pathogens in the studied estuarine waters was unlikely. The wet-weather detections of the dog marker gene suggested this animal faecal material was transported by urban stormwater runoff from the land into the estuary. The detection of the avian marker gene under both weather conditions is probably because the GDF marker represents a wide range of avian species such as coastal birds, waterfowl and land birds including chickens (Green et al., 2012).

A companion sanitary inspection survey conducted after a period of dry weather in the Gymea Bay catchment identified persistent leakage that seeped from a sewer carrier into the freshwater creek that flows into the estuary at Gymea Baths, with highest HFMG concentrations near the carrier and lower concentrations at the site situated at the baths. This trend was observed for both HF183 and CPQ_056 HFMGs from duplex assay testing employed under this survey and provided further confidence in the duplex assay. It also demonstrated the utility of HFMGs to distinguish the presence of a sewer leakage in Gymea Bay from animal associated faecal sources, which was not previously possible with culturable faecal indicator bacteria.

Under dry-weather conditions, HFMGs did detect the presence of leakage from the sewerage trunk main (1200 mm) crossing the creek feeding the Gymea Bay Baths arm of the bay. This provides a tool to supplement current Beachwatch assessments based on the much cheaper culturable faecal indicator bacteria of enterococci. Testing with the more expensive molecular HFMGs could be used to assess for potential issues as illustrated from comparison of Gymea Baths and Davidson Park swimming locations (Figure 3-9). The elevated median lines of the Gymea Baths box plots for dry and the lowest rainfall category (0.1 to 4.9 mm) reflect the seepage detected, while the clear step change documented for Davidson Park appears to reflect gauging records of WWOs commencing after around 10 mm of rainfall in 24 hours.



A. Gymea Baths

B. Davidson Park

Figure 3-9: Enterococci concentrations by rainfall category for: A Gymea Baths; B Davidson Reserve

from State of NSW and Department of Planning and Environment (2021)

Pilot sub-study 4 key outcomes

Monitoring with HFMGs successfully detected human faecal contamination in the receiving waters at the study locations, after WWOs were confirmed by sewer gauges installed at study ERSs.

In addition, a companion sanitary inspection survey conducted, by walking the feeder creek to the baths and collecting spot water samples, demonstrated monitoring with HFMGs were able to distinguish the presence of a leaking sewer impacting on Gymea Bay during dry-weather conditions. This monitoring was able to specifically detect human faecal contamination from dry-weather leakage. Whereas confirmation that this was from a human source was not possible with the current Enterococci culturable faecal indicator bacteria that has been widely used since 1993.

Animal faecal contamination does not appear to be an issue in the estuarine waters of these three locations.

This sub-study demonstrated and provided further confidence in the capability of the MST monitoring approach to understand sources (sewage or animal) of faecal contamination under both dry- and wet-weather conditions to provide useful information upon which to base management actions.

This sub-study recommended further research to obtain information on the decay rates of sewage-associated marker genes in relation to each other, enteric viruses and enterococci. This supports the previous recommendation under pilot sub-study 1 (as outlined above).

3.2.5 Pilot sub-study 5: Urban estuarine sediments as a measure of sewage contamination

Under molecular methods, water column samples are regarded as capturing the most recent event, whereas sediment samples are expected to incorporate a longer record of events. To explore the utility of MST analysis of sediment samples to inform sources of faecal contamination pilot sub-study 5 was undertaken.

Objectives of pilot sub-study 5

The main objective of Ahmed et al. (2020b) was to determine the prevalence and abundance of sewage (human faecal contamination) and animal faecal contamination of sediment at seven urban estuarine locations adjacent to swimming baths (Figure 3-11). Five replicate samples were collected on two occasions from each site.

A graphical outline of this sub-study is provided in Figure 3-10.

To evaluate animal faecal contamination, four animal faecal-associated marker genes assays were used in testing: avian (GFD); dog (DogBact); cow (cowM2); and horse (HoF597). While the assessment of sediment for human faecal contamination was explored with five HFMGs (*Bacteroides* HF183, Lachnospiraceae Lachno3, crAssphage CPQ_056 and *Methanobrevibacter smithii nifH*) and human adenovirus (HAdV), an enteric virus.

It is well understood that the faecal indicator bacteria (FIB) enterococci cannot differentiate between sources of faecal contamination, or with those naturally growing in the environment. Results from the inclusion of the molecular marker gene of enterococci (ENT 23S rRNA) enabled comparisons with other markers of this sub-study.

Text, graphics and citations in this section were drawn from the peer-reviewed journal publication: Ahmed, W., Payyappat, S., Cassidy, M., Harrison, N., Marinoni, O., Besley, C., 2020b. Prevalence and abundance of traditional and host-associated fecal indicators in urban estuarine sediments: Potential implications for estuarine water quality monitoring. Water Res, 184, 116109. https://doi.org/10.1016/j.watres.2020.116109

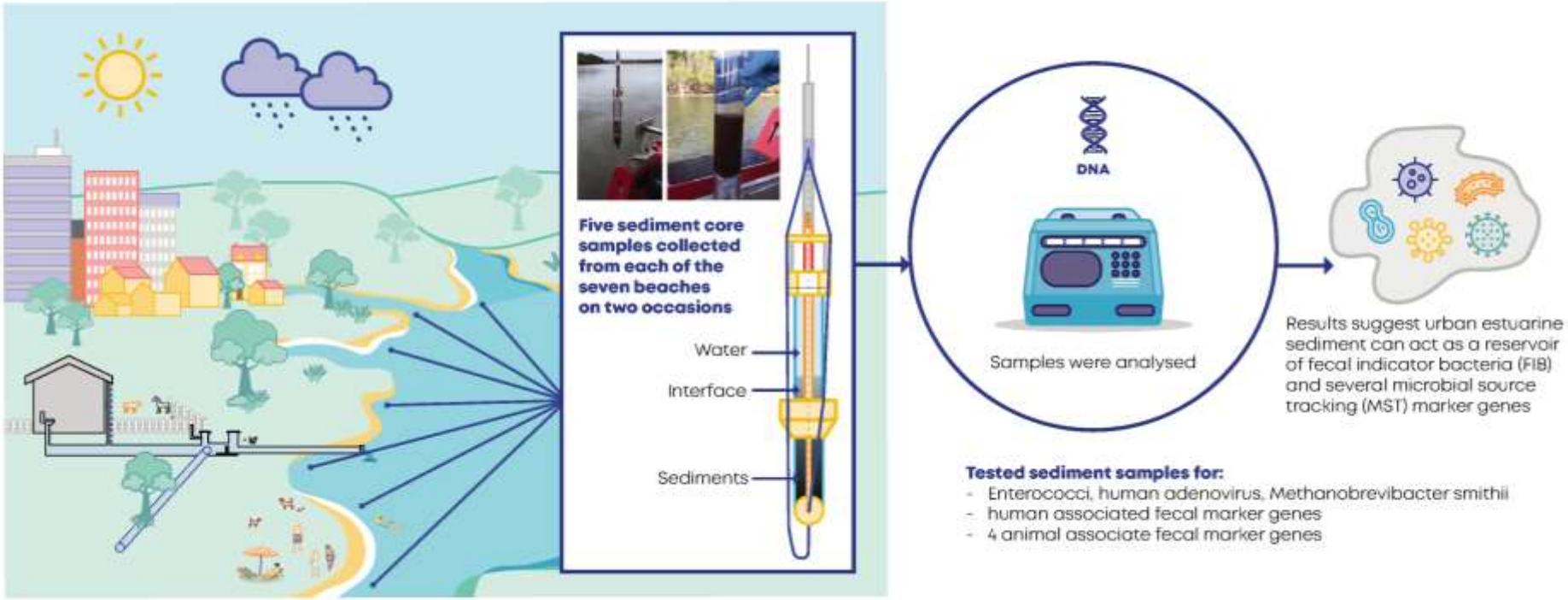


Figure 3-10: Graphical abstract outlining MST in urban estuarine sediments to assess the utility of sediments as a measure of sewage contamination

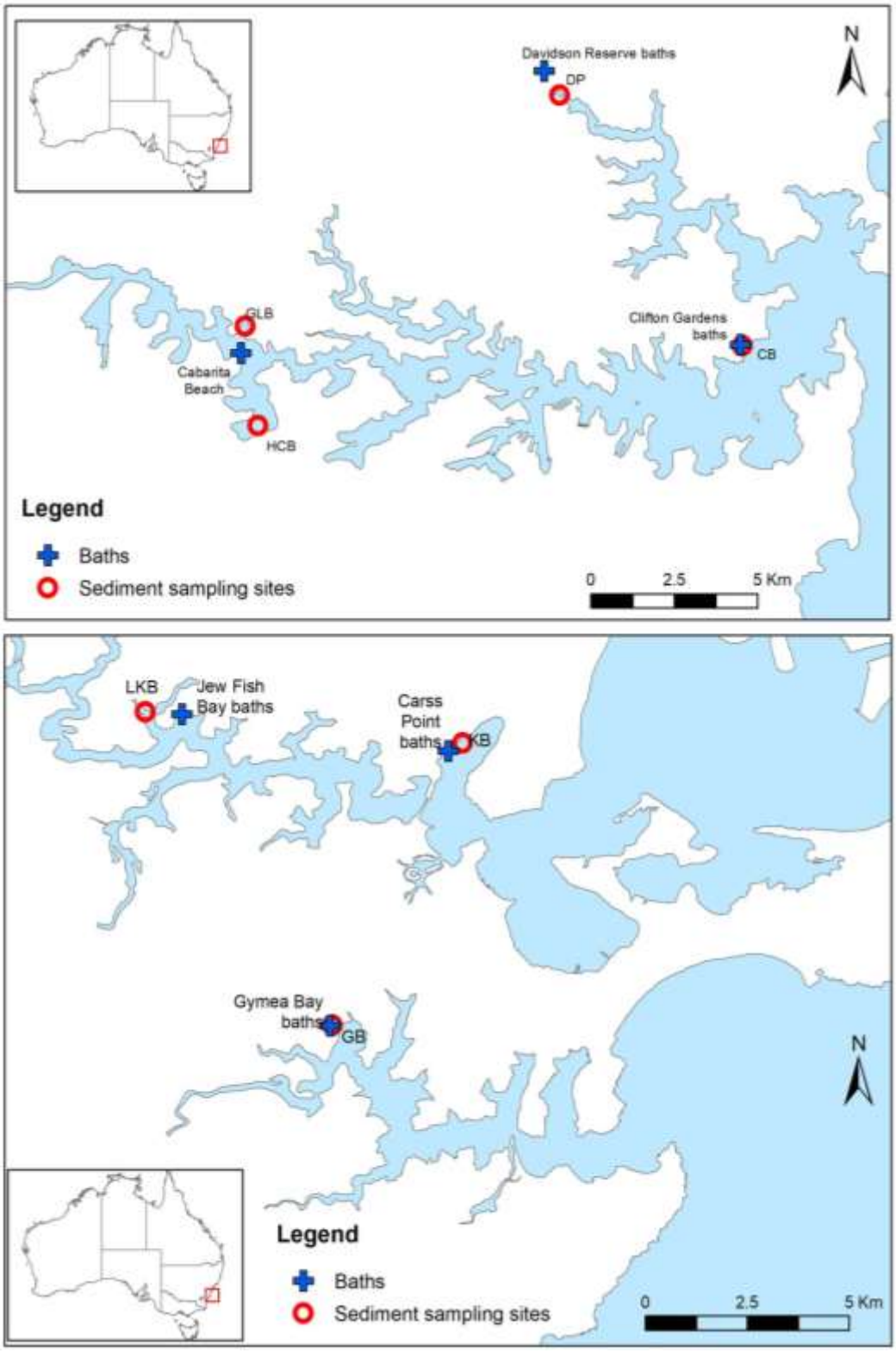


Figure 3-11: Maps showing sediment sampling locations and baths

(A) upper Sydney Estuary, Parramatta River, Sydney Harbour, Middle Harbour; (B) lower Georges River and Port Hacking

Main findings of pilot sub-study 5

The avian-associated marker gene GFD had the highest mean abundance of the markers tested in this sub-study and was widely detected (69.6%) in sediment samples from all seven locations compared to the other animal faeces-associated marker genes. The cowM2 marker gene could not be quantified and the DogBact and HoF596 marker genes across the seven urban estuarine locations were not detected. This high abundance and widespread detection of the GFD marker gene in sediment samples seems to be a sensible outcome as water column testing under sub-study 4 detected the GFD marker in about 80% of samples collected under both dry- and wet-weather conditions (Ahmed et al., 2020a).

Among the HFMGs, CPQ_056 was most prevalent in 63.8% of sediment samples. Lachno3 was detected in 30.4% of samples and HF183 was prevalent in 40.6% of samples, while *niFH* was not detected. The detection of at least two of the HFMGs occurred in 44.9% of sediment samples. In comparison to HFMGs, the overall incidence of HAdV was lower at 17.4%.

A total of 45 sediment samples were positive for the enterococci (ENT 23S rRNA FIB) marker gene, with 40% of these quantifiable. Similarly, 42.9%, 38.1%, 59.1%, and 43.8% of positive samples could be quantifiable for HF183, Lachno3, CPQ_056, and GFD marker genes, respectively. None of the HAdV PCR positive samples were quantifiable. This may reflect the concentrations of the HAdV in raw sewage are approximately 1 to 3 orders of magnitude lower than HF183, CPQ_056, and Lachno3 (Hughes et al., 2016).

The non-quantifiable detection of the HAdV marker gene did not always align with the detection of two or more sewage-associated marker genes. Moreover, the most frequent WWO exposure occurred at locations that did not have a consistent pattern of detection of the HFMGs and the enteric virus HAdV, suggesting sediments may not be a suitable measure of recent sewage contamination.

At an individual sample level, several sediment samples were negative for enterococci but positive for multiple marker genes suggesting that monitoring enterococci alone may not be useful to indicate source-specific faecal contamination in the sediment environment.

Based on the above evidence, unsurprisingly, enterococci (ENT 23S rRNA FIB) and avian (GFD) markers did not show any significant associations with overflow occurrence. This suggested sources other than WWOs and or dry-weather overflows (or leakage) when tested against overflow records for up to six weeks before or 12 weeks before sediment sample collection. Significant associations were determined between HF183, Lachno3, HAdV and cowM2 marker genes with WWOs occurrence in the preceding 6 and 12 weeks, while CPQ_056 was also significant when tested against the preceding 12 weeks.

The marker genes assessed in this sub-study belong to different groups of microorganisms, they decay differently in the environment and their abundances are different in contaminating sources. These factors need to be considered when applying these markers for faecal contamination tracking. Another consideration is that qPCR detects signals from both viable and non-viable cells. Moreover, lack of decay data for HFMGs in sediment makes it difficult to assess potential risks. A recent study reported that sewage-associated marker genes might persist in the sediment longer than the water column (Ahmed et al., 2019d cited in Ahmed et al., 2020a). More persistence data

will be required as the decay of these microorganisms and their DNA in sediment may be influenced by several factors such as porosity, grain size, clay content, oxygen conditions, reduced predation by protozoans, organic carbon, etc. It is also important to know what fraction of sediment borne MST marker genes are being released into the water due to resuspension. Such information will help to clarify whether MST marker genes indicate more recent or longer-term exposure to sewage contamination.

Pilot sub-study 5 key outcomes

Results from this sub-study showed that urban estuarine sediment can act as a reservoir of faecal indicator bacteria (FIB), HFMGs, animal marker genes, and human adenovirus.

This work suggested that sediments may not be a suitable matrix to measure relatively recent sewage contamination with HFMGs. However, further evaluation of variable decay rates of the different HFMGs in sediment may inform comment on potential frequency of sewage exposure at sites if differing decay rates occur between bacterial and viral-based marker genes. This aspect may provide a suitable approach to rank sites in capital solution planning.

To further explore the utility of sediments in understanding sewage contamination further studies seem justified to explore the role of:

- decay of HFMGs along with the decay of the avian animal marker gene (GDF) in sediment
- resuspension from sediment (during storm events to the water column) as a source of faecal contamination for both the avian animal marker gene and HFMGs.

3.2.6 Pilot sub-study 6: Interlaboratory accuracy and precision

Ebentier et al. (2013) argued that results obtained for environmental water samples must be reproducible in interlaboratory comparisons for their wider acceptance by water quality managers and for regulatory uses. Under this sub-study it was important to identify factors that may alter MST results. Such information is vital for the transfer of MST tools from research laboratories to major utilities and other water quality laboratories, particularly given that these laboratories can operate at regional scale, processing relatively large numbers of samples.

Objectives of pilot sub-study 6

An interlaboratory comparison was conducted between the CSIRO Land and Water laboratory at the Ecosciences Precinct, Dutton Park, Brisbane and the Sydney Water Laboratory services laboratory at West Ryde, NSW.

The main aim of this study (Ahmed et al., 2020c) was to assess interlaboratory precision and accuracy between two laboratories on a large number of duplicate environmental water samples collected from three estuarine and one freshwater site by using *Bacteroides* HF183, crAssphage CPQ_056, and pepper mild mottle virus (PMMoV) assays with different instruments and analyst experience levels. A graphical outline of this sub-study is provided in Figure 3-12.

Text, graphic and citations in this section were drawn from the peer-reviewed journal publication: Ahmed, W., Payyappat, S., Cassidy, M., Harrison, N., Besley, C., 2020c. Interlaboratory accuracy and precision among results of three sewage-associated marker genes in urban environmental estuarine waters and freshwater streams. Sci. Total Environ. 741, 140071. https://doi.org/10.1016/j.scitotenv.2020.140071

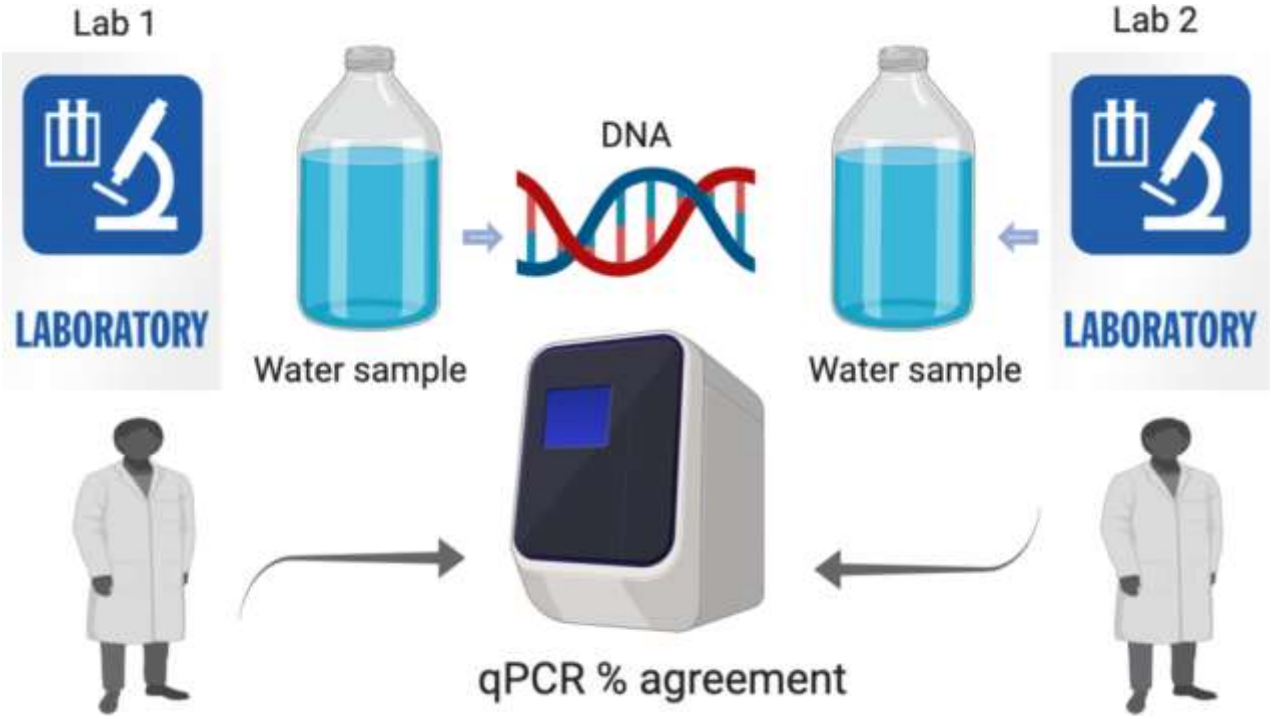


Figure 3-12: Graphical abstract of interlaboratory precision and accuracy study conducted between CSIRO and Sydney Water

Main findings of pilot sub-study 6

Precision

The relative percent difference of duplicate samples for all three assays represented a measure of precision. Fair to moderate agreements were documented that had acceptable relative percent difference values. Some discrepancies were observed on the quantitative detection of sewage-associated marker genes between two laboratories. In general, the CSIRO laboratory reported greater concentrations (<1 $\log_{10}$) of the sewage-associated markers in quantified pooled water samples compared to the Sydney Water laboratory.

Accuracy

As the basis of a measure of accuracy between laboratory testing, qPCR results of samples were designated into co-detection and non-codetection categories. Despite several dissimilarities in laboratory testing protocols between the two laboratories, >74% agreements (co-detection and non-codetection) were found for the HF183, CPQ_056 and PMMoV HFMGs for duplicate water samples collected from urban environmental estuarine waters and freshwater streams.

Correlations were assessed between both laboratories for each HFMG to establish the relationship between HFMG concentrations for all water samples. Significant moderate positive correlations were obtained for HF183 and PMMoV, while a strong positive correlation was returned for CPQ_056. This represented another measure of accuracy.

Factors contributing to discrepancies and improvement strategies

Several factors such as standards, qPCR platforms, PCR inhibitors, nucleic acid extraction efficiency and low levels of targets in some samples may have contributed to the observed discrepancies. Standardised protocols and equipment may improve detection and better align quantitation of the HFMGs in environmental water samples. In particular, digital PCR (dPCR) could offer a good additional precision and accuracy on the abundance of a target gene.

Some of these methodological limitations can be overcome by using multiple markers to obtain confirmatory results:

- testing at least duplicate samples when the contamination level is low
- increasing the volume of nucleic acid in a qPCR reaction
- increasing the number of qPCR technical replicates.

For scenarios, where qPCR is deemed imprecise, digital PCR (dPCR) could offer additional precision and accuracy on the abundance of a target gene. There are several advantages of dPCR such as:

- standards not being required
- highly tolerant to PCR inhibitors, and
- provides a linear response to the number of copies present to allow for small fold change differences to be detected.

Pilot sub-study 6 key outcomes

Under this sub-study, there was good overall agreement between the two laboratories for three HFMGs for duplicate water samples. To further improve accuracy and precision of results between laboratories the results presented in this sub-study highlight the importance of:

- a standardised protocol
- enhanced laboratory equipment (such as digital PCR)
- sample processing strategies
- appropriate quality controls.

Since this sub-study was conducted both laboratories have acquired digital PCR platforms.

3.2.7 Pilot sub-study 7: Concentrations of human faecal marker genes and enteric viruses in influent under dry weather conditions

An understanding of maximal concentrations of HFMGs and enteric viruses in untreated influent was investigated under pilot sub-study 7 as an input into the QMRA.

Objectives of pilot sub-study 7

The objectives of Ahmed et al. (2022) were to:

(i) measure the concentrations of four HFMGs (*Bacteroides* HF183, Lachnospiraceae Lachno3, crAssphage and pepper mild mottle virus (PMMoV), and three reference enteric viruses (HAdV 40/41, HNoV GI + GII and enterovirus) in untreated wastewater (influent) samples from inflows of two (Bondi and Cronulla) wastewater treatment plants (WWTPs) in Sydney, Australia using qPCR and RT-qPCR assays

(ii) assess correlations that may exist among HFMGs and enteric viruses in untreated influent

(iii) estimate ratios between HFMGs and enteric viruses and generate probability distribution functions for qPCR and RT-qPCR quantified enteric viruses to inform the QMRA

To document representative maximal concentrations of HFMGs and enteric viruses within influent, samples were collected on 12 occasions within a 12-month period with the criteria of no rainfall in the 24 hours preceding sample collection and 5 mm or less rainfall within 48 to 72 hours of sample collection. A graphical outline of this sub-study is provided in Figure 3-13.

Text, graphic and citations in this section were drawn from the peer-reviewed journal publication: Ahmed, W., Bivins, A., Payyappat, S., Cassidy, M., Harrison, N., Besley, C. 2022. Distribution of human fecal marker genes and their association with pathogenic viruses in untreated wastewater determined using quantitative PCR. Water Res, 226, 119093. https://doi.org/10.1016/j.watres.2022.119093

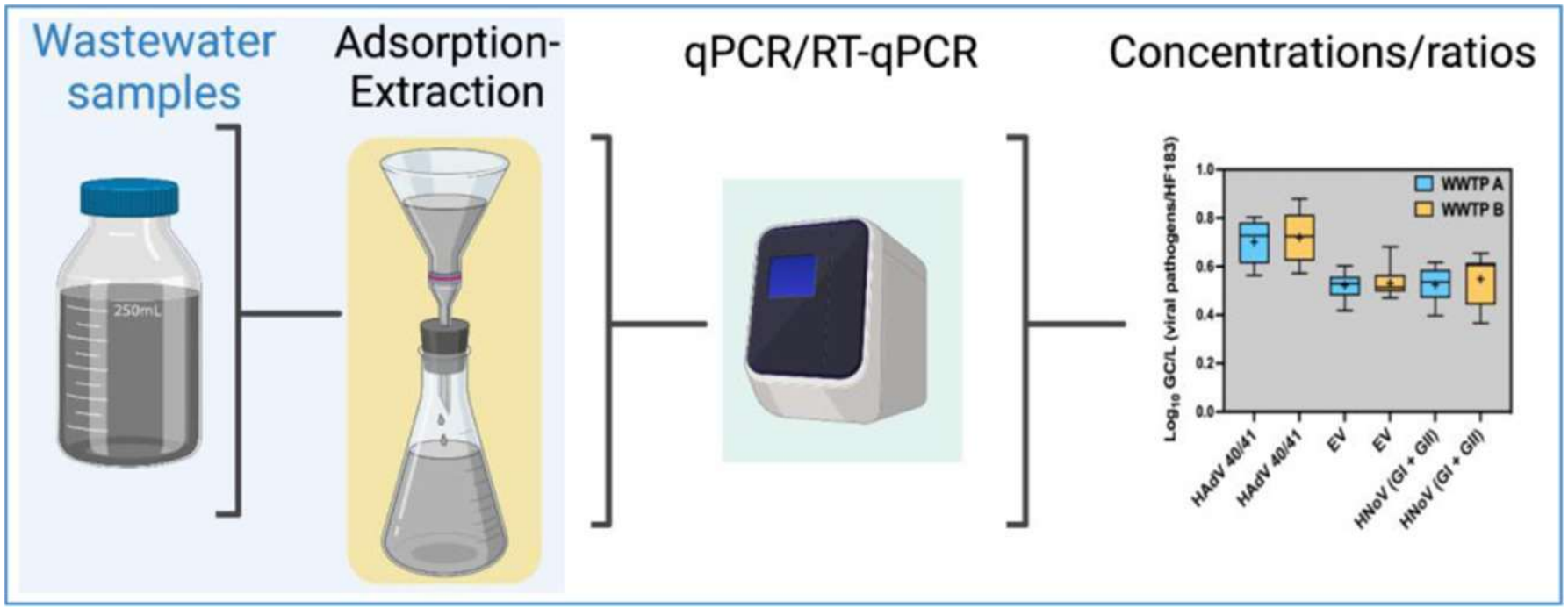


Figure 3-13: Graphical abstract outlining assessment of HFMG and enteric viruses in untreated sewer influent from two wastewater treatment plant locations collected under dry weather conditions

Main findings of pilot sub-study 7

In this study, the mean concentrations of Lachno3 HFMG was the greatest in untreated influent followed by crAssphage, HF183 and PMMoV in samples collected from inflows to both WWTPs.

The concentrations of HFMGs were quite stable over the course of the study with little variations observed within and between WWTPs. In contrast, concentrations of enteric virus gene copies were highly variable suggesting variable prevalence of the diseases in the community.

Over the course of the study, enteric viruses were 3–5 orders of magnitude lower than HFMGs in untreated influent.

Among the HFMGs, HF183, crAssphage and PMMoV correlated well with enteric virus gene copies, whereas weak or negative correlations were observed between Lachno3 and enteric virus gene copies. These correlation trends for Lachno3 may influence less robust outcomes under QMRA. While this may be an undesirable outcome, it does reflect the overarching strategy of this research in studying four HFMGs to end up with at least two HFMGs as lines of evidence for use in future monitoring.

While the two assessed WWTPs had dissimilar population service sizes, the ratios between $\log_{10}$ transformed enteric virus gene copies and HFMGs demonstrated similar central tendency and variability for the same combinations between WWTP A and WWTP B with no difference between the WWTPs. This suggests the widespread presence of these HFMGs in both populations serviced by these two WWTPs.

Pilot sub-study 7 key outcomes

Results of this sub-study (Ahmed et al., 2022) suggest the widespread presence of HFMGs in both populations serviced by the two WWTPs assessed, and this may also suggest consistency in presence of HFMGs within influent carried in the broader Sydney Water sewerage network.

The overall objective to raise probability distribution functions for HFMGs and reference enteric viruses was successfully achieved from correlations and ratios calculated from maximal concentrations of gene copies measured from influent samples collected under dry weather conditions across a 12-month period.

From the data obtained, Sydney specific probability distribution functions for HFMGs were raised to aid in applying the quantitative microbial risk assessment (QMRA) approach to prioritise sewage-contaminated receiving waters. The development of a Sydney region-specific QMRA is the end point of this series of sub-studies of human health risk from WWOs.

3.2.8 Pilot sub-study 8: Expanded testing in urban estuarine receiving waters along with assessment of host sensitivity and specificity

The initial estuarine receiving water study (Section 3.2.4, Sub-study 4, Ahmed et al., 2020a) established a link between WWOs and human faecal contamination of estuarine waters in the Sydney region. This study also demonstrated the capability of the MST monitoring approach to understand the sources of faecal pollution. That is, human or from an animal. However, host sensitivity and specificity of HFMGs and animal scat associated marker genes were not determined by Ahmed et al. (2020a; Section 3.2.4, Sub-study 4).

Objectives of pilot sub-study 8

The main objective of this study was to compare the host sensitivity and specificity, and concentrations of five HFMGs, two enteric viruses, and six animal scat-associated marker genes across faecal samples of influent collected from wastewater treatment plants (WWTPs) and from animal scats of urban catchments in the mega-coastal city of Sydney, Australia. The five HFMGs tested were *Bacteroides* HF183 (HF183), cross-assembly phage (crAssphage), Lachnospiraceae Lachno3 (Lachno3), *Methnobrevibacter smithii nifH* (*nifH*), and pepper mild mottle virus (PMMoV). The two enteric viruses were human adenovirus (HAdV), human polyomavirus (HPyV). The six-animal scat associated marker genes assessed were for: dogs *Bacteroides* BacCan-UCD and *Bacteroides* DogBact; ruminants *Bacteroides* BacR and CowM2; horses *Bacteroides* HoF597; and birds Helicobacter spp. GFD.

A secondary objective of this study was to assess urban estuarine water samples to understand differences in faecal pollution sources using HFMGs and enteric viruses for human faecal sources and animal faecal sources at ten urban estuarine locations. HFMGs in this testing were: HF183; crAssphage; Lachno3; and PMMoV. Enteric viruses tested were human adenovirus type 40/41 (HAdV 40/41) and enterovirus, while three animal scat-associated marker genes assessed were for dogs DogBact, horses HoF597 and birds GFD. Under estuarine water quality monitoring, several marker genes were omitted from those used to assess the main objective after considering their performance criteria, sanitary survey results and budgetary constraints.

A graphical outline of this sub-study is provided in Figure 3-14.

Text, graphics and citations in this section were drawn from the peer-reviewed journal publication: Ahmed, W., Payyappat, S., Cassidy, M., Harrison, N., Besley, C., 2023a. Microbial source tracking of untreated human wastewater and animal scats in urbanized estuarine waters. Sci. Total Environ. 877, 162764. https://doi.org/10.1016/j.scitotenv.2023.162764

Brown, K. I., Graham, K.E. Boehm, A.B. Risk-based threshold of gull-associated fecal marker concentrations for recreational water. Environ. Sci. Technol. 4, 44–48 (2017)

NHMRC, 2008. National Health and Medical Research Council. Guidelines for Managing Risks in Recreational Water. Australian Government Publication Services ISBN 1864962666

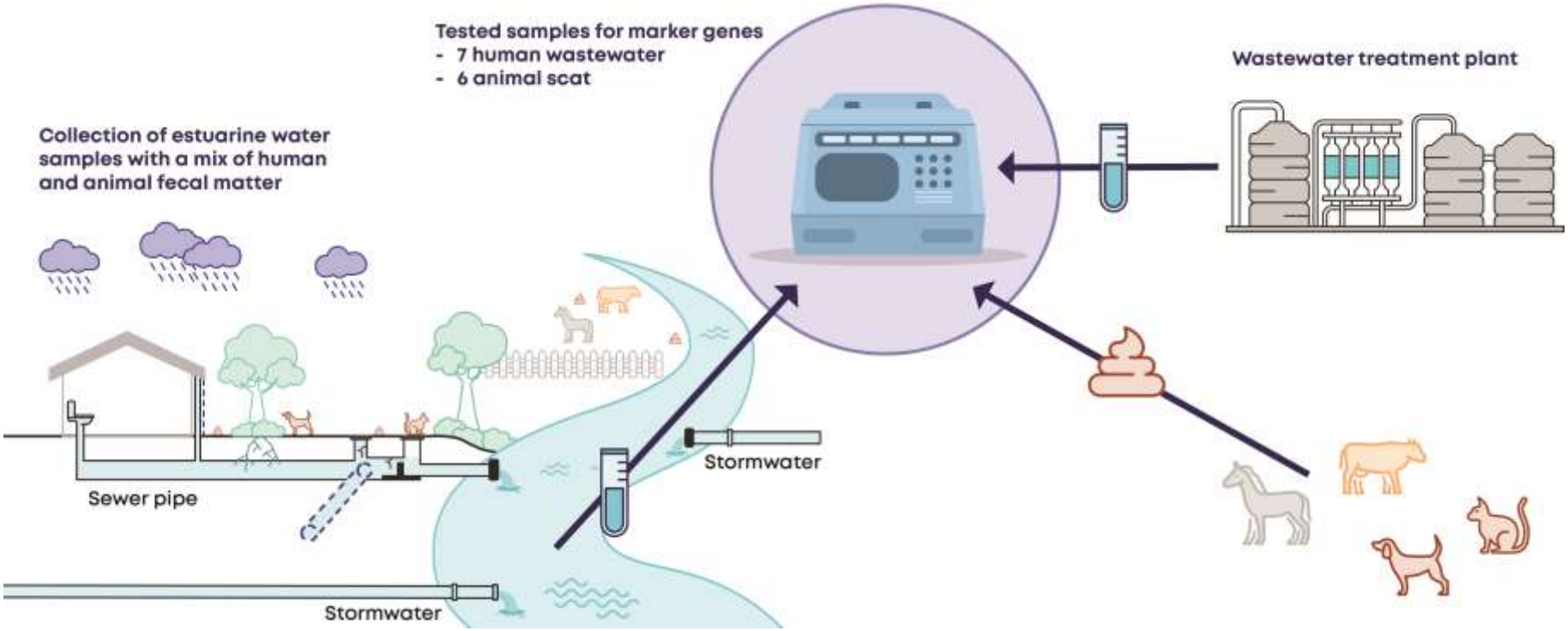


Figure 3-14: Graphical abstract outlining assessment of untreated sewer influent and animal scats for host sensitivity and specificity and the companion study that applied MST to distinguish faecal sources in urbanised estuarine waters at ten locations

Sample collection for testing under objective 1

To determine the host sensitivity and specificity of human- and animal faecal-associated marker genes, ten untreated influent samples were collected from ten WWTPs of the Sydney region and 97 fresh individual animal scat samples were collected from cats, dogs, ruminants (cows), horses, and birds (chickens).

Expanded estuarine pilot study locations for testing under objective 2

To understand differences in faecal pollution sources an additional ten estuarine swimming locations (Figure 3-15) were sampled across dry and wet-weather events with associated WWOs. A total of 135 water samples were collected. These locations encompass a range of microbial assessment categories (A, B, C and D) reported by State of NSW and Department of Planning, Industry and Environment (2021) based upon culturable enterococci concentrations. Swimming locations with microbial assessment categories that ranged from B to D (Table 3-2) were likely impacted from WWOs. The other criterion used in selection of these locations was the availability of already installed sewer gauging or the ability to install sewer gauge/s within the lead-in timeframe of the sub-study.

Table 3-2: Microbial Assessment Categories of expanded estuarine pilot study locations

Beach	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21	2021-22
Hayes Street Beach	B	B	C	C	C	B	B
Clontarf Pool	B	B	B	A	A	B	B
Gurney Crescent Baths	B	B	B	B	B	B	B
Rose Bay Beach	B	C	C	C	C	B	C
Tambourine Bay	B	C	B	B	B	C	C
Carss Point Baths	B	C	C	C	B	B	C
Yarra Bay	C	C	B	B	B	B	C
Oatley Bay Baths	B	C	C	C	B	B	C
Foreshores Beach	D	D	D	D	C	C	D
Kyeemagh Baths	B	B	B	B	B	B	C

The Microbial Assessment Categories are determined from a modified 95th percentile from an enterococci dataset of at least 100 data points using Enterotester developed by Dr. Richard Lugg (DPIE, 2020). Category: A ≤40 cfu/100 mL, GI illness risk < 1%, AFR illness risk < 0.3%; B 41–200 cfu/100 mL, GI illness risk < 1–5%, AFR illness risk: < 0.3–1.9%; C 201–500 cfu/100 mL, GI illness risk < 5–10%, AFR illness risk < 1.9–3.9%; D > 500 cfu/100 mL, GI illness risk > 10%, AFR illness risk > 3.9%; GI = gastrointestinal illness risk per 100 exposures; AFR = acute fever and rash illness risk per 1000 exposures. See page 72-73 of NHMRC (2008) for notes associated with use of these threshold values

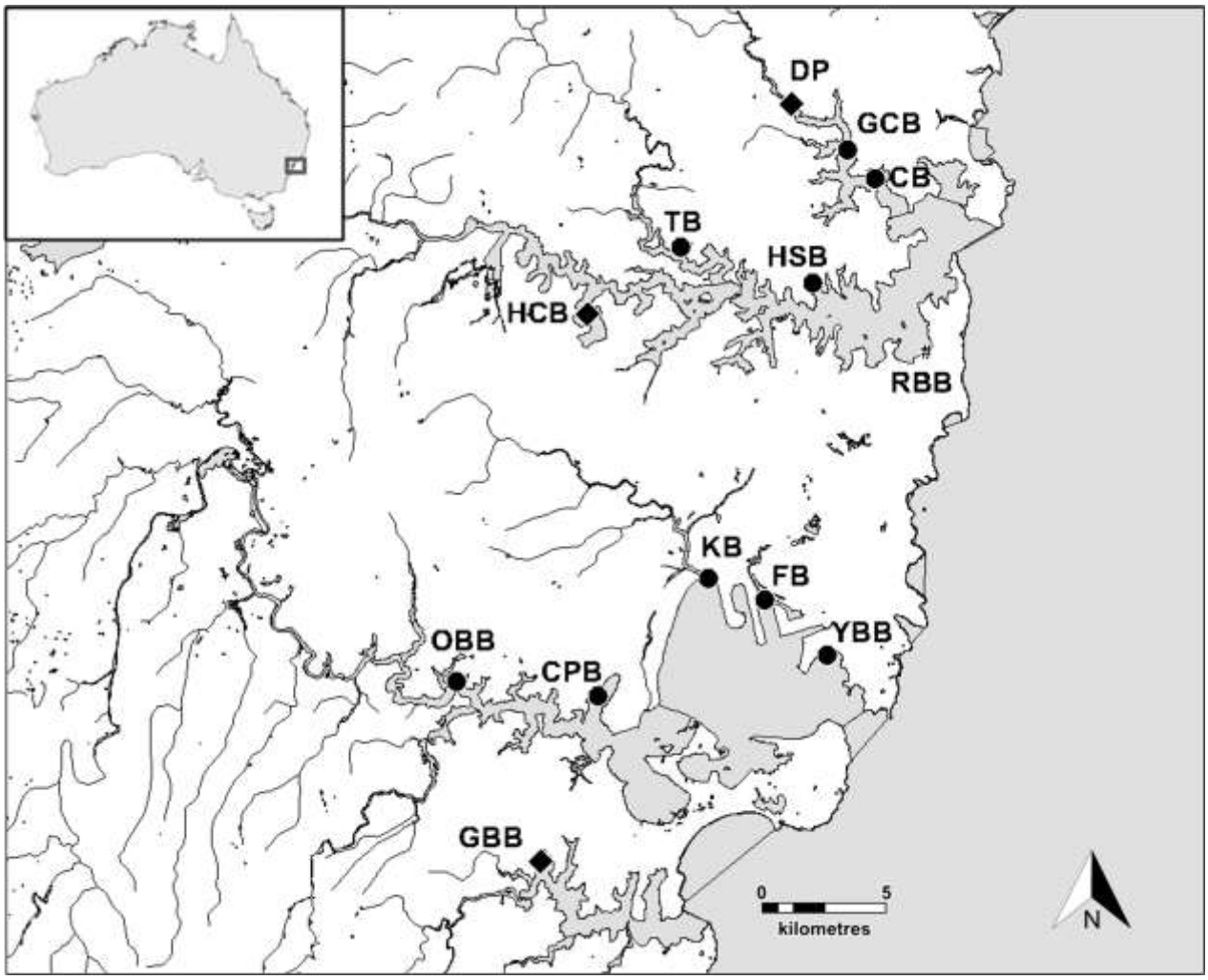


Figure 3-15: Additional 10 estuarine study locations.

Location coding on map: CPB = Carss Park Baths; CB = Clontarf Baths; FB = Foreshores Beach; GCB = Gurney Crescent Baths; HSB = Hayes Street Beach; KB = Kyeemagh Baths; OBB = Oatley Bay Baths; RBB = Rose Bay Beach; TB = Tambourine Bay; YBB = Yarra Bay Beach. Depth profile location PP = Pulpit Point.

Main findings of pilot sub-study 8

Host sensitivity

Absolute host sensitivity was exhibited across three criteria used to assess seven human faecal-associated marker genes of cross-assembly phage (crAssphage), human adenovirus (HAdV), *Bacteroides* HF183 (HF183), human polyomavirus (HPyV), Lachnospiraceae (Lachno3), *Methnobrevibacter smithii nifH* (*nifH*) and pepper mild mottle virus (PMMoV). In contrast, only the horse scat-associated marker gene *Bacteroides* HoF597 (HoF597) exhibited absolute host sensitivity.

Host specificity

The absolute host specificity value (of 1.0) was returned for the human faecal-associated marker genes of HAdV, HPyV, *nifH* and PMMoV for each of the three applied host specificity calculation criteria, while values of >0.9 were returned for crAssphage and Lachno3. Ruminants and cow scat-associated marker genes of BacR and CowM2, respectively exhibited the absolute host specificity value of 1.0.

Among the animal scat-associated marker genes, BacR, CowM2, and HoF597 exhibited high host specificity and appear to be suitable for ruminant, cow and horse faecal pollution. Since DogBact and GFD marker genes showed low host specificity, they should be synergistically used with other marker genes from dog and avian hosts.

Concentrations of WWTP influent and animal scats

Concentrations of Lachno3 were greater in most influent samples (collected from wastewater treatment plants) followed by crAssphage, HF183, *nifH*, HPyV, PMMoV and HAdV, respectively. Among the animal scat-associated marker genes BacCan-UCD was highest followed by DogBact, BacR, CowM2, GFD, and HoF597.

HFMGs were detected in several scat samples from cats and dogs, suggesting the use of at least two HFMGs under concordant sampling of animal scat-associated marker genes will be required to assist in interpretation of faecal sources in environmental waters.

Faecal pollution sources at ten estuarine locations

The presence of multiple HFMGs and enteric viruses in ten estuary locations demonstrate that estuarine waters were impacted by WWOs after the sampled storm/post storm events. Sewer gauging recorded WWOs during these storms. This outcome provided a similar wet-weather data pattern to those observed at the three locations under sub-study 4 (Section 3.2.4, Ahmed et al., 2020b), and provided further confidence in HFMGs as a tool for detecting human faecal contamination from WWOs.

Results of the current sub-study document the HFMGs of PMMoV and crAssphage had a greater prevalence, together with several samples with greater concentrations compared with the other two HFMGs suggesting more reliable detection of human faecal pollution from rainwater diluted influent spilt from WWOs into estuarine waters.

Comparison of quantifiable concentration data to literature illness risk threshold values reflected crAssphage and PMMoV were relatively well aligned across the various storm events sampled at sites with WWOs recorded. This suggests that the DNA- and RNA-based markers of crAssphage and PMMoV may be potentially more reliable indicators of illness risk than was illustrated by the quantifiable concentrations of the bacterial-based markers of Lachno3 and HF183 marker genes.

Under sub-study 4 (Section 3.2.4, Ahmed et al., 2020b), both the stated frequency of detection and concentrations of PMMoV marker gene in storm event samples were lower than HF183, Lachno3 and crAssphage marker genes from a study of three other estuarine locations in the Sydney region. That sub-study suggested that the PMMoV marker gene may be a less sensitive tool to detect diluted human faecal contamination in estuarine water compared to HF183, Lachno3 and crAssphage marker genes. However, in the current study, PMMoV was prevalent in all storm/post storm event (n = 92) samples while crAssphage, Lachno3 and HF183 had three, 11 and 18 less detections, respectively. This suggests PMMoV HFMG should be further assessed as a tool for detection of diluted human faecal contamination in estuarine waters.

All samples collected from the first two storm events (23 to 63 mm of rainfall) had detections of the enterovirus in contrast to a few sporadic detections of the HAdV 40/41. The third storm event sampled was a major east-coast low rain event (with up to 258 mm of rainfall), as the rainfall increased through this event, the prevalence of the enterovirus became more sporadic. Under other storm events, enterovirus was detected more often than the HAdV 40/41, which had only sporadic detections. Under more extensive rainfall conditions detection of the enterovirus became difficult compared to the HFMGs. This reduced detection is due to the fact the concentrations of HFMGs are 2–3 orders of magnitude greater in receiving waters than those of enteric viruses (such as enterovirus and HAdV 40/41).

In the current study for the eight locations sampled under dry-weather conditions, 43 of 45 samples were without detection of the four HFMGs or two enteric viruses. This same dry-weather data pattern was observed at the Davidson Park location under sub-study 4 (Section 3.2.4, Ahmed et al., 2020b), where WWOs were also not active. In the other two samples HF183 was detected. In one of these two samples, HF183 was recorded along with the avian marker (GFD), which may represent cross reactivity. The other HF183 detection was sampled at the Clontarf location situated nearby four marinas with sewage pump out facilities. This may have been the possible source of human faecal contamination. These dry weather outcomes provide further confidence in adoption of HFMGs for microbial source tracking.

The magnitude of animal faecal contamination in all estuarine locations was low compared to human faecal contamination. The current sub-study (Ahmed et al., 2023c) observed the same pattern seen in Ahmed et al. (2020b; sub-study 4 Section 3.2.4) for the GFD marker to be the most frequently detected animal marker under both dry-and storm-event conditions with 38 % prevalence in storm/post storm samples and 12 % of samples collected under dry-weather conditions. Our results are supported by Sala-Comorera et al. (2021) who documented the observation of localised seabird faecal pollution in a relatively small area compared to homogeneously distributed human faecal pollution identified with HF183 and crAssphage marker gene from an active wastewater outfall discharging into urban marine waters.

Pilot sub-study 8 key outcomes

Monitoring with HFMGs successfully detected human faecal contamination in the receiving waters of the study locations, after WWOs were confirmed by installed sewer gauges. These detections were observed repeatedly across storm events.

The detection of enteric viruses compared to HFMGs became difficult as more extensive rainfall occurred as was illustrated from results of the three storm events assessed under this sub-study.

Animal faecal contamination does not appear to be of concern in the estuarine waters when compared to the gull faeces associated *Catellicoccus* marker gene of Brown et al. (2017) at these three locations under either dry- or wet-weather conditions. Under wet-weather conditions the magnitude of human faecal contamination was much higher.

The study of an additional 10 estuarine locations under the current sub-study built further confidence to that previously obtained under sub-study 4 (Section 3.2.4, Ahmed et al., 2020a) of three estuarine locations. The current sub-study further demonstrated the capability of the MST monitoring approach to understand sources (human sewage or different animals) of faecal contamination under dry-weather and varying storm conditions, illustrating the depth of information that can be gathered to inform management actions.

The differing magnitude PMMoV results under the current sub-study (Ahmed et al., 2023a) and sub-study 4 (Section 3.2.4, Ahmed et al., 2020a) warrant further investigation to better understand this HFMG. A limitation of taking samples on a single day after a WWO (as conducted sampled under the current sub-study and sub-study 4) precluded assessment of reduction and decay of HFMGs across multiple days after a WWO into receiving waters. Research into the reduction and decay rates of HFMGs in relation to enteric viruses was conducted under sub-studies 9 to 11.

3.2.9 Pilot sub-study 9: Reduction of human faecal marker genes and enteric viruses in estuarine waters tracked across multiple days post WWOs

This sub-study (Ahmed et al., 2023b) presents an evaluation of the potential impact of sewage pollution on the quality of estuarine waters in Sydney with a specific focus on tracking the reduction of faecal indicator bacteria (FIB) enterococci, HFMGs and enteric viruses over multi-day sampling campaigns after WWOs had occurred.

Objectives of pilot sub-study 9

The main objective of this sub-study was to determine the presence and concentration of culturable enterococci, four HFMGs (*Bacteroides* HF183, Lachnospiraceae Lachno3, crAssphage and pepper mild mottle virus (PMMoV)) and four enteric viruses (human adenovirus 40/41 (HAdV 40/41), enterovirus, human norovirus GI (HNoV GI) and GII (HNoV GII)) at two estuarine locations in Sydney, NSW.

A graphical outline of this sub-study is provided in Figure 3-16.

Samples were systematically collected from two depths at each of three spatially separated subsites within the two estuarine locations over two separate six-day sampling campaigns, post a WWO event at each location. The collected samples were then analysed to determine the concentrations of the target HFMGs and enteric viruses. The results of the analyses were then compared to established gastrointestinal (GI) risk benchmarks, which have been calculated based on the number of expected GI illnesses per 1000 primary contact recreators.

Text, graphic and citations in this section were drawn from the peer-reviewed journal publication: Ahmed, W., Payyappat, S., Cassidy, M., Harrison, N., Besley, C. 2023b. Reduction of microbial source tracking marker genes and enteric viruses in Sydney estuarine waters receiving WWOs. Sci. Total Environ. 896, 165008. https://doi.org/10.1016/j.scitotenv.2023.165008

To support this sub-study, and sub-study 11, buoys were installed at Hen and Chicken Bay and at Davidson Park in Middle Harbour after regulatory approvals were in place from NSW government departments. Under this sub-study a Tinytag Aquatic 2 TG-4100 submersible temperature data logger was attached at 0.5 m water depth to each moored buoy.

The sampling campaign at Davidson Park was conducted during the mid-point of the swimming season while the sampling campaign at Hen and Chicken Bay was conducted between the start and mid-point of the swimming season. Median surface-water (0.5 m) temperature captured with submersible data loggers from the seven days preceding each of the wet-weather events were comparable to historical data for the swimming season periods between 1996 and 2007 (Section 3.1 of Ahmed et al., 2023b). The swimming season in the Sydney region is defined by Beachwatch as October 1 to April 30.

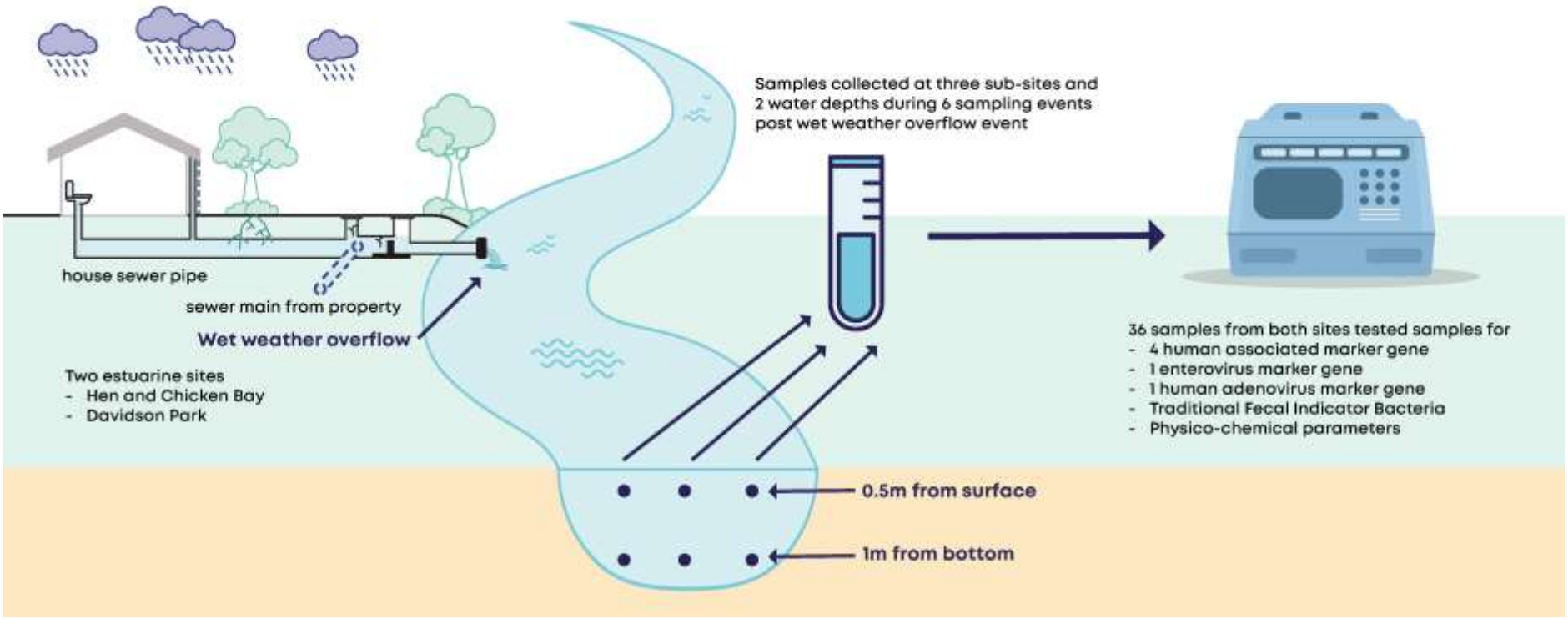


Figure 3-16: Graphical abstract overview of the multiday sampling campaign study to track reduction of HFMGs and enteric viruses in estuarine waters

Main findings of pilot sub-study 9

Among the four enteric viruses neither HNoV GI nor GII was detected, while 13.9 % of estuarine water samples had detections of enterovirus. Quantifiable concentrations for HAdV 40/41 were returned from 65.2 % of samples collected across the two locations and two depths with about two-thirds of these quantifications recorded in the surface layer samples. In contrast the presence of all four HFMGs (HF183, Lachno3, crAssphage, and PMMoV) was observed in all 36 (100 %) estuarine water samples collected from the surface layer from both locations. Detection frequencies of these markers were slightly lower at 1 m above the bottom surface.

WWOs were transported by stormwater into the receiving waters of the two estuarine study locations where the reduction of human faecal markers and enteric viruses was observed to differ at most subsites between the two water depths particularly on day 1. This was most likely influenced by environmental factors such as tidal movement and mixing. The concentrations of enterococci, HF183, Lachno3, and crAssphage markers were greater on days 1 and 2 at both Hen and Chicken Bay and at Davidson Park and steadily decreased over days 7 and 12, respectively.

The $\log_{10}$ reduction values obtained from the analysis exhibited substantial variability across different subsites. Water samples collected at a depth of 0.5 m below the water surface generally exhibited a greater reduction in microbial targets than those collected at a depth of 1 m above the bottom surface. Log_{10} reduction values are an accepted way to assess change in microbial indicators at each step of wastewater treatment. Whereas in this study, these values expressed change from maximal water column concentrations to the documented concentrations on the last day of each sampling campaign.

The concentrations of the human faecal markers were compared to established gastrointestinal (GI) risk benchmarks (Ahmed et al., 2023b). The concentrations of HF183, Lachno3 and crAssphage marker only exceeded the GI risk benchmark until day 3. While concentrations of PMMoV marker were indicative of exceedance of the GI risk benchmark on day 7 post WWOs that was much longer than indicated by culturable enterococci concentrations which were within this GI risk benchmark by day 2 and day 4 for the two locations, respectively.

Studies in the literature have calculated the decay rates of enterococci, human faecal markers and enteric viruses in estuarine waters by conducting microcosm/mesocosm or in-situ decay studies (Kay et al., 2005; Zhang et al., 2015; Mattioli et al., 2017). However, the present study poses difficulty in determining whether the observed decline in microbial targets stems from genuine decay or if it is a consequence of tidal impact that may have diluted the microbial targets or led to their deposition into sediments. Interestingly, we observed that the reduction of PMMoV was notably lower compared to other human faecal markers. This suggests that PMMoV may persist longer in environmental waters. Recorded median surface (0.5 m) water temperature in the seven days preceding the two sampling campaigns were 21.3°C for Hen and Chicken Bay and 25.9°C for Davidson Park. Specifically, it has been shown that PMMoV RNA can remain detectable in seawater for up to 1 week at temperatures between 31 and 33°C, and in river water (freshwater) for approximately 10 days at 25°C and up to 21 days at 4°C (Hamza et al., 2011; Rosario et al., 2009). These findings have important implications for the management of environmental waters, as they suggest that PMMoV may be more resilient to degradation than other human faecal markers in environmental waters.

Pilot sub-study 9 key outcomes

Sub-study 9 (Ahmed et al., 2023b) documented reduction of FIB (enterococci), four HFMGs (*Bacteroides* HF183, Lachnospiraceae Lachno3, crAssphage and pepper mild mottle virus (PMMoV)) and the enteric virus HAdV 40/41 during a sampling campaign spanning seven and 12 days. However, determining whether this decline resulted from natural decay or was influenced by tidal effects, which could have diluted the microbial targets or caused their deposition into sediments, proved challenging.

Results of sub-study 9 highlight the need for further studies to enhance understanding of the mechanisms responsible for the reduction of faecal indicator bacteria, human faecal marker genes and enteric viruses in estuarine waters. Sub-studies 10 and 11 were undertaken to investigate these mechanisms.

3.2.10 Pilot sub-study 10: Decay of human faecal marker genes and enteric viruses in laboratory microcosms

As advocated in sub-studies 1, 3, 8 and 9 (Ahmed et al., 2018; 2019b; 2023a; 2023b) an understanding of the decay kinetics of HFMGs in relation to most notably, enteric viruses were investigated under sub-study 10 to provide Sydney region specific inputs into the QMRA.

Decay of pathogens and other microorganisms refers to the reduction in their numbers over time due to complex processes that depends on many factors, including temperature, pH, sunlight, exposure, salinity, predation, and the presence of organic matter. These factors can affect the survival and growth of pathogens in water, as well as the rate at which they decay (Nelson et al., 2018; Korajkic et al., 2019; Tiwari et al., 2023).

Objectives of pilot sub-study 10

The main objective of Ahmed et al. (2023c) was to determine the decay rates of four HFMGs and four enteric viruses in laboratory microcosms mimicking estuarine water environments in temperate Sydney, NSW, using qPCR and RT-qPCR assays. A secondary objective was to determine the factors that influenced the decay of HFMGs and enteric viruses in the study microcosms. A graphical outline of this sub-study is provided in Figure 3-17.

In the first trial of microcosm experiments (Batch 1 of Figure 3-17), unfiltered (microorganisms present) and filtered (microorganisms likely absent) estuarine water dialysis tubes (filled with estuarine waters and human wastewater) were exposed to shaded condition and two different temperatures in controlled water baths (filled with estuarine water) representing the mean spring (19.4°C) and summer (25.5°C) estuarine water temperatures in Sydney, NSW (Sydney Water, unpublished data). Examination of spot (grab) surface water samples collected between 1996 and 2007 indicated the temperature at the start of the swimming season (in spring) was lower than that at the end of the swimming season (in autumn).

In the second trial of microcosm experiments (Batch 2 of Figure 3-17), unfiltered estuarine water dialysis tubes filled with estuarine water and human wastewater were exposed to artificial sunlight and dark conditions at 19.4°C and 25.5°C.

Text, graphic and citations in this section were drawn from the peer-reviewed journal publication: Ahmed, W., Payyappat, S., Cassidy, M., Harrison, N., Besley, C. (2023c). Comparing the decay of human wastewater-associated marker genes and enteric viruses in laboratory microcosms simulating estuarine waters in a temperate climatic zone. Sci. Total Environ. 908, 167845. https://doi.org/10.1016/j.scitotenv.2023.167845

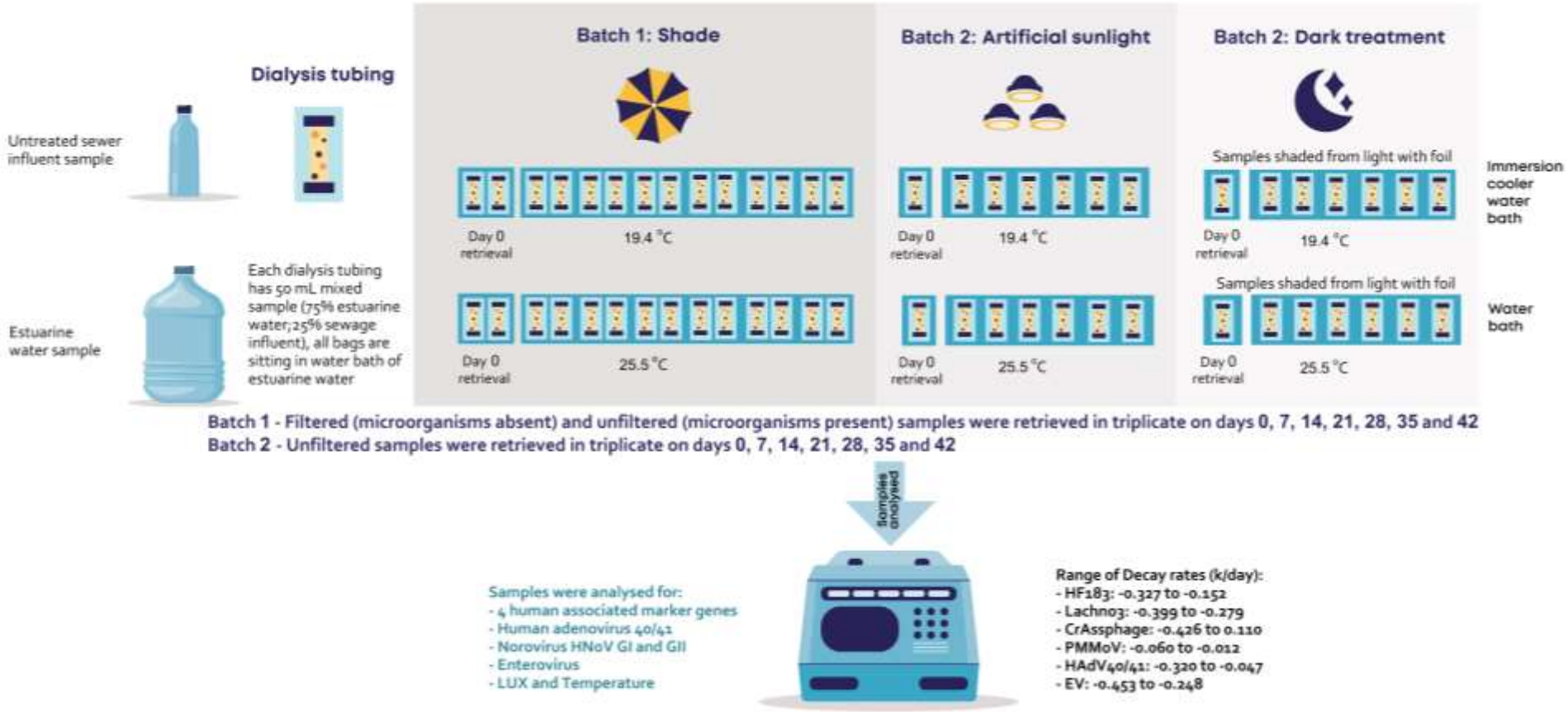


Figure 3-17: Graphical abstract illustrating the laboratory microcosm decay study

Main findings of pilot sub-study 10

Among the HFMGs and enteric viruses, only the viruses NoV GI and GII could not be detected in untreated human wastewater samples and therefore, their decay rates could not be determined in this study.

The study found that HF183, Lachno3, crAssphage, PMMoV, HAdV 40/41, and enterovirus exhibited varying decay rates depending on the experimental conditions. The average T90 values ranged from a few days to several months, indicating the rapid decay or prolonged persistence of these markers and enteric viruses in the estuarine environment.

In a broader context across both experiments, both bacterial HFMGs (HF183 and Lachno3) displayed considerably higher rates of decay when contrasted with the viral HFMGs across all treatment conditions. Although among viral HFMGs, crAssphage decay rates were relatively faster when compared to PMMoV. Enteric viruses tended to decay more rapidly than the documented extended persistence of PMMoV.

Notable outcomes under the Batch 1 experiment found that the presence of microbiota and temperature significantly influenced the decay rates of HF183 and PMMoV. While under the Batch 2 experiment, the exposure to artificial sunlight significantly accelerated the decay rates of bacterial HFMGs (HF183 and Lachno3), viral HFMGs (crAssphage and PMMoV), and enteric viruses (HAdV 40/41, and enterovirus).

Pilot sub-study 10 key outcomes

This sub-study characterises the decay rates of HFMGs and enteric viruses under different experimental conditions that mimicked temperate environmental conditions of the Sydney region. The findings contribute to our understanding of the fate and persistence of these markers in the environment, including the key observations that the exposure to artificial sunlight and elevated temperature were found to expedite the degradation of various HFMGs and enteric viruses. These outcomes highlight that the inclusion of decay rates are crucial in utilising a QMRA framework to consider the human health risks associated with environmental waters contaminated with sewer influent.

From the four HFMGs it could be possibly expected that PMMoV may be less robust as a surrogate indicator of enteric viruses due to the observed extended persistence of the viral HFMG PMMoV. This suggests that the decay of PMMoV may not accurately reflect the presence of enteric viruses associated with human faecal contamination. While this may be an undesirable outcome, it does reflect the benefit from the overarching strategy of this research in studying four HFMGs to carry forward at least two HFMGs as lines of evidence in future monitoring.

3.2.11 Pilot sub-study 11: Decay of human faecal marker genes and enteric viruses in field microcosms

To account for potential variations between decay rates observed under laboratory conditions under sub-study 10 (Ahmed et al., 2023c) and those occurring in natural environments, it was considered advisable to conduct a further in-situ study to assess the decay of HFMGs and enteric viruses within estuarine water. This would provide a more accurate understanding of the degradation processes in real-world conditions and enable a more comprehensive evaluation of the fate and persistence of these markers and viruses in the estuarine environment of the Sydney region.

Objectives of pilot sub-study 11

The main objective of Ahmed et al. (2024a) was to determine the in-situ decay rates of four HFMGs and four enteric viruses in estuarine water environments in temperate Sydney, NSW, using qPCR and RT-qPCR assays. A secondary objective was to compare the decay rates of targets obtained for in-situ decay to the previously published laboratory microcosms decay rates (Ahmed et al., 2023c) to understand how long pathogens can persist in different settings. A graphical outline of this sub-study is provided in Figure 3-18.

Two submersible mesocosm devices were constructed and enclosed with plastic mesh wire to protect dialysis tubes from potential damage caused by floating debris. The small pore size of the dialysis tubes prevents nutrients and microorganisms larger than 6–8 kDa from diffusing across the membrane. Each dialysis tube was filled with 75:25% (v/v) of estuarine water: untreated sewage so that enteric viruses whose concentration are low, can be measured. Regulatory approvals were obtained for placement of buoys into the Sydney estuary at former study locations of Hen and Chicken Bay and Davidson Park to allow fixed placement of these submersible mesocosm devices (Figure 3-18). The higher turbidity levels of Hen and Chicken Bay provide a differing setting to that of Davidson Park.

A sub-objective of sub-study 11 was to examine the impact of sunlight exposure on the decay of HFMGs and enteric viruses. Consequently, each device was divided into two sections: the lower half was covered with a robust black plastic sheet to create a dark treatment environment while the upper half remained uncovered to receive sunlight exposure (Figure 3-18). During deployment, the dialysis bags for the light treatment were submerged approximately 0.5 m below the water's surface, while the dark treatment bags were submerged approximately 1.5 m below the surface.

Text, graphic and citations in this section were drawn from the peer-reviewed journal publication: Ahmed, W., Gabrewold, M., Payyappat, S., Cassidy, M., Harrison, N., Besley, C. 2024a. Assessing the nucleic acid persistence of human wastewater markers genes and enteric viruses in estuarine waters in Sydney, Australia. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2024.171389

Figure 3-18: Graphical abstract illustrating the field mesocosm decay study

Main findings of pilot sub-study 11

When compared to laboratory microcosms replicating estuarine water conditions (Section 3.2.10, Ahmed et al., 2023c), most HFMGs and viruses decayed faster in Davidson Park and Hen and Chicken Bay mesocosms. However, some exceptions, such as crAssphage in sunlight and enterovirus, displayed faster decay in laboratory microcosms.

Sunlight exposure had a notable impact on the decay rates of target organisms in both Davidson Park and Hen and Chicken Bay mesocosms:

- HFMGs and enteric viruses decayed faster under sunlight, with HF183, Lachno3, and enterovirus showing particularly accelerated decay in Davidson Park mesocosms situated in the less turbid settings
- the decay rates of HFMGs and enteric viruses varied significantly between Davidson Park and Hen and Chicken Bay mesocosms. HF183 consistently exhibited faster decay in both sunlight and dark conditions, whereas crAssphage decayed faster in Hen and Chicken Bay mesocosms under sunlight but displayed similar rates in the dark compared to Davidson Park mesocosms
- in general, the decay of enteric viruses was slower than the bacterial targets (HF183 and Lachno3) but faster than the viral targets (PMMoV and crAssphage). Although PMMoV exhibited the slowest decay and was most persistent of all targets tested.

The study concluded that decay rates vary between mesocosms, emphasising the impact of sunlight exposure, which was potentially influenced by the elevated turbidity at Hen and Chicken Bay. This reflects that the deliberate selection of differing settings for these two study locations has provided valuable insights into decay of HFMGs and enteric viruses in Sydney estuarine waters.

The following discussion text from Ahmed et al. (2024a) describes overall context of sub-study 11 findings with other studies. Under sub-study 11, PMMoV persisted longer than any other HFMG or enteric viral target in both mesocosms. This is similar to the observation under sub-study 10 (Section 3.2.10, Ahmed et al., 2023c), and consistent with its persistence through wastewater treatment train (Rosario et al., 2009; Hamza et al., 2011; Kitajima et al., 2014; Symonds et al., 2014). Combination of extended persistence and frequent detection in wastewater and surface waters (Kitajima et al., 2018), make PMMoV a conservative indicator of human faecal contamination in the environment. On the other hand, bacterial HFMGs (HF183 and Lachno3), generally decayed faster than viral HFMGs and viral enteric pathogens. This observation is consistent with our earlier work, and that of others (Ahmed et al., 2023c; Mattioli et al., 2017), and it is not surprising considering the inherent differences between the taxonomic groups. Interestingly, the decay of crAssphage was closest to that of enteric viruses suggesting that it may be the better conservative surrogate for viral pathogens in temperate estuarine environments (such as the Sydney region). Finally, all three enteric viral pathogens decayed similarly irrespective of the study site and exposure to sunlight. This finding is important as it suggests that a singular viral indicator, such as crAssphage, has the potential to be an adequate surrogate for multiple viral pathogens.

Pilot sub-study 11 key outcomes

Sub-study 11 outcomes along with those from sub-studies 9 and 10 provide a weight of evidence that PMMoV is a conservative indicator of human faecal contamination in the environment due to its persistence in the estuarine environment.

In general, the decay of enteric viruses was slower than the bacterial targets (HF183 and Lachno3), but faster than the viral targets (PMMoV and crAssphage). Although the decay of crAssphage was closest to that of enteric viruses suggesting that it may be a better conservative surrogate for viral pathogens in temperate estuarine environments of the Sydney region.

The decay rates of HFMGs and enteric viruses generated in this study will inform Sydney site-specific risk-based thresholds (RBTs) of HFMGs which are derived from the QMRA framework. The Sydney specific QMRA is discussed in Section 3.3.

3.3 Sydney region specific QMRA informed by 11 sub-studies

The application of quantitative polymerase chain reaction (qPCR) based MST marker genes are increasingly being used to identify faecal contaminating sources and inform management decisions. This technique has been the basis of collaborative research with CSIRO conducted under the human health pilot study of the WWOM. Outcomes of this research have been progressively documented (from 2018 to 2024) in high quality peer review journals as summarised above under the 11 sub-studies (Section 3.2). These sub-studies have supported the development of a Sydney-specific QMRA framework that developed risk-based thresholds (RBTs) for each of the four HFMGs studied in estuarine receiving waters.

The QMRA framework is applied to simulate levels of health risks stemming from the ingestion of estuarine recreational waters contaminated with diluted untreated sewer influent from WWOs, that contain enteric pathogens. An outline of the QMRA framework is summarised in Section 3.3.1 (Ahmed et al., 2024b). The key output from the QMRA framework is the simulated health risks of GI (gastrointestinal) illness referred to as RBTs reported for each of the four HFMGs (HF183, Lachno3, crAssphage, and PMMoV) considering site-specific decay rates of HFMGs and the reference pathogen human norovirus (HNoV). RBTs outcomes are documented in Section 3.3.1.

Text, graphics and citations in this section were drawn from the peer-reviewed journal publication: Ahmed, W., Schoen, M.E., Soller, J., Harrison, J.C., Hamilton, K.A., Gebrwold, M., Simpson, S.L., Payyappat, S., Cassidy, M., Harrison, N., Besley, C.H. 2024b. Site-specific risk-based threshold (RBT) values for sewage-associated markers in estuarine bathing waters. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2024.172448

The graphical abstract of this paper is reproduced in Figure 3-19.

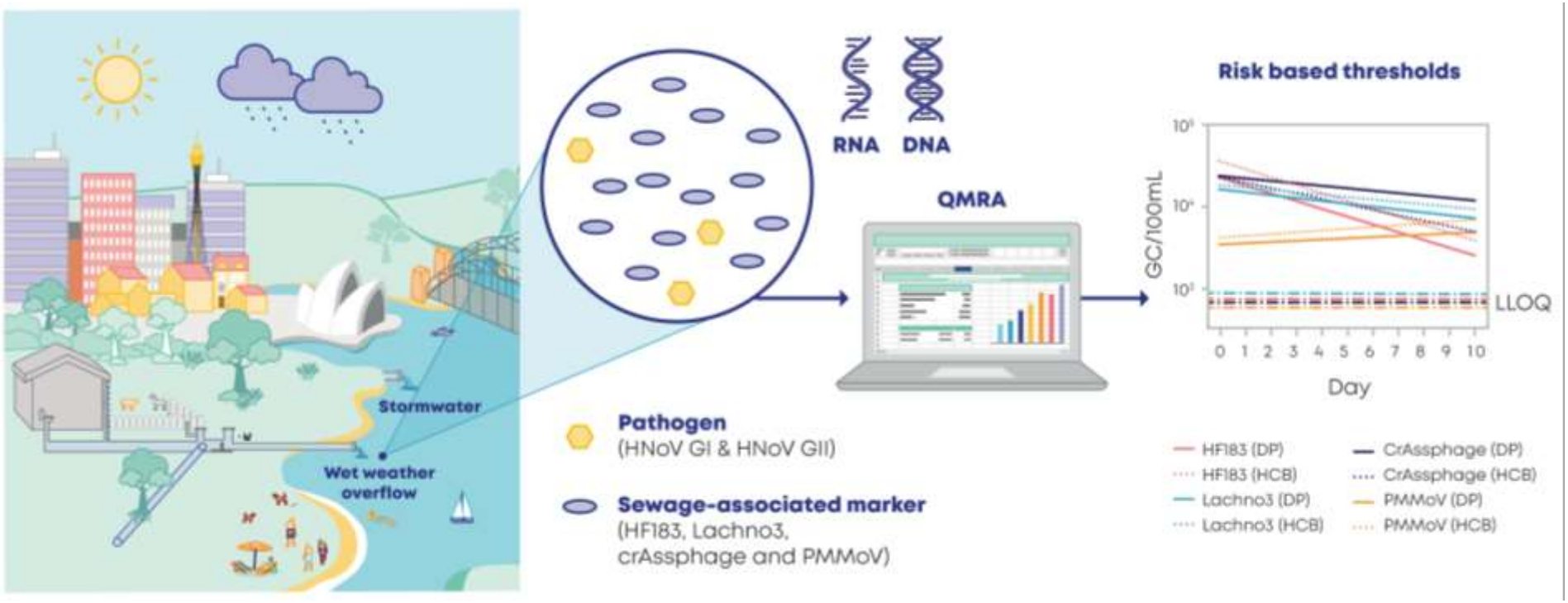


Figure 3-19: Graphical abstract illustrating the QMRA approach

3.3.1 Sydney area specific QMRA for estuarine receiving waters

The QMRA model estimates levels of health risks from ingestion of estuarine recreational waters contaminated with enteric pathogens (simulated using human norovirus as the reference pathogen) from diluted untreated sewer influent from WWOs of a known age. Gauging was installed on ERS to identify when WWOs occurred in study locations to establish a known age of spilled influent.

Several research studies have established a link between concentrations of HFMGs and the predicted risks of illness from pathogens in recreational waters, utilising QMRA. These studies include: Staley et al. (2012); Boehm et al. (2015), Ahmed et al. (2018); Symonds et al. (2016); Brown et al. (2017a); Brown et al. (2017b); Boehm et al. (2018); Crank et al. (2019); McLellan et al. (2018); Boehm and Soller (2020); and Schoen et al. (2020). Ahmed et al. (2024b) provide further discussion of these earlier studies. None of these studies incorporated site-specific HFMG/pathogen concentrations data and decay rates into their QMRA models. For example, Boehm et al. (2015) aggregated data on the concentrations of HF183 and HumM2 (another HFMG) in untreated wastewater across 54 wastewater treatment plants (WWTPs) in the United States. The HNoV concentration in wastewater adopted by Boehm et al. (2015; 2018; 2020) and Schoen et al. (2020) was from a meta-analysis of approximately 850 samples by Eftim et al. (2017).

Additionally, none of the previous RBT studies used the recent HNoV dose-response by Teunis et al. (2020), which combined all available human feeding study data and outbreak analysis in a hierarchal framework. Previous RBTs using HNoV adopted dose-response relationships fitted to a limited set of HNoV feeding study data (Teunis et al., 2008; Van Abel et al., 2017). By incorporating the outbreak data from shellfish consumption into the dose-response analysis, the updated HNoV dose-response relationship captures a range of environmental HNoV in surface waters.

Ahmed et al. (2024b) stated site-specific QMRA can capture the unique environmental conditions of a site that can affect the persistence of microorganisms, distribution of markers/pathogens in faecal sources, and the transportation of markers/pathogens from sources to the environment. Site-specific QMRA tailors understanding of microbial risks, allowing for a more effective strategy for risk management in diverse environments.

A static QMRA analysis was employed to estimate the marker concentration (that is, the RBT) in untreated sewage-contaminated estuarine environments that aligns with the USEPA benchmark of 32 illnesses per 1000 swimmers (USEPA, 2012). Four sewage-associated markers: HF183, Lachno3, crAssphage, and PMMoV were selected due to their established precision in tracking sewage contamination in environmental waters (Harwood et al., 2014; Demeter et al., 2023). These markers have also been utilised to determine the magnitude of sewage contamination in various Sydney estuary locations (Section 3.2.4, Ahmed et al., 2020a; Section 3.2.8, Ahmed et al., 2023a; Section 3.2.9, Ahmed et al., 2023b). Furthermore, the host-specificity of these markers were determined within Sydney estuarine catchments. The specificity values were 0.86, 0.94, 0.92, and 1.00 (maximum value of 1.00) for HF183, Lachno3, crAssphage, and PMMoV, respectively.

Notably, limited cross-reactivity was observed for HF183, Lachno3, and crAssphage, as detailed in the study by Ahmed et al. (2023a; Section 3.2.8).

A QMRA model that estimates a RBT assuming untreated sewage contamination of a single, known age was deemed appropriate and was adopted (Boehm et al., 2018; Schoen et al., 2020). Prevalence and concentrations of sewage-associated markers may vary geographically (Ahmed et al., 2016). Therefore, site-specific HF183, Lachno3, crAssphage and PMMoV concentrations were used. Concentrations of these markers were determined in untreated grab samples of sewage influent collected bi-weekly/monthly over a 12-month period from two WWTPs (WWTPs A and B) in urban catchments in Sydney (Section 3.2.7, Ahmed et al., 2022). The lower limits of quantification (LLOQ) for these markers were estimated from data presented by Ahmed et al., 2022 (Section 3.2.7). The period for outdoor recreational swimming in Sydney is considered to typically span from October 1 to April 30 annually. During this timeframe, concentrations of HF183, Lachno3, crAssphage, and PMMoV in wastewater from two studied WWTPs were combined and utilised to fit Weibull (HF183, crAssphage, and PMMoV) and log-normal (Lachno3 and HNoV GI + GII) statistical distributions.

In this study, HNoV was selected as the reference pathogen because it contributed most to the illness risk in previous studies (Boehm et al., 2018; Schoen et al., 2020) and is thought to be the most important etiologic agent for exposure to human-impacted recreational waters (Soller et al., 2010). HNoV GI + GII concentrations were also obtained from our previous study (Section 3.2.7, Ahmed et al., 2022). The concentrations of HNoV are generally higher in the winter season (non-swimming period) than in the summer season and may vary substantially in untreated wastewater (Nordgren et al., 2009). To account for this variation, HNoV GI + GII data obtained between October 1 to April 30 were used to fit distributions. HF183, crAssphage and PMMoV markers were modelled as a Weibull distribution using maximum likelihood estimation (MLE) within the indicated ranges to account for variable concentrations. Lachno3 and HNoV GI + GII concentrations were modelled as a log-normal distribution using MLE within the indicated ranges using R package fitdistplus (R Core Team, 2021). The empirical and theoretical cumulative distribution function (CDF) of each candidate distribution was visually compared, and the goodness-of-fit statistics, Akaike's Information Criteria, and Bayesian Information Criteria were used to select the best fit. Refer to Ahmed et al. (2024b) for the description of calculations from the scientific literature employed under the QMRA.

3.3.2 Sydney area specific risk-based thresholds (RBT) from QMRA

RBTs were estimated for HF183, Lachno3, crAssphage, and PMMoV HFMGs, aligned to the health risk benchmark of an average of 32 GI illnesses/1,000 swimmers (0.032) (USEPA, 2012). As outlined above RBTs assuming aged sewage were estimated, considering the site-specific decay rates of each marker and HNoV in estuarine receiving waters.

For the known aged sewage contamination scenarios, HF183, Lachno3 and crAssphage RBT concentrations decreased over 10 days in both locations of Hen and Chicken Bay and Davidson Park due to the relatively greater decay rates of these markers compared to reference pathogen HNoV (Figure 3-20, Figure 3-21). This decreasing trend was also found by Boehm et al. (2018; 2020) and Schoen et al. (2020) for these markers (excluding Lachno3, which was not included as a marker in the aged scenario). However, the PMMoV RBT concentrations increased over 10 days

due to slower decay rate of PMMoV in receiving waters compared to HNoV (Figure 3-20, Figure 3-21). The high stability of PMMoV in estuarine receiving waters is corroborated by our recent laboratory estuarine receiving water microcosm study, where we also reported a higher stability of the PMMoV marker compared to HF183, crAssphage, enterovirus, and HAdV (Section 3.2.10, Ahmed et al., 2023c).

The prolonged persistence of PMMoV in estuarine receiving waters prompts the question of its usefulness as a marker for fresh sewage contamination when HNoV is anticipated to decrease over time and there may be existing aged sewage contamination. This persistence and decreasing measured trend of HNoV influenced a gradual increase in calculated RBTs with time (Figure 3-20, Figure 3-21). Nevertheless, the greater persistence of PMMoV implies that the RBT concentration for the fresh sewage scenario could be applicable and can be viewed as conservatively protective of human health across various aging scenarios. Detection of PMMoV along with crAssphage may be particularly informative of sewage contamination, as under Sub-study 11 (Section 3.2.11, Ahmed et al., 2024a) the decay of crAssphage was closest to that of the three enteric viruses studied, and as such appears to be a suitable conservative surrogate for viral pathogens in temperate estuarine environment of the Sydney region.

Site specific trends of the concentrations of HF183, Lachno3, and PMMoV RBTs estimated through site-specific QMRA for DP and HCB locations were similar in the context of fresh and aged sewage contamination scenarios. However, the crAssphage RBT concentration at the HCB location decreased more rapidly compared to the DP location, primarily due to a greater difference in decay rates between crAssphage and HNoV.

To assess the utility of the RBTs from the QMRA, the concentrations of four markers in 92 storm/post-storm estuarine receiving water samples were compared to the estimated RBT concentrations. These samples were collected from ten different locations within Sydney, as detailed in our prior study (Section 3.2.8, Ahmed et al., 2023a). Of the four markers analysed, the concentrations of HF183, Lachno3, crAssphage and PMMoV exceeded the RBTs in 4.34, 1.08, 36.9 and 56.5 % of samples, respectively. Given that each water sample was analysed for all four markers simultaneously and their concentrations in untreated sewage were within an order of magnitude, the underlying reason for this observed discrepancy is not clearly understood.

A potential explanation could be the prolonged persistence of PMMoV compared to other markers in estuarine receiving waters, as suggested in previous studies (Ahmed et al., 2023c; Ahmed et al., 2024a). However, it is noteworthy that the decay rates of HF183 and Lachno3 were comparable to that of crAssphage (Ahmed et al., 2024a), and therefore, it does not account for the lower concentrations of HF183 and Lachno3 in estuarine receiving water samples compared to crAssphage. Another potential factor contributing to the observed discrepancy is the possibility that the adsorption extraction concentration method employed for the simultaneous capturing of bacterial and viral markers may not have efficiently recovered bacterial markers from estuarine receiving water samples compared to viral markers The concentration method utilised in this study was originally designed to concentrate viruses from both wastewater and environmental waters (Ahmed et al., 2015; Ahmed et al., 2020b). The adsorption extraction method involves supplementing estuarine receiving water samples with hydrochloric acid (HCl), which could potentially have adverse effects on bacterial recovery compared to viruses. Even the recovery

rates of different viruses using the same method may exhibit a two-fold difference (Ahmed et al., 2015). It is suggested that modifications to the existing adsorption extraction concentration method may be required to ensure consistent capture of both viruses and bacteria simultaneously.

Utilising the two HFMGs in the duplex assay (the bacterial HFMG HF183 and the viral HFMG crAssphage) together with PMMoV HFMG may be the most cost-effective (Section 3.2.2, Ahmed et al., 2019a) way forward under future monitoring for the broader Wet Weather Overflow Abatement program. This is recommended as the potential adsorption extraction concentration method limitation suggests Lachno3 a bacterial HFMG should be omitted from future work.

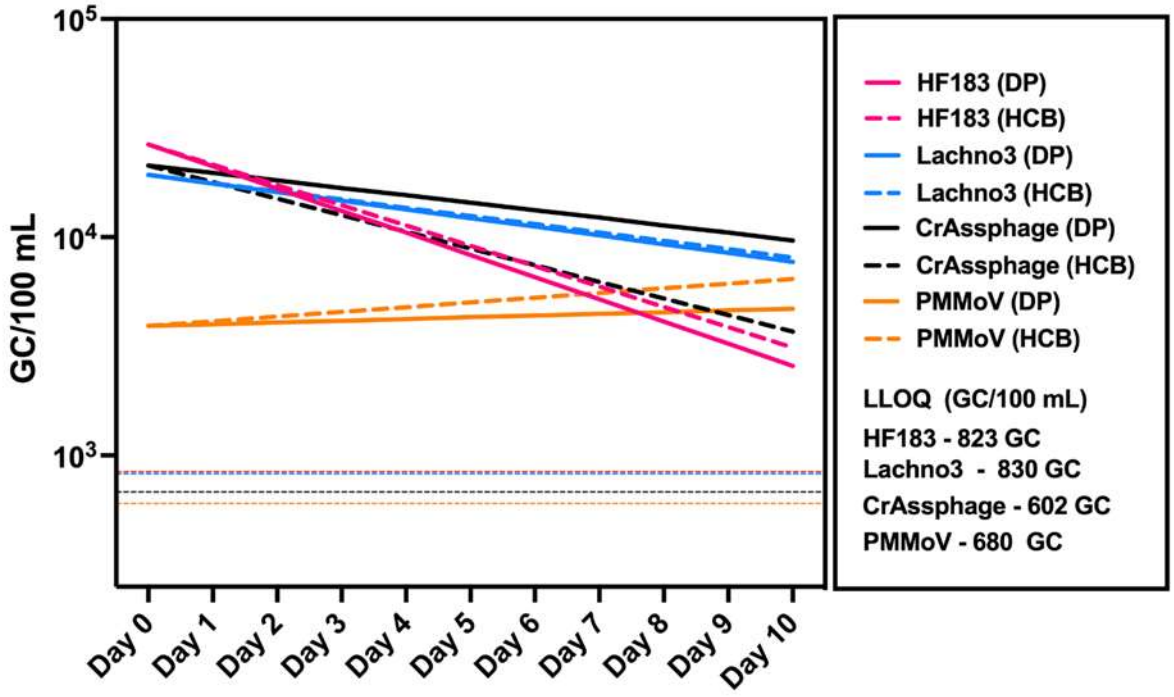


Figure 3-20: HF183, Lachno3, crAssphage and PMMoV RBT median concentrations (GC/100 mL) in Davidson Park (DP) Hen and Chicken Bay (HCB) estuarine waters corresponding to the illness benchmark of 0.032 for fresh (day 0) and aged (day 1 to day 10) sewage contamination

Dotted lines represent the lower limits of quantification (LLOQ) values.

3.3.3 Applying RBTs at other estuarine receiving water locations

Ahmed et al. (2024b) suggested that a potential limitation of site-specific approaches is that the results are not as readily generalisable to other locations, particularly locations with dissimilar characteristics. Although, in the Sydney-specific QMRA the RBT concentrations were estimated using site-specific decay rates and have the potential to be applied to other swimming locations within the similar estuarine environments within the Sydney region where WWOs are the main sources of sewage contamination in light of comparable outcomes for three of the four HFMGs. Selection of crAssphage RBTs calculated from either Hen and Chicken Bay or Davidson Park can be applied to future HFMG data from other urban estuarine receiving water locations based on the physical characteristic of turbidity from documented values of Hen and Chicken Bay or Davidson Park that only marginally varied between these two pilot study locations (Supplementary Table ST2, Ahmed et al., 2023c).

Support for application of RBTs to HFMG data from other estuarine receiving water locations of the Sydney region was illustrated by RBTs applied as outlined in Section 3.3.2 for the 92 samples collected from 10 locations after known WWOs had occurred. The collection of those samples was documented in Section 3.2.8 (Ahmed et al., 2023a,). Based upon those results, this suggests that multiple HFMGs simultaneously measured for each location could provide useful insight to differences between locations to inform a ranking of locations for input into the prioritisation methodology. Further discussion of the application of RBTs and associated monitoring to provide HFMG data is provided in Section 3.4.

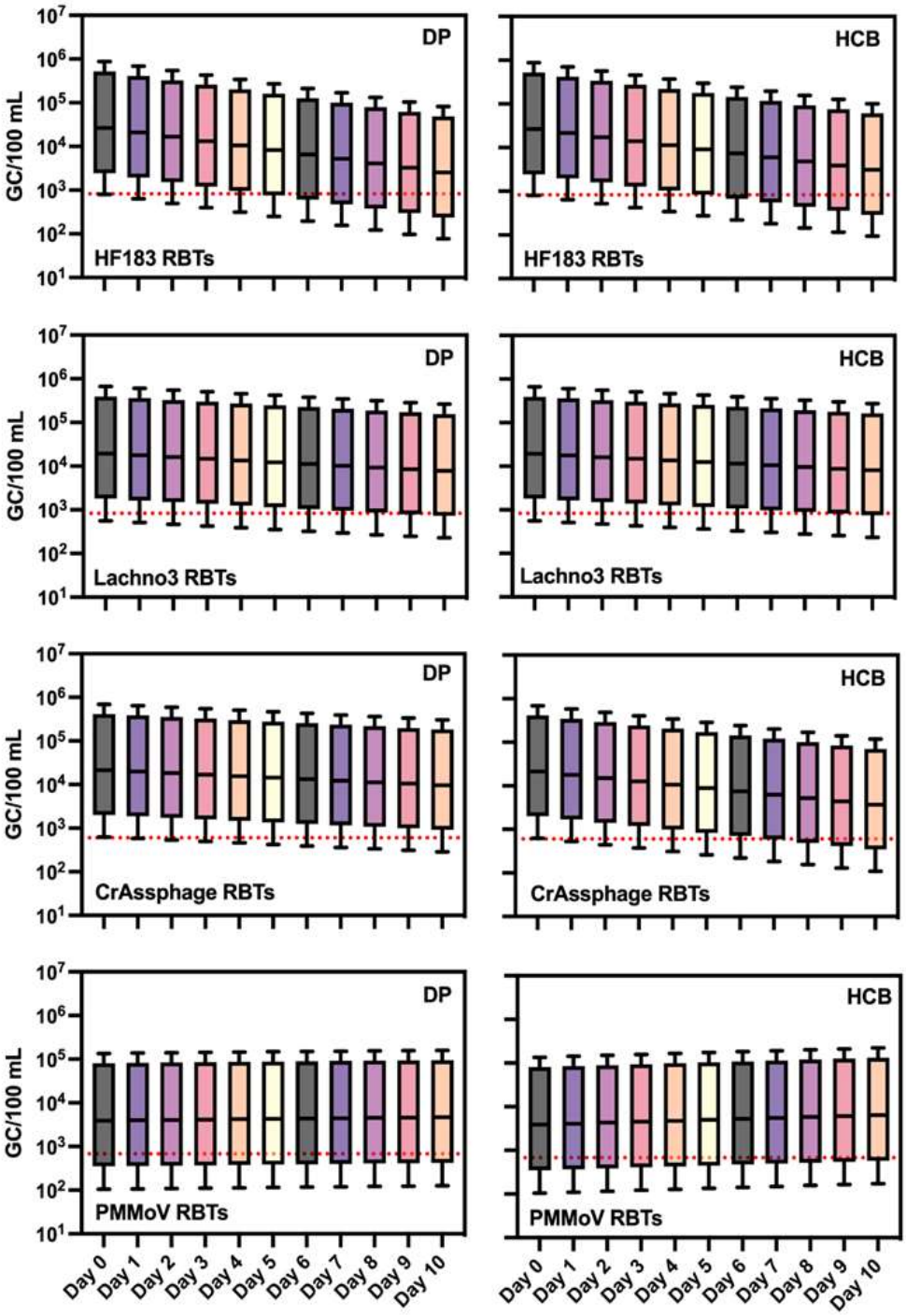


Figure 3-21: HF183, Lachno3, crAssphage and PMMoV RBT concentrations (GC/100 mL) in Davidson Park (DP) Hen and Chicken Bay (HCB) estuarine receiving waters corresponding to the illness benchmark of 0.032 for fresh (day 0) and aged (day 1 to day 10) sewage contamination

Red dotted lines represent LLOQ values. The LLOQ values were 823, 830, 602 and 680 GC/100 mL for HF183, Lachno3, crAssphage and PMMoV, respectively. The lower and upper boxes denote 25th and 75th percentiles. The lower and upper bars represent the 10th and 90th percentiles.

3.4 Application of HFMGs in future management in the wet-weather overflow abatement program

In the Sydney region, MST with HFMGs (HF183, Lachno3, crAssphage, and PMMoV) has successfully detected human faecal contamination in urban estuarine waters at 13 swimming locations after WWOs from the sanitary (separate to the stormwater system) sewerage system (Section 3.2.4 Ahmed et al., 2020a; Section 3.2.8 Ahmed et al., 2023a). In a companion sanitary inspection survey conducted after a period of dry weather that was documented by Ahmed et al. (2020a), the testing of samples with the HFMG duplex assay (Section 3.2.2 Ahmed et al., 2019a) successfully identified persistent leakage of influent that seeped from a trunk main sewer carrier into the freshwater creek that flows into the estuary at Gymea Baths. The highest HFMG concentrations were found near the carrier and lower concentrations were detected at the site situated at the baths. This outcome illustrated the utility of HFMGs to distinguish the presence of sewer influent in dry weather. These outcomes indicate that monitoring for HFMGs under both dry- and wet-weather conditions provides an additional tool to supplement information from Beachwatch assessments that are based on the much cheaper culturable faecal indicator bacteria of enterococci.

As discussed in Section 3.2.4, the charts of enterococci concentrations by rainfall category contained within State of NSW and Department of Planning and Environment (2021) reporting may represent a useful source of information. Charts with elevated enterococci concentrations under the ‘dry’ and lowest rainfall category (0.1 to 4.9 mm) (as displayed in Figure 3-9A) may represent swimming locations with similar issues as identified with HFMGs for Gymea Bay or may represent abundant animal populations. Assessing those locations with HFMGs and animal faecal-associated marker genes would allow an understanding of the faecal sources contained within these enterococci data patterns. Although the Beachwatch enterococci indicator is considerably cheaper, determining the source of faecal contamination would enable cost-effective remediation at those locations with detected human faecal contamination. Hence it would be prudent to incorporate sanitary inspection surveys as a first pass based on locations selected from State of NSW and Department of Planning and Environment (2021) reporting to identify this type of leakage issue.

Under moderate to heavy rainfall, stormwater entry into the wastewater system is 5 to 10 times more than in dry weather and exceeds the designed capacity of 3 times average dry-weather flow as outlined under Section 1.1. These conditions allow rainwater ingress to cause diluted sewer influent that exceeds dry-weather capacity to initate overflow spills from ERS into waterways. Under these conditions ERSs are a necessary component of the sanitary sewerage system to protect public health by preventing sewage backing up into homes and businesses (Bickford et al., 1999).

The current focus of capital solution planning under under the 2024 to 2030 period for the broader WWOA program is to minimise infiltration and inflow into the sewerage system across the four main coastal sewerage systems (North Head, Bondi, Malabar and Cronulla). Each sewerage system is divided into Sewer Catchment Asset Management Plans (SCAMPs), Bondi – 9; Cronulla – 11; Malabar – 74; and North Head – 61.

Remediation works focusing on source control of infiltration and inflow into the sewerage system (theoretically aimed at restoring three times dry weather flow capacity of pipes) are most likely to abate WWOs under light rainfall conditions from this restored pipe capacity together with corrected dry-weather leakage issues should reduce illness risk on some swimming days currently impaired by human faecal contamination.

To avail of learnings from raising the QMRA (Section 3.3, Ahmed et al., 2024b) and to also apply the CSIRO recommendation of using two or more HFMGs (Section 3.1), the duplex assay (crAssphage and HF183) together with PMMoV are included in the work going forward. If future continuous improvement work (as outlined in Section 3.5) is successful in defining another HFMG with better performance and acceptable RBTs, then that HFMG could be implemented in place of PMMoV.

To support the WWOA program's overarching aim to select where source control works are implemented in a cost-effective manner, the following recommendations are made for campaign monitoring:

1. **Under dry-weather conditions**, conduct surface water column monitoring with HFMGs and animal-associated faecal marker genes (for at least two avian faecal-associated marker genes) as a first pass based on locations selected from State of NSW and Department of Planning and Environment (2021) reporting that have similar data patterns to those of Gymea Bay. The assessment for avian faecal contamination is recommended as detection of avian faecal contamination was documented under both dry- and wet-weather conditions in receiving water studies of the 13 estuarine sites (Section 3.2.4, Ahmed et al., 2020a; Section 3.2.8, Ahmed et al., 2023a).
2. **Undertake surface water-column monitoring with HFMGs under a few differing magnitude storm events** for each SCAMP, ideally with:
    - 10 to 20 mm of rainfall in the preceding 24 hours
    - greater than 20 mm of rainfall in the preceding 24 hours
    - greater than 35 mm of rainfall in the preceding 72 hours for context of protracted storm events at a location

    Under each event, turbidty data must also be collected to allow comparison of crAsspahage results to RBTs calculated from DP or HCB, described in Section 3.3.3.

    The understanding gained under the WWOM of WWO spill behaviour suggests these rainfall amounts represent conditions to base prioritising locations for source control that is most likely to provide information toward restoration of the designed dry-weather capacity of the sewerage system
3. Once HFMG concentrations have been returned from laboratory testing, those concentrations are compared with the Sydney specific RBTs established through the site-specific QMRA modelling conducted under the WWOM project
4. **Rank swimming locations by numerically ordering greater concentration departures from RBTs** to least outcomes from concentration departures from RBTs for each

respective HFMG assessed. Another consideration in ranking locations accounts for Dr Ahmed's advocation, for example, if three assessed HFMGs return RBTs at a concentration of concern, then that location would be prioritised over a location where only two, one or no HFMG returned RBTs at a level of concern

5. in some cases, a finer dissection of locations is required and **further assessment under sanitary inspection surveys with HFMGs** may assist in establishing problem ERSs under low magnitude rainfall events could be undertaken

These numerically ordered outcomes can then be inputted into Sydney Water's risk prioritisation methodology to inform capital solution planning to provide a most cost-effective restoration of the designed dry weather capacity of the sewerage system.

Working with existing human resourcing to conduct storm event sampling

The advantage of applying the QMRA calculated RBTs by Day 0 to Day 10 is that they provide flexibility to monitor across a range of days post storm event. This advantage helps overcome the limited human resources available for reactive event monitoring that can be completed in a workday by conducting monitoring within the 11 days window after a wet-weather event. Competing regulatory programs also have reactive wet-weather monitoring to be performed, such as keeping the drinking water system safe. The inherent flexibility of RBTs allows results collected from this 11-day window to be compared to the appropriate RBT day threshold.

It should be noted that this approach is only appropriate when rainfall has not occurred within those 11 days. If rainfall did occur within the 11 days post a storm event, then comparison to RBTs would be limited across the days up to the next rainfall event to provide a conservative assessment.

Continuous improvement obligation

As outlined in Section 3.3.2 the potential absorption extraction concentration method limitation suggests Lachno3 (a bacterial HFMG) should be omitted from future work as the more cost-effective duplex assay developed under sub-study 2 (Section 3.2.2, Ahmed et al., 2019a) provides results for the bacterial HF183 HFMG and the crAssphage viral HFMG. The combination of the duplex assay and PMMoV (a viral HFMG) may be the most cost-effective way forward under future monitoring for the broader WWOA program. The prolonged persistence of PMMoV in estuarine receiving waters prompts raises the question of its usefulness as a marker for fresh sewage contamination. After considering outcomes of the QMRA, the greater persistence of PMMoV may imply that the RBT concentration for the fresh sewage scenario could be applicable and conservatively protective of human health across various aging scenarios.

Additional investment in applied research to assess the suitability of other novel emerging marker assays is recommended. This investment would conform to the pollution study (PS307) objective of continuous improvement to the prioritisation methodology. Using extracted DNA that have been stored (frozen at -80^{o}C) from pilot study samples to develop HFMG aliquots, will provide a cost-effective capacity to assess potential marker assays as discussed in Section 3.5 below.

3.5 Refinements to the existing MST toolbox for monitoring surface waters

Reasoning informing recommendations for refinement of MST toolbox

The collaborative work between CSIRO and Sydney Water (under the WWOM project) aimed to understand the utility of the MST toolbox for monitoring surface waters to determine faecal sources (human or animal) and to guide sound capital solution planning decisions for future management of the wastewater network.

In the initial stages of the WWOM project, CSIRO advised the use of two or more HFMGs assays in confirming a human faecal source. The WWOM project managers commended CSIROs investigation of three HFMGs which included a bacterial (*Bacteroides* HF183) and two viral marker genes cross-assembly phage (crAssphage)] and pepper mild mottle virus (PMMoV). The bacteriophage HF183 is an obligate anaerobe that is highly abundant in the human gut. CrAssphage is a highly abundant DNA virus in the human gut and is an integral part of the normal human gut virome. While PMMoV is an RNA virus that infects peppers (capsicums) and its presence in human faeces originates from the consumption of infected peppers and or in other food products that contain spices with peppers as an ingredient.

In 2016, PMMoV was one of the few viral marker genes available for developing MST, and crAssphage was then a novel MST marker gene. Studies of PMMoV were advancing, and by 2018 PMMoV was considered a promising surrogate for infectious enteric viruses (Symonds et al., 2018). The status of crAssphage as a surrogate for enteric viruses was yet to be established. During 2016-2018, the bacterial Lachno3 human faecal-associated MST marker gene assay became available, and it was subsequently included into studies of the WWOM project.

CSIRO assessed the performance of these four human faecal-associated MST marker genes, HF183, crAssphage, Lachno3 and PMMoV. That assessment was based upon host sensitivity (using three criteria) which was unable to separate the four MST marker genes, as an absolute value host sensitivity was returned for all marker genes (Ahmed et al., 2023a). In contrast, separation of these marker genes was provided by assessment of host specificity (based upon three criteria), with an absolute value of 1.0 for the human faecal-associated MST marker gene of PMMoV, values of >0.9 for crAssphage and Lachno3, and a value of 0.76 for HF183 (Ahmed et al., 2023a). The study outcome affirmed advice to employ two or more human faecal-associated MST marker genes to confirm a human faecal source.

In further assessment, results of recent reduction studies (Ahmed et al., 2023b, 2023c) suggest PMMoV may represent an overestimation of potential human health risk in receiving waters. This is supported by Gyawali et al. (2021) who determined that PMMoV significantly overestimated the norovirus contamination in shellfish. These findings suggest viral PMMoV HFMG is a less suitable surrogate for enteric viral pathogens in estuarine waters as discussed above. PMMoV is the HFMG most likely to be replaced from outcomes of future continuous improvement studies conducted in the Sydney region. As outlined above, the bacterial Lachno3 HFMG was not recommended to be taken forward into future monitoring.

Promoting the duplex assay HFMGs of HF183 and crAssphage, which are bacterial- and DNA viral-based markers, respectively, and replacing PMMoV an RNA viral-based marker, with another

bacterial- and DNA viral-based marker, may present a potential weakness. Research on this aspect is currently lacking in the literature, and therefore it is yet to be determined if this is a weakness or not.

After a further eight-years (2016-2024), some additional novel MST marker genes have emerged for use in the MST toolbox, and we describe several here:

- the **tomato brown rugose fruit virus (ToBRFV) Mo gene** is a plant RNA virus (Natarajan et al., submitted). Initial testing (outside of Australia) suggests the ToBRFV Mo gene is more prevalent and abundant than the plant based PMMoV RNA virus and matches prevalence of crAssphage. Further assessment of ToBRFV Mo marker gene as a replacement for PMMoV seems warranted to determine if it has better applicability as a pathogen surrogate to reflect risk of RNA infectious enteric viruses such as enterovirus and norovirus in receiving waters.
- **bacteriophage Bifidobacterium (Bifi)**. Initial testing of a Bifi qPCR assay has demonstrated 100% host-specificity against non-human samples and was detected in all untreated and treated wastewater samples examined by Shenzhen and Brisbane (Li et al., 2022). Concentrations of Bifi in wastewater have been documented to be lower than that of HF183 and crAssphage CPQ_056. The initial host-specificity (100% against untreated wastewater samples) results suggest the Bifi may be another promising human-specific marker gene, and as such further assessment is warranted to understand receiving water concentrations and potential utility to detect WWOs. Specifically, there is a need to determine if the reported lower wastewater concentrations are a limiting factor on detection in receiving waters. The concentration of the Bifi marker is similar to enteric viruses in water and therefore could potentially act as a better surrogate for enteric viruses. Answering these questions is important to determining the usefulness of this potential marker.
- a **culturable bacteriophage GB-124** is being developed by the USEPA into a qPCR assay (personal communication between Dr. Orin Shanks, USEPA, and Dr. Ahmed, CSIRO). Such an assay has the potential to considerably improve risk assessments, as the qPCR assay and culture-based method will provide information on the viability of GB-124 in environmental samples. This information can be extrapolated to determine the fraction of viable norovirus in estuarine waters. Collaborative testing of the GB-124 novel marker gene may also be informative and help align to a potential future recognised USEPA method. As part of the collaborative testing, a study to evaluate the viability of GB-124 from the culturable method to the qPCR HFMG assays may contribute some understanding to the consideration of the viability of pathogenic viruses from analyses based on qPCR.

Ahmed et al. (2023) has recommended further research to better understand the dynamics of enteric viruses in environmental waters, and to develop more sensitive detection methods that can accurately quantify viral loads in samples with varying levels of dilution such as in receiving waters. Evaluating the above emergent human faecal-associated MST marker genes would be a prudent continuous improvement step for the broader WWOA program. In particular, the RNA virus ToBRFV Mo human associated marker gene may provide a suitable replacement for PMMoV. The

availability of stored DNA and RNA from the WWOM project, that was the basis of previous publications, affords the opportunity to explore these emergent novel human faecal-associated MST genes. These new marker genes will be evaluated, and based on the performance, may be included in the existing toolbox to further strengthen our WWOA program.

Recommendations for refining the MST toolbox

It is recommended that ToBFRV Mo, Bifi and GB-124 human faecal-associated marker genes are assessed for prevalence, abundance and host specificity. Samples already collected (from the WWOM project) and stored DNA and RNA will allow comparisons back to the datasets of the previous WWOM collaborative paper outcomes, along with developing RBTs for the most promising marker from this work. That comparative information may enable selection of an additional marker gene for inclusion in future monitoring to maintain three HFMGs (potentially developed as a triplex assay). Maintaining three HFMGs takes into account Dr Ahmed's advocation, for example, if three assessed HFMGs return RBTs at a concentration of concern, then that location would be ranked higher than a location where only two, one or no HFMG returned RBTs at a level of concern. The decay rates of the newly proposed markers will also need to be evaluated. This should significantly improve the quality of information used as the basis of management decisions in capital solution planning given potential cost implications of such work. Development of a triplex assay would create a most cost-effective monitoring tool.

As discussed in Section 3.3.2, modifications to the existing adsorption extraction concentration method should be explored to see if more consistent capture of both viruses and bacteria can be obtained simultaneously.

Lastly, as both the CSIRO and Sydney Water laboratories have since acquired dPCR platforms and opportunity now exists, as suggested in sub-study 6 (Ahmed et al., 2020c), to investigate the accuracy and precision of sewage-associated marker gene assays by analysing not only environmental waters but also water samples seeded with sewage in a double-blind study.

References

Ahmed, W., Bivins, A., Payyappat, S., Cassidy, M., Harrison, N., Besley, C., 2022. Distribution of human fecal marker genes and their association with pathogenic viruses in untreated wastewater determined using quantitative PCR. Water Res. 226, 119093, https://doi.org/10.1016/j.watres.2022.119093

Ahmed, W., Payyappat, S., Cassidy, M., Harrison, N., Besley, C., 2023a. Microbial source tracking of untreated human wastewater and animal scats in urbanized estuarine waters. Sci. Total Environ. 877, 162764. https://doi.org/10.1016/j.scitotenv.2023.162764

Ahmed, W., Payyappat, S., Cassidy, M., Harrison, N., Besley, C., 2023b. Reduction of human fecal markers and enteric viruses in Sydney estuarine waters receiving wet weather overflows. Sci. Total Environ. 896, 165008. https://doi.org/10.1016/j.scitotenv.2023.165008

Ahmed, W., Gabrewold, M., Payyappat, S., Cassidy, M., Harrison, N., Besley, C. 2024a. Assessing the nucleic acid persistence of human wastewater markers genes and enteric viruses in estuarine waters in Sydney, Australia. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2024.171389

Gyawali, P., Devane, M., Scholes, P., Hewitt, J. 2021., Application of crAssphage, F-RNA phage and pepper mild mottle virus as indicators of human faecal and norovirus contamination in shellfish. Sci. Total Environ. 783, 146848 https://doi.org/10.1016/j.scitotenv.2021.146848

Li, X., Ahmed, W., Wu, Z., Xia, Y., 2022. Developing a novel Bifidobacterium phage quantitative polymerase chain reaction-based assay for tracking untreated wastewater. Sci. Total Environ. 838, 155815. https://doi.org/10.1016/j.scitotenv.2022.155815

Natarajan, A., Fremin, B.J., Schmidtke, D.T., Wolfe, M.K., Zlitni, S., Graham, K.E., Brooks, E.F., Severyn, C.J., Sakamoto, K.M., Lacayo, N.J., Kuersten, S., Koble, J., Caves, G., Kaplan, I., Singh, U., Jagannathan, P., Rezvani, A.R., Bhatt, A.S., Boehm A.B., (2023) The tomato brown rugose fruit virus Movement Protein gene is a novel microbial source tracking marker. Appl. Environ. Microbiol. 89:e00583-23. https://doi.org/10.1128/aem.00583-23

Symonds, E.M., Hguyen, K.H., Harwood, V.J., Breitbart, M., 2018. Pepper mild mottle virus: A plant pathogen with a greater purpose in (waste)water treatment development and public health management. Water Res. 144, 1-12. https://doi.org/10.1016/j.watres.2018.06.066

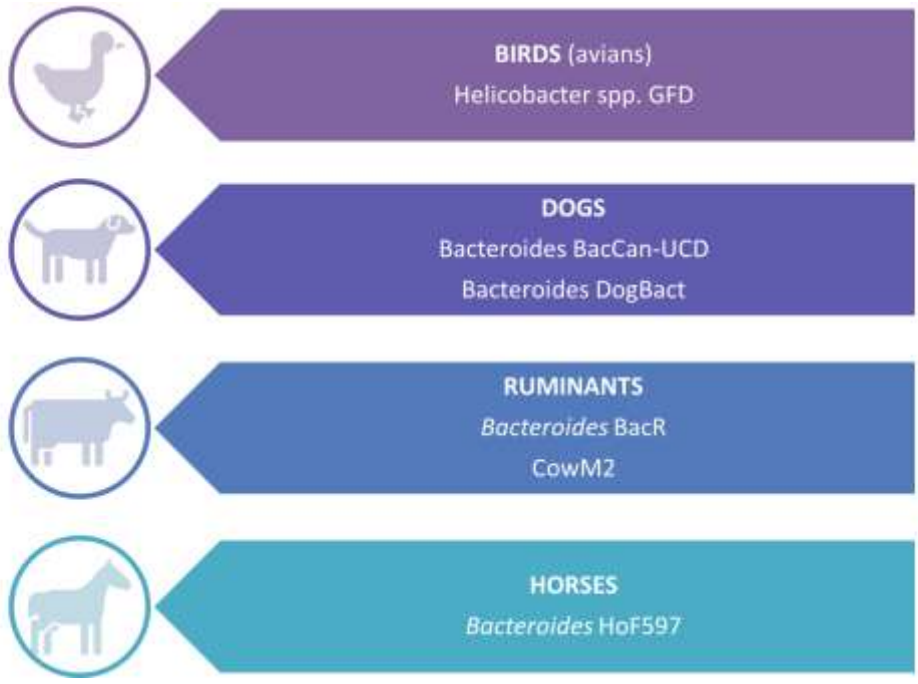
3.6 Expansion of the MST toolbox with animal marker gene assays

Under the WWOM, CSIRO conducted analysis of animal scat-associated marker genes, as Sydney Water did not yet have this capacity. The utility of animal scat-associated marker genes in the above sub-studies 3, 4, and 8 (Sections 3.2.3 Ahmed et al., 2019a; Section 3.2.4 Ahmed et al., 2020a; and Section 3.2.8 Ahmed et al., 2023a) successfully demonstrated the ability to identify animal faecal sources in receiving waters.

Developing in-house capability with animal scat-associated marker genes would provide molecular tools that can be employed under both dry- and wet-weather assessments of faecal sources in receiving waters. These tools under dry weather sewer overflow investigations would be most valuable in post remediation work situations where high enterococci concentrations are detected. Under those circumstances water samples would be analysed for both HFMGs and appropriate animal scat-associated marker genes to understand sources of faecal contamination. When animal faecal sources are implicated in the absence of human faecal sources then this post remediation assessment would categorise repair work as successful, and human resourcing could then be deployed elsewhere to correct other network issues.

3.6.1 Testing of animal scat-associated marker genes under WWOM

Under sub-study 8 (Section 3.2.8, Ahmed et al., 2023a) host specificity and host sensitivity was assessed for the following six animal scat-associated marker genes:



Host sensitivity

Among the animal scat-associated marker genes, HoF597 (horse) marker gene showed absolute host sensitivity (1.00), while host sensitivity of the BacR (ruminant) marker gene was high (0.94) and CowM2 (ruminant) was slightly lower (0.82) with the same values returned for all three test-criteria (Ahmed et al., 2023a).

Variations in values across the three test criteria occurred for the other three animal scat-associated marker genes. BacCan-UCD (dog) showed absolute host sensitivity using PCR detection results, however, the values were less for qPCR/g scat (0.82) and/ng nucleic acid (0.75). The sensitivity values of DogBact (dog) marker genes were 0.81 using both PCR detection data and qPCR data/g of scat. However, the value reduced to 0.75 when assessed with qPCR data/ng criterion. The host sensitivity value of the GFD (bird) marker gene was greater (0.86) for PCR detection compared to qPCR/g scat (0.76) and qPCR/ng of DNA (0.53). Nonetheless, with the exception of one criterion for the GFD marker, these animal-scat associated marker genes showed host sensitivity values >0.75 which can be considered excellent (Ahmed et al., 2023a).

Host specificity

BacR, and CowM2 (ruminant) marker genes showed absolute host-specificity (1.00) for all three criteria used to calculate host specificity. While the HoF597 (horse) marker gene showed absolute host specificity (1.00) for qPCR/g scat and /ng of nucleic acid, but the specificity value was slightly lower for PCR detection data (0.98) (Ahmed et al., 2023a). These high host specificity values suggest these marker genes are suitable to assess ruminant (including cows) and horse faecal pollution.

In contrast DogBact (dog) and GFD (bird) marker genes showed lower host specificity and suggest they should be synergistically used with other marker genes from dog and avian hosts (Ahmed et al., 2023a).

The following outlines the returned specificity values across the three test criteria:

- The host specificity of BacCan-UCD (dog) was the lowest (0.23) using PCR detection criteria, but the values were greater for qPCR/g of scat (0.30) and /ng of DNA (0.58). Based on the high prevalence of the BacCan-UCD markers in other animal-scat samples, Ahmed et al. (2023a) did not recommend using the BacCan-UCD marker alone for tracking dog faecal pollution in Sydney region.
- DogBact (dog) host specificity values also showed similar patterns like BacCan-UCD, but the values were greater (ranging from 0.60 to 0.90).
- The host specificity value of GFD (bird) marker gene was greater for qPCR/ng of DNA (0.84) compared to other two criteria (0.74 using PCR detection and 0.81 for qPCR/sample mass or volume).

Based upon the specificity findings of Ahmed et al. (2023a) the exploration of other animal scat-associated marker genes for confirming dog and avian faceal contamination is warranted.

3.6.2 Other potential animal scat-associated marker gene assays

Zlender and Rupnik (2023) reviewed 55 animal scat-associated marker gene assays published between 2005 and 2017. They suggest due to geographical instability, diagnostic measures of a marker should always be validated before its application in a new geographical area. Our abovementioned study has undertaken this testing for six animal scat-associated marker genes outlined above.

In addition to bacterial animal markers, Dr Ahmed of CSIRO has also recommended testing of viruses linked to animals such as pigs (porcine adenovirus), cattle (bovine adenovirus, polyomavirus), and chicken (chicken parvovirus), as viruses are recognised for their greater host specificity compared to bacterial markers.

Birds

Zlender and Rupnik (2023) stated that birds are known to carry human pathogens that are excreted with faecal waste and include enteric bacteria (Salmonella, E. coli, and Campylobacter), protozoans (Cryptosporidium, Giardia) and microsporidia (Enterocytozoon, Encephalitozoon) (Vlahović et al., 2004; Graczyk et al., 2008). There are two main potential origins of bird faecal pollution: poultry farms and wild birds. While wild bird droppings appear to harbour less abundant and fewer pathogenic bacteria than poultry their contribution to faecal contamination of receiving waters should not be neglected as they are fundamental components of the aquatic ecosystem (Benskin et al., 2009; Boukerb et al., 2021). Although specific context for the Sydney region is absent in this literature.

The presence of the GFD avian marker gene in the estuarine water column under both dry- and wet-weather conditions was observed across our three receiving water sub-studies (Sections: 3.2.3, 3.2.4 and 3.2.8) and also in sediment where the GFD avian marker gene was widely detected in 63.8% of samples collected from seven disparate estuarine locations across three estuaries of the Sydney region (Section 3.2.5 Ahmed et al 2020b). A recent post remediation assessment observed waterfowl in a turbid wetland that had high measured enterococci concentrations along with detection of HFMGs. This clearly suggests that developing an in-house capability to assess avian faecal contamination is a necessary component of the microbial source tracking toolbox to help better interpret sources and contributions of faecal contamination.

Zlender and Rupnik (2023) found five assays for tracking faecal contamination originating from birds in general. All of them target the 16S rRNA gene of different bacteria. Given that birds are the most diverse land vertebrates and can be endemic to certain geographic locations (Chiappe, 2009), the selection of broadly specific markers and detecting bird faeces in general can be difficult. The primary challenge lies in the sensitivity of assays, which may further decrease by validating on a broader range of wild bird species. GFD and Av4143 were the most frequently validated assays for identifying faecal contamination originating from birds and showed the highest performance when compared to the remaining three markers (GFB, GHC, Av163F) (Green et al., 2012; Ohad et al., 2016). However, their performance varied greatly among different validation studies (Ohad et al., 2016; Symonds et al., 2017; Vadde et al., 2019; Schiaffino et al., 2020; Zhang

et al., 2020; Rytkönen et al., 2021) and within-study comparisons show contradictory results on which assay results in highest specificity and sensitivity (Vadde et al., 2019; Rytkönen et al., 2021).

Zlender and Rupnik (2023) outlined that some assays were designed to detect gull faeces. Among them Gull2 and Gull4 were the most frequently validated. According to the results of Ryu et al. (2012), Gull4 assay tends to be more specific and less sensitive than Gull2. GHC and GFB assays for the detection of gull faeces were validated when first published by Green et al. (2012). The specificity of these assays was very high, however they detected only 64 and 26% of gull faeces, respectively.

For detecting faecal contamination of waterfowl, Ohad et al. (2016) developed three assays with relatively low sensitivity but high specificity: Av13, Av24 and Av216. Using the comparative analysis of the 16S rRNA gene, one swan-associated (Swan_2) marker was developed and resulted in sensitivity of 75% and specificity of 90% when tested on faecal samples (Boukerb et al., 2021). They go on to state, among assays developed for the detection of faecal contamination originating from domestic birds, one assay aims to detect poultry feces in general (LA35) (Weidhaas et al., 2010; Weidhaas and Lipscomb, 2013; Schiaffino et al., 2020). Another more recently developed assay (Av43) detects chicken feces with a high degree of specificity (Ohad et al., 2016), whereas two assays based on mtDNA (ND5, CytB) aim to detect both chicken and duck feces (Zhuang et al., 2017; Schiaffino et al., 2020).

Dog

Zlender and Rupnik (2023) described dogs as an important part of the urban environment. They indicated that the BacCan-UCD marker was the most validated with host sensitivity always above 75% while host specificity ranged from 47 to 100% in different studies. They suggested other unvalidated dog marker gene assays including DF113F-DF472R, DF53F-DF606R and DF53F-DF606R targeting the 16S rRNA gene of Bacteroidales (Hussein et al., 2014) and the DogBac assay that we have since tested as outlined above. Future testing in the Sydney region of these Hussein et al. (2014) dog-marker genes may help identify a suitable second dog marker gene to pair with the DogBac marker gene to implement the Ahmed et al. (2023a) recommendation for use of more than one marker gene to confirm dog faecal detection.

Pig

Of the pig-associated faecal marker genes reviewed by Zlender and Rupnik (2023), they noted that Pig-2-Bac was the most validated and proved superior to other assays (Pig-Bac1, Pig-Bac2, *L. amylovorus*, P-CytB and P-ND5) in multiple studies in terms of sensitivity and specificity.

Native animals

For native animals a possum assay has been developed and may be also worth considering as an addition to the microbial source tracking toolbox.

The above text outlining other potential animal faecal associated marker gene assays is drawn from the following review paper by Zlender, T., Rupnik, M., 2023. An overview of molecular markers for identification of non-human fecal pollution sources. Front. Microbiol. 14:1256174. https://doi.org/10.3389/fmicb.2023.1256174

3.6.3 Development of an in-house capability across ten animal marker gene assays

Extensive experience gained from post remediation work, where enterococci concentrations have remained elevated and animal faecal contamination was suspected, guided the following recommendations for adoption of animal scat-associated marker gene assay targets:

Horse

The ability to assess potential horse faecal contamination with the HoF597 marker gene may be of benefit in peri-urban areas of greater Sydney including along the Hawkesbury-Nepean River and near horse-event facilities such as the Centennial Park area and adjacent to racecourses of Sydney. Similarly, the availability of the ruminant marker BacR may also be of benefit in peri-urban areas, although for cows the specific assay CowM2 may also be advantageous. As noted above these three assays have already undergone Sydney specific testing and performed very well being recommended for assessment of these faecal sources as simplex assays. The bovine adenovirus and polyomavirus could also be assessed for the Sydney region as suggested by Dr Ahmed.

Pig

A Pig-2-Bac marker gene assay may be also advantageous to have available for testing in peri-urban areas of greater Sydney near pig farms. However, Sydney area specific testing will need to be performed (as conducted by Ahmed et al., 2023a) to confirm the reported good performance outlined by Zlender and Rupnik (2023) review of animal marker gene assays. Of the viruses recommended by Dr Ahmed, the porcine adenovirus could also be assessed for the Sydney region.

DogBact and GFD marker gene assays showed lower host specificity, these markers should be synergistically used with other marker genes from dog and avian hosts. Hence, the DogBact marker gene assay should be paired with another dog specific assay as recommended by Ahmed et al. (2023a).

Dog

The following three assays described above by Zlender and Rupnik (2023) DF113F-DF472R, DF53F-DF606R and DF53F-DF606R should be tested with the same methods as under Ahmed et al. (2023a) to find a replacement assay for the poor host-specificity performance of BacCan-UCD in the Sydney region.

Birds

Alternative bird marker gene assays outlined by Zlender and Rupnik (2023) include the generic Av4143, the waterfowl specific Av13, Av24, Av216 or poultry generic LA35 or chicken specific Av34. Of the viruses recommended by Dr Ahmed the chicken parvovirus could also be assessed for the Sydney region. Adoption of some of these assays to pair with GFD marker will be subject to outcomes of suitable host sensitivity and specificity from Sydney specific testing as outlined by Ahmed et al. (2023a).

3.6.4 Development of a multiplex animal scat-associated marker gene assay

Development of a multiplex animal scat-associated marker gene assay would be desirable from a cost-efficiency management-lens. This approach is expected to mirror that of the HFMG duplex assay developed under sub-study 2 (Ahmed et al., 2019a). Candidate assays for inclusion in the multiplex assay should be selected based on outcomes of Sydney region specific testing and potential compatibility.

As outlined above, the availability of in-house capability with animal marker genes would help determine situations when animal faecal sources are implicated in the absence of human faecal sources. Then post remediation assessment would categorise repair work as successful, and human resourcing could then be deployed elsewhere to correct other network issues.

Recommendations for animal scat-associated marker gene assays

It is recommended the 10-animal scat-associated marker gene assays be implemented into the microbial source tracking toolbox, with assay inclusion subject to outcomes of Sydney region specific testing that is yet to be completed for some of the assays.

- ruminants
  - Bacteroides BacR
  - CowM2
  - bovine adenovirus, subject to outcomes of Sydney-specific testing
  - polyomavirus subject to outcomes of Sydney-specific testing
- horses
  - Bacteroides HoF597
- pigs
  - PIG-2-Bac, subject to outcomes of Sydney-specific testing
  - Porcine adenovirus, subject to outcomes of Sydney-specific testing
- dogs
  - Bacteroides DogBact
  - plus another dog marker assay from amongst DF113F-DF472R, DF53F-DF606R and DF53F-DF606R. Assay selection subject to outcomes of Sydney-specific testing
- birds (avian)
  - *Helicobacter spp.* GFD generic bird assay
  - plus the Av4143 generic bird assay subject to outcomes of Sydney-specific testing
- waterfowl
  - Av13, Av24, Av216, selection of one of these assays subject to outcomes of Sydney-specific testing
- poultry
  - LA35 poultry in general, Av34 chicken specific, inclusion of one of these two assays subject to outcomes of Sydney-specific testing
  - Chicken parvovirus, subject to outcomes of Sydney-specific testing

Development of a multiplex animal scat-associated marker gene assay is important for cost-efficiency. This approach should mirror that of the HFMG duplex assay developed under sub-study 2 (Ahmed et al., 2019a). The availability of in-house capability with animal marker genes would help determine situations when animal faecal sources are implicated in the absence of human faecal sources. Then post remediation assessment would categorise repair work as successful, and human resourcing could then be deployed elsewhere to correct other network issues.

4 Contaminants that may pose a risk of adverse ecological effect

A key aim of this part of the WWOM study is the identification of contaminants of concern in WWOs and evaluation of their potential to adversely impact biota.

Ammonia has been previously identified together with chlorine as the most important toxicants immediately downstream of effluent (treated-sewage) discharges (Davis, 1997). Environment Canada (2001) also noted that unionised ammonia was the most frequent cause of toxicity from wastewater effluent. Camargo and Alonso (2006) suggested that ammonia, nitrite, and nitrate can contribute to direct toxicity of aquatic organisms, and Quijano et al. (2017) stated that combined sewer overflows are a major source of carbonaceous biochemical oxygen demand and ammonia.

Metals in influent are also contaminants of concern. Drozdova et al. (2015) suggested that the occurrence of metals in influent depends on the number of connected inhabitants and on the human activities from industries and households. A literature review of sources of metal contaminants in domestic wastewater (untreated sewage influent) from household studies in Australia indicated that major inputs were from the metals lead, zinc, and copper, with arsenic, nickel, and mercury near detection limits. Inputs of lead appeared to originate from the laundry and bathroom, while zinc mainly originated from the bathroom, and the major sources of copper were from plumbing and water supply (Tjadraatmadja and Diaper, 2006). The above citations reflect that ammonia, metals and biochemical oxygen demand should be evaluated in sewer influent flowing from properties into the Sydney Water wastewater system.

In evaluations undertaken under the WWOM, the pilot studies aimed to establish whether chemicals are contaminants or pollutants. Chapman (2007) states contamination is simply the presence of a substance where it should not be or at concentrations above background and goes on to further state that pollution is contamination that results in or can result in adverse biological effects to resident communities. All pollutants are contaminants, but not all contaminants are pollutants.

References

Chapman, P.M. 2007. Determining when contamination is pollution – Weight of evidence determinations for sediments and effluents. Environ. Int. 33, 492-501.

Camargo, J.A., Alonso, Á. 2006. Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: a global assessment. Environ. Int. 32, 831-849.

Davis, J.R. 1997. Revitalization of a northcentral Texas river, as indicated by benthic macroinvertebrate communities. Hydrobiologia. 346, 95-117.

Drozdova, J., Raclavska, H., Skrobankova, H. 2015. A survey of heavy metals in municipal wastewater in combined sewer systems during wet and dry weather periods. Urban Water J. 12, 131-144.

Environment Canada. 2001. The state of municipal wastewater effluents in Canada / Issued by the Indicators and Assessment Office. https://publications.gc.ca/site/eng/104183/publication.html. (Accessed 21 Dec 2023).

Quijano, J.C., Zhu, Z., Morales, V., Landry, B.J., Garcia, M.H. 2017. Three-dimensional model to capture the fate and transport of combined sewer overflow discharges: A case study in the Chicago Area Waterway System. Sci. Total Environ. 576, 362-373.

Tjadraatmadja, G. Diaper, C. 2006. Sources of critical contaminants in domestic wastewater – a literature review.CSIRO: Water for a Healthy Country National Research Flagship Report, Canberra, ACT. https://doi.org/10.4225/08/59b9805b74db1

4.1 Previous work identifying contaminants of concern in the Sydney region

A risk assessment of chemicals in wet-weather discharges from wastewater treatment plants and spills from 11 major ERSs of the Sydney region was undertaken by Bickford et al. (1999), with 114 chemicals of potential concern assessed. Results of laboratory toxicity testing conducted as part of that assessment indicated that influent collected from sewer carrier pipes under WWO conditions was toxic, while no conclusions could be drawn on toxicity in downstream receiving waters (Bickford et al., 1999). This risk assessment documented ammonia and nitrate as contaminants of concern in WWOs. Another 12 contaminants of concern were identified by this work, with 85% of the load of those chemicals contained in stormwater (Bickford et al., 1999), which suggests secondary contributions from WWOs. Among these 12 chemicals were five organochlorine chemicals, three metals (copper, silver and zinc), two organophosphates that are broad-spectrum insecticides (diazinon and chlorpyrifos), along with one metabolite of a chlorophenoxy compound that is a broad leaf herbicide (2,4-dichlorophenol).

The deregistration of organochlorine chemicals for general use in Australia in the mid-late 1980s (Dept. of Environment, 1997), with remaining uses banned in 1992 (Radcliffe, 2002), has likely seen their diminished (if not ceased) disposal through the sewerage system as stockpiled supplies were used up. This diminished use is supported by the below detection limit results for 18 organochlorine chemicals reported by Besley and Birch (2019) from that assessment of sediment near two deepwater ocean outfalls off Sydney from 2002 to 2016. These two outfalls are the end point discharges of primary-treated effluent from the North Head and Malabar sewer catchments, and these catchments are two of the four main sewer catchments under study in the WWOM. As such, this suggests that organochlorine chemicals are highly unlikely to have posed a risk to biota through the 2016 to 2023 study period of the WWOM. Hence, these were not assessed under the WWOM.

Birch and Taylor (1999) investigated sources of heavy metals in sediments of Sydney Harbour and concluded that the highest concentrations of sedimentary heavy metals occur at the headwaters of embayments and tributaries in the estuary. High total suspended solids and elevated concentrations of heavy metals in sediment and in particulates within canals draining large catchments are evidence that drainage from these areas is a major source of contaminants to this estuary. Another possible important point source of heavy metals is leachates from reclamation

areas. Of the diffuse sources, atmospheric contributions may be substantial, but stormwater drains with small catchments and WWOs had no observable impression on the regional distribution of heavy metals in surficial sediments (Birch and Taylor, 1999).

Birch and Taylor (1999) further stated that the two regions of the Port Jackson catchment, which are least affected by human impact are the mainly forested sub-catchments of Darling Mills and Upper Middle Harbour Creeks. Although these areas are the closest to pristine in the Sydney Harbour catchment, heavy metal concentrations of fluvial sediments are higher than the pre-anthropogenic heavy metal data obtained from estuarine cores. They suggest that higher fluvial heavy metal concentrations are probably due to atmospheric deposition and contributions from sewer overflows, which affect even the most undeveloped areas in the catchment. Birch and Taylor (1999) cited a Water Board (1993 now Sydney Water) report which identified that of the approximate 740 ERSs spilling into the Sydney Harbour catchment, a third spill into the Middle Harbour (88) and into Lane Cove River (158). Preliminary results of modelling estimate 1000 t of suspended solids and 1.3 t of lead are spilled into the harbour via WWOs annually (Water Board, 1993). Birch and Taylor (1999) stated, that if these values are accurate, lead loading from WWOs was negligible considering the mass of lead contained in sediments of the estuary (approx. 3500 t Pb). In 1993, the two largest (by volume) WWOs spilt into Sydney Harbour from an ERS at Long Bay (3689 ML/year) and another ERS in the Lane Cove River (2740 ML/year) constituted 64% of the total sewage influent load to Sydney Harbour. Birch and Taylor (1999) commented that an absence of observable enrichment in heavy metals adjacent to these ERSs indicated that WWOs do not contribute significant metallic contaminants to this estuarine system. Since 2001, under all but extreme wet-weather conditions, spills from these two ERSs are now captured by the Northside Storage Tunnel (Figure 3-2) with up to 480 ML able to be contained from a single WWO event.

Birch (2024) recently identified road-derived metals as the chief contributor of metals to stormwater from a review and critical assessment of over three decades of research supplemented by global studies. Roads comprise almost 25% of a typical urban catchment and generate a considerable metal load from highly effective impervious surfaces which is transported directly to the adjacent receiving waterways (Birch, 2024). Within this review it is evident that copper, lead and zinc are commonly investigated in studies of road-derived metals. These metals are also commonly detected in sewer influent (Besley et al., 2023, Table 6). Hence the occurrence of copper, lead and zinc in both stormwater and in sewer influent inhibits potentially meaningful comparisons of WWOM study results of sanitary (separate from the stormwater system) WWOs with studies of combined sewer overflows (CSOs) from an intermixed sewer and stormwater system.

References

Besley, C.H., Batley, G.E., Cassidy, M. (2023). Tracking contaminants of concern in wet-weather sanitary sewer overflows. Environ. Sci. Pollut. Res. 30, 96763-96781. https://doi.org/10.1007/s11356-023-29152-x

Besley, C.H., Birch, G.F. 2019. Deepwater ocean outfalls: A sustainable solution for sewage discharge for mega-coastal cities (Sydney, Australia): Influence of deepwater ocean outfalls on shelf sediment chemistry. Mar. Pollut. Bull. 145, 707-723, https://doi.org/10.1016/j.marpolbul.2019.06.009

Bickford, G., Toll, J., Hansen, J., Baker, E., Keessen, R. 1999. Aquatic ecological and human health risk assessment of chemicals in wet weather discharges in the Sydney region, New Sout Wales, Australia. Mar. Pollut. Bull. 39, 1-12.

Birch, G.F. 2024. Review and assessment of road-derived metals as a major contributor of metallic contaminants to urban stormwater and the estuarine environment (Sydney estuary, Australia). J. Hazard Mater. 465, 133096. https://doi.org/10.1016/j.jhazmat.2023.133096

Birch, G.F., Taylor, S. 1999. Source of heavy metals in sediments of Port Jackson estuary, Australia. Sci. Total Environ. 227, 123-138.

Dept. of Environment. 1997. Some Previous Uses of Organochlorine Pesticides in Australia. Scheduled Wastes Fact Sheet Number 4. Department of Environment, Canberra, Australia.

Radcliffe, J.M. 2002. Pesticide Use in Australia. Australian Academy of Technological Sciences and Engineering. Parkville, Victoria.

4.2 Pharmaceuticals and other organic chemicals

Birch et al. (2015) investigated estuarine waters from 30 sites adjacent to stormwater outlets across the entire Sydney estuary after a dry period. They concluded that the presence of eight (of 59 tested) pharmaceuticals and acesulfame (an artificial sweetener) in water from all parts of the Sydney estuary, suggested untreated sewer influent is leaking from the sewerage system into the stormwater network in Sydney estuary sub-catchments. The eight detected pharmaceutical organic chemical compounds were codeine (an analgesic), paracetamol (acetaminophen a very common analgesic), tramadol (another analgesic), venlafaxine (an antidepressant), propranolol (antiarrhythmic and antihypertensive), fluoxetine (another antidepressant), iopromide (an iodinated contrast agent) and carbamazepine (an anti-convulsive). Birch et al. (2015) also identified 7 (of 38 tested) pesticides reflecting stormwater transport to the estuary.

An initial assessment of the presence of pharmaceuticals at proposed WWOM study sites was conducted in 2016 (unpublished data of the WWOM). Dry-weather water column testing used a pharmaceutical scan of 13 chemicals that were available from laboratory testing already in place at the Sydney Water laboratory. Results of this scan were assessed against sewer network records of repair jobs. Outcomes of this testing confirmed reported issues with network faults and identified a few additional sites. That work suggested that 66% (27of 41) of sewer catchments were intact and without dry-weather leakage. This understanding helped inform the selection of sites for the ecology studies of the WWOM. Five of the 13 chemicals tested: acetaminophen, ibuprofen, naproxen (an analgesic), diclofenac (an anti-inflammatory), and carbamazepine; were also documented in this assessment. These findings along with the Birch (2015) findings informed the decision to further explore a broader suite of pharmaceuticals and personal-care products to understand if any of these contaminants are potentially of concern in sewer overflows.

Since the WWOM study commenced in 2016, tracers of sewage have been identified in other regional studies. Currens et al. (2019) used acetaminophen and sucralose (a sweetener) as co-analytes to track sewage in receiving waters, while White et al. (2019) employed acetaminophen and ibuprofen to trace untreated sewage inputs in receiving waters and determined sucralose to be

an excellent tracer of both treated and untreated sewage inputs. These studies suggest that concentrations of organic chemicals are detectable in the water column. But are they contaminants or at concentrations that represent adverse ecological risk? These two studies did not comment upon this question.

References

Birch, G.F., Drage, D.S., Thompson, K., Eaglesham, G., Mueller, J.F. 2015. Emerging contaminants (pharmaceuticals, personal care products, a food additive and pesticides) in waters of Sydney estuary, Australia. Mar. Pollut. Bull. 97, 56-66.

Currens, B.J., Hall, A.M., Brion, G.M., Fyar, A.E. 2019. Use of acetaminophen and sucralose as con-analytes to differentiate sources of human excreta in surface waters. Water Res. 157, 1-7.

White D., Lapworth, D.J., Civil, W., Williams, P. 2019. Tracking changes in the occurrence and source of pharmaceuticals within the River Thames, UK; from source to sea. Environ. Pollut. 249, 257–266.

4.3 Examining a broad suite of organic chemical markers of sewage spills or stormwater inflows to receiving waters

Chemical markers (tracers) of sewage can be grouped into three categories:

a) produced in the human body (for example faecal sterols)

b) ingested and pass through the body (for example pharmaceuticals and food additives)

c) associated with washing and laundering (for example detergents and whitening agents).

To evaluate the presence of organic chemicals from the abovementioned categories a collaborative study was undertaken with the Centre for Aquatic Pollution Identification and Management (CAPIM) of the University of Melbourne with laboratory work conducted in conjunction with Emeritus Professor Kiwao Kadokami of the Institute of Environmental Science and Technology, University of Kitakyushu, Japan. The latter study enabled the simultaneous screening of 1392 organic chemicals using the combination of both gas chromatography/mass spectrometry (GS-MS) and liquid chromatography/quadrupole time-of-flight mass spectrometry (LC-QTOF-MS) database methods developed by Professor Kadokami and his team (Kadokami et al., 2005, Kong et al., 2015, Kadokami and Ueno, 2019).

Objectives of the pilot study investigating contaminants of concern in urban receiving waters

In this study, the Chemcatcher™ and Polar Organic Chemical Integrative Samplers (POCIS) passive samplers were chosen to investigate trace organic chemical residues in urban receiving waters. Two types of Empore disks (SDB-XC and RPS) used with the Chemcatcher™ system, and Oasis HLB sorbent were deployed in a POCIS sampler. Companion water samples were also collected (Figure 4-1).

The main aim of this study was to establish the number, type and concentrations of chemicals present at study locations and to derive in-situ sampling rates (R_s) values for detected chemicals.

This was achieved by screening extracts of passive sampler disks/sorbents and composite water samples by multi-residue GS-MS and LC-QTOF-MS AIQS database methods outlined above.

A second aim was to assess deployment timeframes for each passive sampler system by comparing the linear uptake phase for chemicals detected across each of the six deployment weeks.

Text, graphic and citations in this section were drawn from the peer-reviewed journal publication: Allinson, M., Cassidy, M., Kadokami, K., Besley C.H. 2022. In situ calibration of passive sampling methods for urban micropollutants using targeted multiresidue GC and LC screening systems. Chemosphere, 311, 136997. https://doi.org/10.1016/j.chemosphere.2022.136997. Text and citations in this section were drawn from this peer-reviewed journal publication and the following citations.

Liu, W., Xue, J., Kannan, K. 2017. Occurrence of and exposure of benzothiazoles and benzotriazoles from textiles and infant clothing. Sci. Total Environ. 592, 91–96.

Shi, Z., Liu, Y., Xiong, Q., Cai, W., Ying, G. 2019. Occurrence, toxicity and transformation of six typical benzotriazoles in the environment: a review. Sci. Total Environ. 661, 407–421.

Zhao, X., Zhang, Z.F., Xu, L., Liu, L.Y., Song, W.W., Zhu, F.J., Li, Y.F., Ma, W.L. 2017. Occurrence and fate of benzotriazoles UV filters in a typical residential wastewater treatment plant in Harbin, China. Environ. Pollut. 227:215–222.

Receiving water pilot study locations and deployed equipment

This study was conducted at three locations, with two freshwater streams sampled and an estuarine channel (Figure 4-2A). At the start of each round of sampling six passive sampler canisters (Figure 4-2B) were deployed with each containing two types of Empore disks (SDB-XC and RPS) used with the Chemcatcher™ system, and Oasis HLB sorbent deployed in a POCIS sampler (Figure 4-1). At the end of each week a passive sampler cannister was retrieved along with collection of the composite water sample that had been collected across the week by an autosampler (Figure 4-1). Three rounds of sampling were conducted.

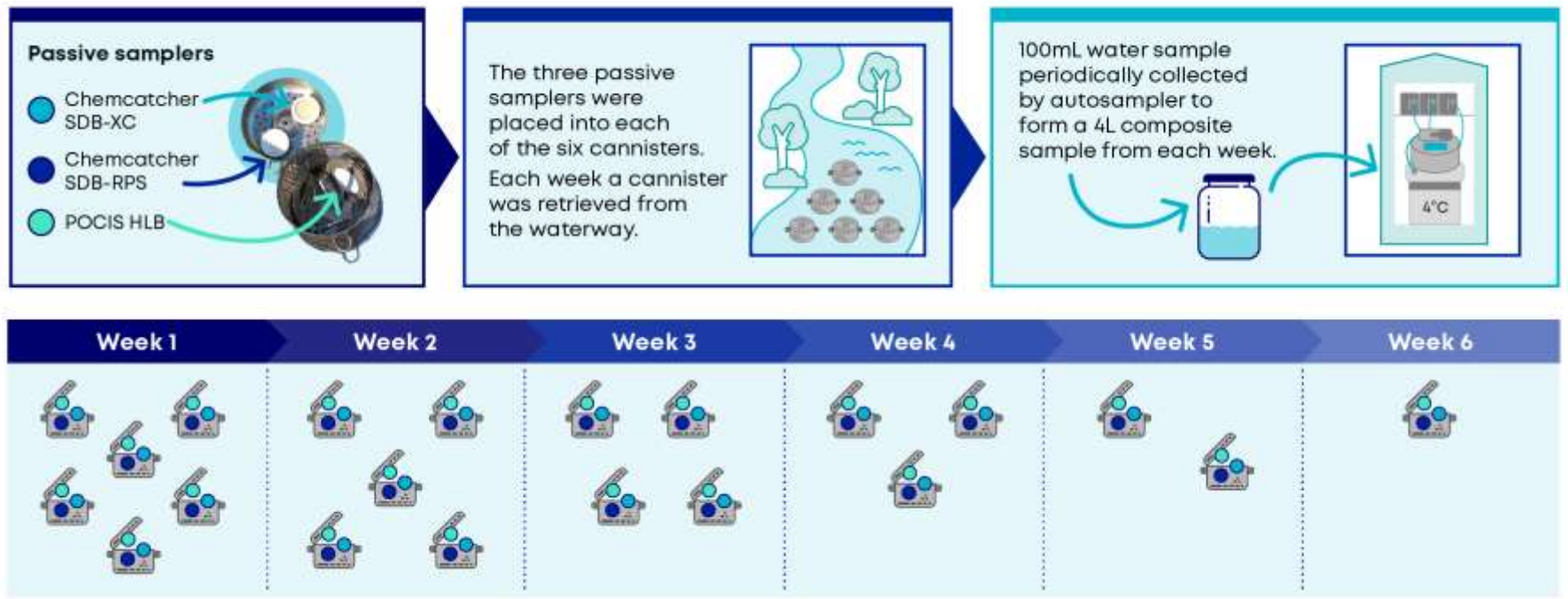


Figure 4-1: Graphical abstract showing the in-situ passive sampler calibration study

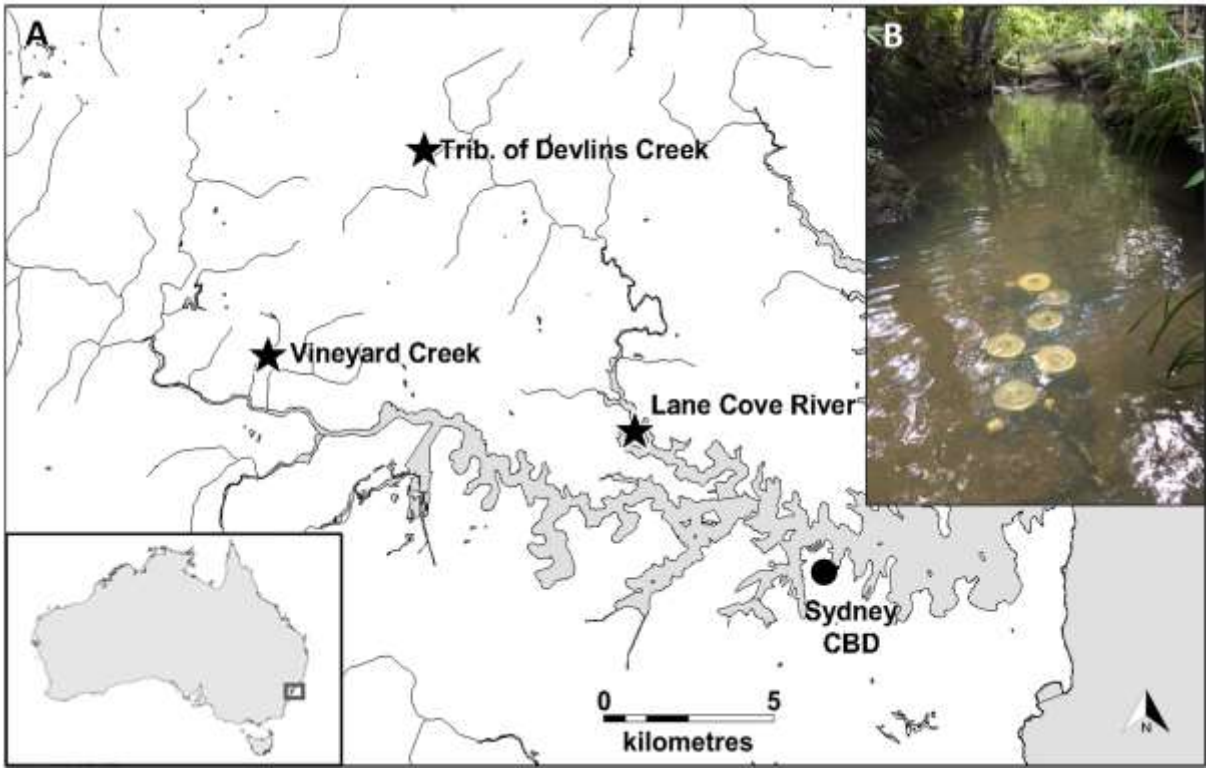


Figure 4-2: Location map showing in-situ passive sampler calibration study sites (A) and a photo showing initial deployment of six passive sampler canisters (B)

Main findings

A total of 254 organic chemicals were detected out of the simultaneously screened 1392 organic chemicals assessed for at the three urban receiving water study sites of the Sydney region. Water samples were assessed for stormwater transported contaminants and contaminants delivered by WWO spills. The most frequently detected compounds under GC/MS analysis were aliphatic hydrocarbons, pesticides, phenols, polycyclic aromatic hydrocarbons (PAHs), sterols and fatty acid methyl esters. While from LC/QTOF-MS analysis pesticides, pharmaceuticals, and personal-care products were mainly detected.

The study found that employing sorbents of Chemcatcher^{TM} SDB-RPS and POCIS Oasis HLB yielded a relatively similar set of persistently detected organic chemicals. Results from regression analysis of these two passive sampler sorbents suggested, for a given chemical and for a given sampling period, there was better agreement between different passive sampler sorbents deployed within a site than between passive sampler sorbents deployed across sites. That is, the type of sorbent used matters less than the site at which the passive sampler sorbent is placed. In contrast, Chemcatcher^{TM} SDB-XC yielded a different set of persistently detected organic chemicals with few chemicals suitable for regression analysis.

Regression analysis found that optimum deployment durations were as short as two weeks or up to six weeks for some chemicals (Table 5, Allinson et al., 2023). This variation in deployment durations together with differing sampling rates (R_s) across persistently detected chemicals suggested that if quantitative data were an objective, it would be advantageous to conduct a site-

specific pilot study to inform designing a longer-term monitoring study based upon passive sampling techniques. Depending on the target chemicals of interest, either Chemcatcher™ SDB-RPS or POCIS Oasis HLB would better reflect periods with relatively higher stormwater runoff that was likely to have transported contaminants into the receiving water column, compared to Chemcatcher™ SDB-XC. Salinity did not seem to influence ranges of sampling rates (R_s) with one exception collected with Chemcatcher™ SDB-XC, which suggested that either Chemcatcher™. SDB-RPS or POCIS Oasis HLB would be preferable passive sampler sorbents to deploy if a mix of freshwater and estuarine sites were to be examined.

Selected from this study was a suite of 33 organic chemical markers of sewage based on potential sole human usage. While 19 organic chemical markers of stormwater were selected based upon knowledge of disposal not being via the sewer (Section 4.4). Chemicals with potentially ubiquitous sources of disposal via the sewer and present in urban stormwater runoff were not considered for further WWOM study. The exception to this were benzotriazoles which have numerous potential sources in sewage including as corrosion inhibitors in dishwashing detergents (Shi et al., 2019), as ultraviolet filters in personal-care products (Zhao et al., 2017), along with their incorporation into textiles (Liu et al., 2017).

Key sub-study outcomes

The passive sampler sorbents of Chemcatcher™ SDB-RPS or POCIS Oasis HLB were preferable to deploy to assess a mix of freshwater and estuarine sites, provided they captured the target sewage marker chemicals (tracers) of interest. Application of this outcome is discussed further in Section 6.2.

Out of the potential 1392 organic chemicals, 254 chemicals were detected with a suite of 33 organic chemicals selected as markers (tracers) of sewage presence in receiving waters. While another 19 organic chemicals were identified as markers of stormwater inflows into receiving waters. Application of this outcome is discussed further in Section 4.4, which outlines the study to evaluate whether they pose a threat to ecosystem health.

4.4 Tracking contaminants of concern in the water column

Prior to this sub-study commencing, the existing pharmaceutical scan of 13 chemicals available at the Sydney Water laboratory was expanded to accommodate the suite of sewer marker chemicals determined under the previous study (Section 4.3). Similar pre-work was also required to implement capability to also analyse the suite of stormwater marker chemicals. This pre-work demonstrated another aspect of the WWOM facilitating capability uplift within Sydney Water. Proprietary patents on an optical brightener (FB71) prevented Sydney Water from sourcing the relevant chemical standard, as such this chemical was omitted as a tracer of sewage. Sydney Water's Laboratory Services already had in place a well-established capability to assess ammonia and metals, which was also utilised under this study.

The suite of 19 stormwater marker chemicals discerned from the prior study (Section 4.3), were not detected in the water column (freshwater and estuarine) from any of the storm events sampled in this subsequent sub-study that tracked organic contaminants of concern. The lack of detection in the water column occurred despite sampling efforts that included targeting of low rainfall events (without ERS spills) to examine first flush conditions where entry of turbid water first commenced. Hence, these stormwater marker chemicals are not discussed further in this section. Detections of stormwater marker chemicals under passive sampling are discussed in Section 6.2.

Of the suite of sewer marker chemicals, only 18 (Table 2, Besley et al., 2023) were detected within the sewer influent water column and these were the focus of the following pilot study.

Objectives of the pilot study tracking contaminants of concern in WWOs

This pilot study was conducted to assess 18 detected organic contaminants together with ammonia and metals in the sewerage system. The study had five objectives:

(i) to examine the types of contaminants detected in the water column of the sewer and in associated downstream receiving waters

(ii) to determine if the water-column-detected organic contaminants were similar across the four sewer carriers and associated five receiving water sites

(iii) to assess dilution of influent in the receiving waters

(iv) as an additional line of evidence to assess dilution using companion human faecal-associated marker gene (HFMG) microbial source tracking (MST) data.

(v) to review the aquatic toxicity of water-column-detected contaminants

A graphical outline of this sub-study is provided in (Figure 4-3).

Text, graphic and citations in this section were drawn from the peer-reviewed journal publication: Besley, C.H., Batley, G.E., Cassidy, M. 2023. Tracking contaminants of concern in wet-weather sanitary sewer overflows. Environ. Sci. Pollut. Res. 30, 96763-96781. https://doi.org/10.1007/s11356-023-29152-x

Pilot study locations

ERSs at four locations in the sewerage system were selected for study after consideration of safe access for field staff under various weather conditions. These ERS overflowed (spilled) to Gymea Bay, Gymea; Darling Mills Creek, Baulkham Hills; Buffalo Creek, East Ryde; and Vineyard Creek, Dundas (Figure 4-4). They represented one low, two medium, and one high volume and frequency WWOs, respectively (as outlined in Supporting information Table S1 of Besley et al., 2023). An autosampler was installed to collect samples from the sewer carrier at each of these locations along with an autosampler installed to allow simultaneous water column sampling at companion downstream receiving water sites. Two autosamplers were installed at the estuarine Gymea Bay location, to allow sampling of the receiving waters as ERSs spilled into each of the two arms of this embayment (Figure 4-4). This study was undertaken between October 2018 and February 2020, please see the 'Field sample collection' section of Besley et al. (2023) for details of the various sampling events.

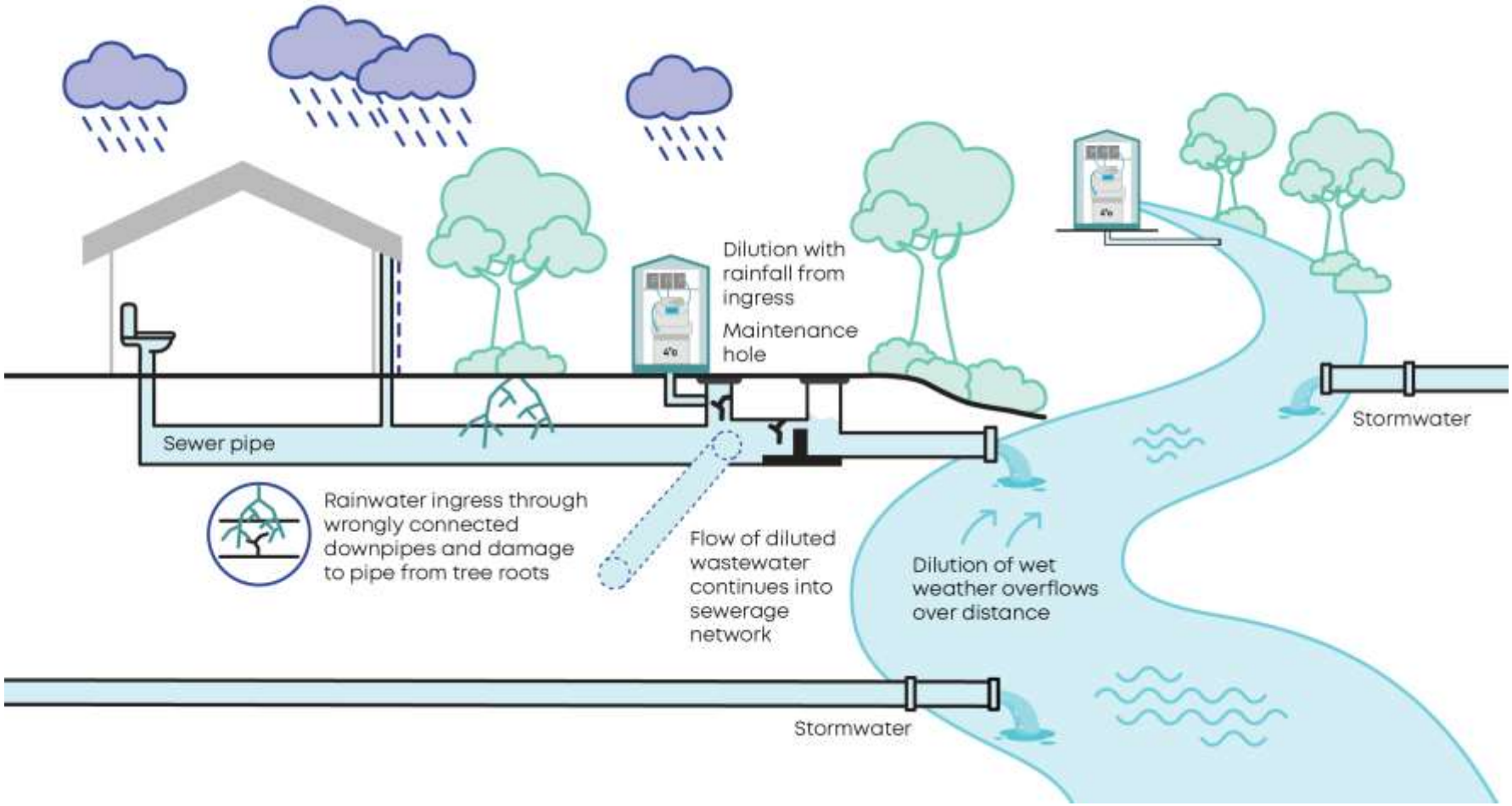


Figure 4-3: Graphical abstract for the tracking of contaminants of concern in wet-weather sanitary sewer overflows study

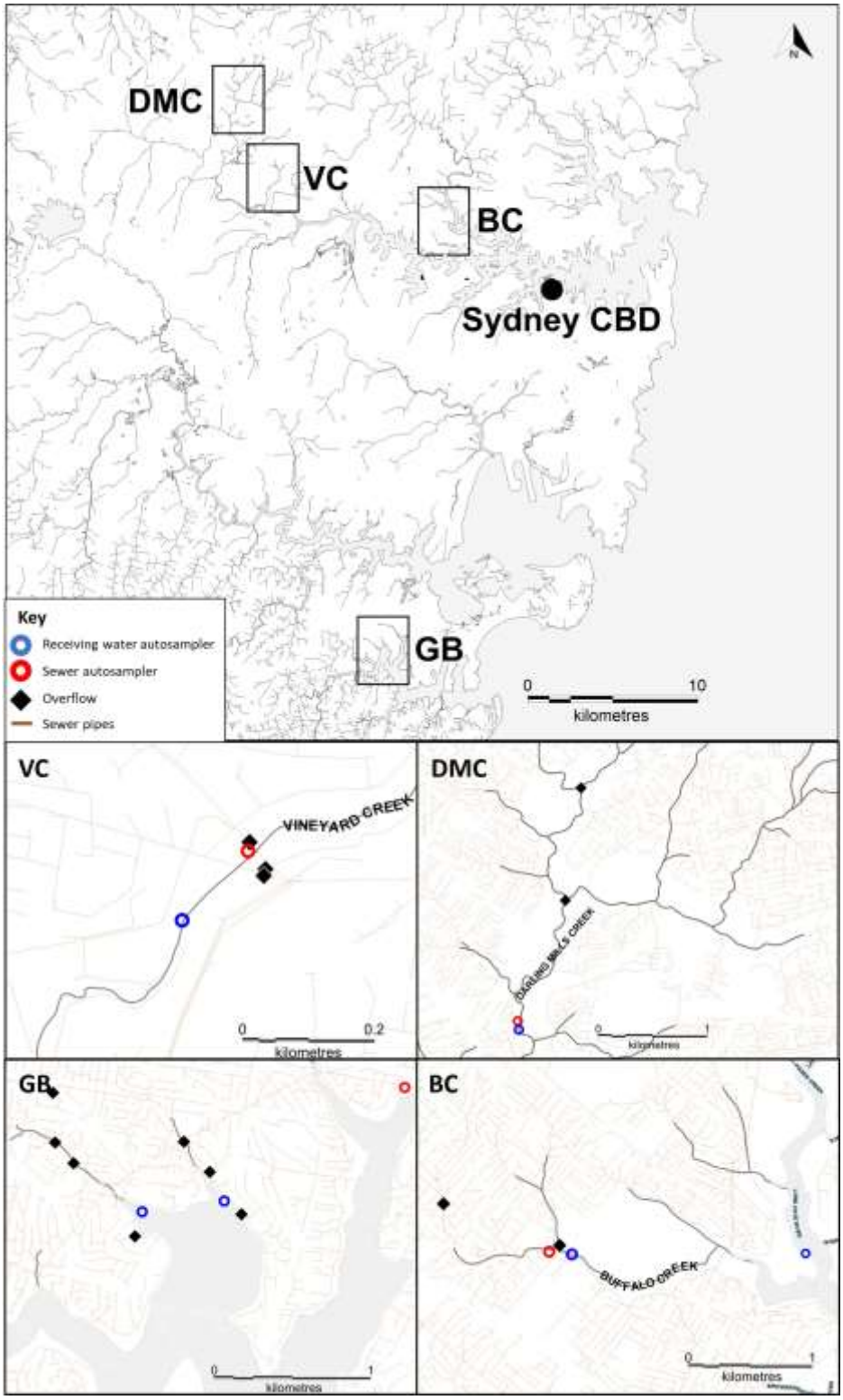


Figure 4-4: Overall location map and detail of each of the five study sites for tracking contaminants VC = Vineyard, DMC = Darling Mills; GB = Gymea (two sites at this location), and BC Buffalo. Red circles represent sewer autosamplers, blue circles represent downstream receiving water autosamplers, black diamonds indicate position in sewer catchments of gauged ERS (sewer overflow points), and brown lines represent sewer main pipes of these urban study locations.

Study objectives (i) and (ii): Detected contaminants

The 18 organic compounds tracked were found in concentrations above laboratory quantitation limits. This suite of 18 compounds was analysed for in-sewer influent, both in dry weather and in wet-weather under rainfall ingress conditions when overflow spills were occurring from the ERSs. Not surprisingly, the greatest number of detectable contaminants was seen in the dry-weather samples, where all 18 organic chemicals were in the autosampler-collected influent samples. The detected compounds had the following order of concentration:

acetaminophen > metformin > theobromine > sucralose >ibuprofen > three benzotriazoles

with others close to quantitation limits. Under wet-weather conditions, the same order prevailed with only 12 (Table 2, Besley et al., 2023) of the dry-weather chemicals detectable. A very similar pattern of chemical concentrations was seen in receiving waters downstream of ERSs when WWO spills were occurring. In the receiving waters, the number of detectable chemicals reduced to those eight chemicals listed above in the highest concentrations.

A view across the five receiving water sites revealed a contrasting pattern of organic chemical concentrations compared to the relatively uniform pattern observed across the four sewer carriers, when samples were collected under diluted rainfall ingress conditions.

Analysis of receiving water samples (Supplementary Table S6, Besley et al., 2023) from the Darling Mills site and from both Gymea Bay receiving water sites provided results:

- all at the quantitation level except for 1H-benzotriazole as shown at Gymea Bay Site 1
- only four contaminants (ibuprofen, metformin, sucralose, and theobromine) were at measurable concentrations in samples from the Vineyard receiving water site
- acetaminophen and the three benzotriazoles, were above quantitation limits in both Buffalo and Vineyard receiving water samples
- concentrations of the benzotriazoles were relatively similar between Buffalo and Vineyard Creeks
- acetaminophen was at relatively higher concentrations in Vineyard Creek.

The above data patterns may be partly explained by the placement of downstream receiving water sites. The Buffalo and Darling Mills sites had the same carrier diameter, but safe access considerations influenced the placement of Buffalo Site 50 m downstream of the ERS, compared to the more distant placement at Darling Mills some 950 m downstream (Figure 4-4). At Gymea Bay, site placement catered for the two urban streams, draining into this estuarine bay, that both receive sewer overflows (Figure 4-4). Hence, a site was placed at the mouth of each stream within the bay, with the nearest ERS to each bay arm situated at 400 m and 150 m to these stream mouths (Figure 4-4). The ERS with the relatively highest modelled volume for the Gymea study location was nearest to Gymea Site 2. All ERSs under study at the Buffalo, Darling Mills, and Gymea locations were gravity fed. Safe access to Vineyard Creek occurred 140 m downstream from the three ERSs that spill within 10 m of each other into the same waterway reach (Figure 4-4). The highest modelled volume ERS at this location was a siphonic overflow discharge structure that conveyed spills from a 2100-mm diameter sewer main, while on the opposite side of

the creek the other two ERSs were gravity fed and spilt from a 1830-mm sewer main, although a small connecting pipe allowed one of these two ERSs to also spill influent from the 2100-mm sewer main. Hence, the Vineyard location was atypical with three ERSs situated very close (clustered) together (Figure 4-5), compared to the other three study locations where those sewer trunk and associated reticulation pipes that drained urban areas had ERSs that were spatially separated along the sewerage system (Figure 4-4).

Measured overflow spill volumes do not appear to explain the detection of measurable concentrations in the receiving waters. For example, a relatively higher median spill volume occurred into Darlings Mills Creek (4.8 ML) when compared against the relatively lower median spill volume into Buffalo Creek (0.3 ML). The opposite pattern to that recorded for detection of measurable organic chemical concentrations would have been expected for these two sites. A more plausible explanation is the dilution over distance (50 m for Buffalo versus 950 m for Darling Mills) from an ERS source of a WWO spill, as the explored patterns of measurable concentration detection across sample events seem supportive given there were no measurable detections for any of the eight organic chemicals at Darling Mills, while four organic chemicals were recorded at measurable concentrations in the receiving waters of the Buffalo site.

Figure 4-5: Three adjacent ERSs spilling influent into Vineyard Creek and stormwater flows

The photograph shows ERS AGN 8723124, SO0014 (at left of the photograph) in a siphonic discharge state, while on the right-hand side of the photograph gravity fed ERS AGN 1376595, SO0016 and AGN 1374307, SO0017 (at far right) were also spilling.

Study objectives (iii) and (iv): Evaluating dilutions

Pooling influent samples collected from all four sewer carriers allowed a comparison of median concentrations of organic contaminants in sewer influent collected in dry weather and under wet-weather rainfall ingress diluting conditions that coincided with overflow spills. Pooled samples had a relatively tight range of initial dilutions from 4.9 times (×) to 9× for five of six contaminants assessed. The sixth contaminant, sucralose, exhibited a dilution of 23×. The companion line of evidence that compared dry- to wet-weather concentrations of molecular markers of microbial contaminants (HFMGs) measured in sewer influent returned dilutions of 3.8× for Bacteroides HF183 and 12.4× for crAssphage CPQ_056, with both results supportive of dilutions determined from organic contaminants. These trends are displayed in Table 4 of Besley et al. (2023).

For WWO events, evaluation of sewer influent and companion receiving water samples of a study location revealed differing dilutions. For example, companion measurements of HFMGs collected during wet-weather overflow spills on February 22 and March 1, 2022, from Darling Mills and Buffalo locations, respectively, provided evidence of dilution in the receiving waters. The Buffalo receiving waters were sampled 50 m downstream of the spilling ERS returned dilutions of 4.9× for Bacteroides HF183 and 6.4× for crAssphage CPQ_056. Higher dilutions were observed at the more distant receiving water location for Darling Mills, some 950 m downstream of the ERS, with dilutions of 28× for Bacteroides HF183 and 65× for crAssphage CPQ_056. These trends are displayed in Table 4-1, reproduced from Table 5 of Besley et al. (2023).

Table 4-1: Dilutions of organic chemicals and human-associated MST marker genes from measurements in downstream receiving waters (RW) and sewer influent collected under wet-weather rainfall

Contaminant	Darling Mills February 2, 2022			Darling Mills January 24, 2023			Buffalo March 1, 2022			Buffalo July 4, 2022		
	influent	RW	Dilution	influent	RW	Dilution	Influent	RW	Dilution[a]	Influent	RW	Dilution[a]
Ammonia, mg total ammonia N/L	1.9	0.1	19x	2.79	0.1	28x	0.8	0.2	4x	0.8	0.2	4x
Acetaminophen, µg/L	<1	<1	-	ns	ns	-	6	<1	>6x	ns	ns	-
Metformin, µg/L	12	<1	>12x	ns	ns	-	13	2	6.5x	ns	ns	-
Theobromine, µg/L	3	<1	>3x	ns	ns	-	2	<1	>2x	ns	ns	-
Number of samples	1	1		5	5		1	1		7	7	
Human-associated MST marker	**geometric mean**			**geometric mean**			**geometric mean**					
	influent	**RW**	**Dilution[b]**	**influent**	**RW**	**Dilution[b]**	**Influent**	**RW**	**Dilution[b]**	**Influent**	**RW**	**Dilution[b]**
	log_{10} GC/L			**log_{10} GC/L**			**log_{10} GC/L**					
CrAssphage CPQ_056	8.70	6.89	65x	8.38	6.47	81x	8.63	7.83	6.4x	ns	ns	-
Bacteroides HF183	8.49	7.04	28x	8.32	6.57	57x	8.77	8.08	4.9x	ns	ns	-
Number of samples	1	1		2	2		1	1				

Contaminant	Vineyard October 5, 2022			Vineyard October 6, 2022			Vineyard October 9, 2022		
	influent	RW	Dilution[a]	influent	RW	Dilution[a]	Influent	RW	Dilution
Ammonia, mg total ammonia N/L	9.4	2.3	4.1x	4.9	5.6	0.9x	1.1	1.0	1.1x
Acetaminophen, µg/L	ns	ns	-	ns	ns	-	<1	<1	-
Metformin, µg/L	ns	ns	-	ns	ns	-	38	33	1.2x
Theobromine, µg/L	ns	ns	-	ns	ns	-	1	<1	>1x
Number of samples	5	5		5	5		1	1	
Human-associated MST marker	**geometric mean**			**geometric mean**			**geometric mean**		
	influent	**RW**	**Dilution[b]**	**influent**	**RW**	**Dilution[b]**	**Influent**	**RW**	**Dilution[b]**
	log_{10} GC/L			**log_{10} GC/L**			**log_{10} GC/L**		
CrAssphage CPQ_056	8.91	8.61	2.0x	8.73	8.82	0.8x	7.99	7.91	1.2x
Bacteroides HF183	9.02	8.74	1.9x	8.98	9.02	0.9x	8.36	8.20	1.4x
Number of samples	2	2		2	2		1	1	

ns = not sampled
[a] dilutions based on mean values
[b] dilutions based on back-transformed geometric means
GC/L = gene copies / L

During the weather event of July 4, 2022, a mean concentration of 0.2 mg total ammonia N/L was recorded in the receiving water downstream of the Buffalo ERS in samples collected while a WWO spill occurred. Companion measurement in the influent during this event recorded 0.8 mg total ammonia N/L. These influent and receiving water results mirrored total ammonia N/L from March 1, 2022, and suggested that over the 50 m distance from the ERS to the downstream autosampler, a dilution of 4× had occurred in the receiving waters at the Buffalo location. This dilution of 4× is supportive of the above-mentioned HFMG results as was the metformin dilution of 6.5× recorded at this location. The trend of relatively higher dilutions observed in MST results recorded at the more distant receiving water location of Darling Mills was also observed for total ammonia N/L where dilutions of 19× and 28× were recorded for the February 2, 2022, and January 24, 2023, samples, respectively. These diluted ammonia concentrations measured in the receiving waters for both the Buffalo and Darling Mills sites were at least 4× less than the ANZG (2018) water quality guideline value (0.79 mg total ammonia N/L). This suggests that the risk to ecosystem health posed by influent concentrations of ammonia was abated by dilution of influent in receiving waters at the distances that sites were situated in the current study. These trends are displayed in Table 4-1, reproduced from Table 5 of Besley et al. (2023).

The ERSs under study at the Buffalo and Darling Mills locations were gravity fed with influent transported by their respective sewer main pipes from the local urban area, with the Buffalo and Darling Mills locations being more representative of urban areas in Sydney where spatially separated ERSs spill to freshwater urban streams (Figure 4-4). In contrast, the Vineyard location is atypical of the local sewerage system with the merger of trunk main pipes that transport influent from a larger geographic area and three ERSs that spill within 10 m to the same reach of a small urban stream (Figure 4-4 and Figure 4-5), with one of these ERSs capable of siphonic discharge flows. The siphonic discharge is initiated by a gravity flow, which then primes a siphonic pipe that accelerates the spill rate. Examples of siphonic and gravity-fed spill rates are available in Besley and Cassidy (2021, Section 2.2). The combination of the accelerated siphonic flows from one ERS together with the two-gravity fed ERS spills at the same point into this relatively small urban stream may help explain the limited receiving water dilutions of around 1 observed at this location on October 6 and 9, 2022. In contrast, at the start of the October weather event on October 5, 2022, dilutions of 2× to 4× were recorded for HFMGs and ammonia. This suggests that later in a protracted spill event at the Vineyard location, the flow in the receiving waters of Vineyard Creek comprised mostly of rainwater ingress diluted influent (Table 4-1, reproduced from Table 5 of Besley et al., 2023).

Sampling of these three October 2022 collection events across this wet-weather period illustrated that the influent within the sewer pipe became most highly diluted from rainwater ingress as cumulative rainfall increased (Table 6 of Besley et al., 2023). This in turn suggests that the ecological risk posed in receiving waters from ammonia diminished as spill duration from the ERSs increased at the Vineyard location. An ecological risk was potentially present from recorded ammonia concentrations in receiving waters due to the relatively low dilution confirmed by estimates from companion lines of evidence (Table 4-1, reproduced from Table 5 of Besley et al., 2023). The few other organic chemicals with detected concentrations in the sewer influent were not present in measurable concentrations within the receiving waters, and as such did not yield reliable dilution values.

Three differing storm events that coincided with WWOs from one, three or all ERS were sampled for the estuarine Gymea Bay location with ammonia and HFMG lines of evidence assessed. Very high receiving water dilutions of 68× to 300× were suggested from results for ammonia and from HFMGs with over 500× dilution. This suggests that an ecological risk is potentially not posed by gravity-fed ERS spills into the tidally flushed Gymea Bay estuarine receiving water environment from the lines of evidence assessed in the current study.

Thus, the additional line of evidence to assess dilution provided by companion HFMG data was supportive of dilutions determined with the two lines of evidence provided by organic chemical and ammonia data.

Study objective (v): Toxicity of detected contaminants and the risk of adverse ecological impacts

In this study, the toxicity of organic chemicals was compared against toxicity values identified in scientific literature as set out in Table 7 of Besley et al. (2023). Looking at the measured concentrations of the identified contaminants of potential concern in Table 2 (of Besley et al., 2023), it would appear that none of the chemicals are likely to pose concerns for ecosystem health before dilution, and even less likely after dilution in the receiving waters (Table 2 of Besley et al., 2023), when compared to toxicity data outlined in Table 7 of Besley et al. (2023). It should be noted that there were chemicals for which no toxicity data could be obtained, but these chemicals were not present in high concentrations in the overflow waters.

The companion measurements of bioavailable (dissolved) concentrations of zinc and copper in wet-weather ingress diluted influent were above water quality guideline values (ANZG, 2018). However, any concern for ecosystem health from these two metals was potentially ameliorated by relatively high dissolved organic carbon concentrations present in the influent. Although iron concentrations in these samples exceeded the guideline value (ANZG, 2018), there is evidence that it is likely to be complexed by natural organic ligands and not necessarily bioavailable or toxic (Nagai et al. 2007).

Of greater concern are the high ammonia concentrations from the Vineyard sewer (Table 6 in Besley et al., 2023). The guideline value (ANZG, 2018) exceedances were, respectively, 84× (in dry-weather), 20× (intermediate wet-weather), 6×, 13×, 7×, and 1.5× for the wet-weather samples. The May 6, 2021, wet-weather event (19 mm in 24 h to 9 am on day of sampling) sampled from Vineyard with 20× exceedance of the guideline value is potentially representative of the least dilution under rainfall ingress conditions as sampling was conducted the day before an overflow spill occurred from this sewer main as sampling was inadvertently triggered by an ERS situated on another sewer main that joins with this main. Samples in the second wet-weather event of July 4, 2022, from Vineyard were correctly collected after the nearest ERS was confirmed to be spilling. The 6× guideline value exceedance reflects the potentially greater rainfall ingress that had occurred during this event, as 154 mm of rain fell in 120 h to 9 am on day of sampling (Table 6 in Besley et al., 2023). The highest rainfall recorded between 2018 and July 2022 was 346 mm in February 2020 (over 120 h). Diminishing trends in ammonia and for aluminum, iron, and manganese were recorded in influent across the three wet-weather collection events in early October, 2022, from Vineyard (Table 6 in Besley et al., 2023). The sample event conducted on October 5 was collected during the first hour of overflow spills that continued for 20 h into October 6, when the second collection was undertaken during the eighteenth hour of this continuous spill.

The third collection on October 9 was also collected in the fourteenth hour of a 24 h spill that commenced in the evening of October 8, 2022. Cumulative rainfall totals increased across October 5 to October 9, 2022, with 20 mm (October 5), 43 mm (October 6), and 96 mm (October 9) up to the respective sample times for these events from the start of this wet-weather period. Increased rainwater ingress into the sewerage system is likely to have occurred during this period and may explain the diminishing trend recorded from these three sampling events conducted at the Vineyard location in October, 2022 (Table 6 in Besley et al., 2023).

The lowest ammonia concentrations of 0.8 mg total ammonia N/L (pH of 7.0 and 6.9) measured in the Buffalo sewer influent under wet-weather ingress conditions on two occasions were just above the ANZG (2018) guideline value (of 0.79 mg total ammonia N/L). The low ammonia concentrations recorded in the Buffalo sewer were also approached in the highly diluted rainwater ingress influent of the Vineyard sewer on October 9, 2022, when an ammonia concentration of 1.1 mg total ammonia N/L was recorded (Table 6 in Besley et al., 2023).

Identifying suitable sewage marker chemicals and monitoring application

Of the organic chemicals assessed, none were detected in receiving water samples collected under light rainfall conditions where autosamplers were triggered after 2 mm of rainfall occurred in 15 min. This lack of detection, together with gauging of ERSs that recorded no WWOs occurring during these light rainfall events, suggests that the origin of these contaminants is from a sewer source. As such, these organic contaminants provide a line of evidence to track WWOs. The eight chemicals (acetaminophen, ibuprofen, metformin, sucralose, theobromine and three benzotriazoles) detected in the receiving waters during overflow spill events, as well as in both dry-weather influent and wet-weather ingress diluted influent, are potential chemical markers (tracers) to indicate sewage contamination. As outlined in Section 4.2, since the WWOM commenced, three of the eight identified organic chemicals of the WWOM study have been recommended as tracers of sewage by Currens et al. (2019) and White et al. (2019).

A consideration for monitoring with these chemicals is that they may be limited by the short half-lives of some chemicals in receiving waters. The half-life for acetaminophen in river waters due to photolysis has been reported as between 56 and 77 h (Yamamoto et al., 2009) with a similar biodegradation half-life (50 h), indicating that it should persist in a river system between discharge and dilution. The rate of formation of any more toxic degradation products would therefore be slow. Few data are available for the half-lives of other overflow constituents in fresh waters. Of those found from a literature search, the in-stream decay half-life for metformin was 1 day (Caldwell et al., 2019). Chung et al. (2018) reported photolysis half-lives of 44 and 25 h, respectively, for 1-H-benzotriazole and 5-methyl-1H-benzotriazole. Based on the above half-life information, this suggests that monitoring with these marker chemicals as a line of evidence of sewage contamination may be best limited to a window of up to two days after an ERS spill.

Pilot study key findings and two recommendations

An assessment of the toxicity of the 18 organic chemicals of concern tracked in this study indicated that none appeared to pose concerns for ecosystem health before wet-weather ingress dilution, and this was even less likely after dilution in the receiving waters.

The companion measurements of dissolved concentrations of zinc and copper in wet-weather ingress diluted influent were above water quality guideline values.

A risk to ecosystem health was posed by ammonia concentrations in influent but this risk was abated by dilution in receiving waters related to the distance's sites were situated in the current study. Receiving water dilution diminished this risk at four of the five receiving water locations studied. However, at the relatively small urban stream of Vineyard Creek receiving water site, ammonia was a contaminant of potential concern. The sewerage system at Vineyard Creek has an atypical sewer main carrier convergence and ERSs spills comprise most of the flow, particularly under longer spill durations, hence ammonia may pose a risk to ecosystem health due to the limited receiving water dilutions achieved. Although, ammonia concentrations were observed to diminish as cumulative rainfall totals increased as inflow and infiltration increased into the sewerage system.

At the estuarine Gymea Bay study location, the very high dilutions of ammonia observed suggest the risk of potential adverse ecological effect is likely to be very low, and unlikely to be detected by toxicity testing. This informed the decision to focus the laboratory toxicity testing aspect of the WWOM on water samples collected from the freshwater Vineyard, Buffalo and Darling Mills pilot study locations.

Recommendations

Measurement of ammonia should be undertaken in monitoring as a key line of evidence to evaluate the risk of potential adverse ecological effect of sewage contamination of receiving waters. Evaluation of risk is made against the guideline value set out in ANZG (2018).

It is also recommended that the suite of all eight organic chemicals (acetaminophen, ibuprofen, metformin, sucralose, theobromine and three benzotriazoles) should be used together as a line of evidence in future WWOA investigations as markers (tracers) for the presence of sewage contamination up to two days after an ERS spill, based on half-life considerations. This recommendation to use this suite of all eight chemicals is influenced by the lack of uniformity of detection of these chemicals across receiving water sites after ERS spills.

4.5 Ecotoxicological pilot studies

Ecotoxicological studies under this phase of the WWOM project were conducted at the Vineyard, Buffalo and Darling Mills study locations. In 2021, in collaboration with Dr Anu Kumar of CSIRO, ecotoxicological pilot studies were commenced. Investigations of the Vineyard study location are reported in Section 4.5.1 with investigations of the Buffalo and Darling Mills study locations reported in the Section 4.5.3. A discussion of WWO spill durations is contained in Section 4.5.2 that has helped inform a modified testing approach applied in evaluation of the Buffalo and Darling Mills locations.

Ecotoxicological investigations developed an understanding of toxicity of influent under dry-weather conditions, that is, without rainfall ingress into the sewerage system. This was followed with investigations during wet-weather when rainfall ingress diluted sewer influent exists within the sewerage system, to understand how dilution may have altered toxicity of influent. Wet-weather influent samples were collected from the respective sewer carriers when WWOs were spilling to receiving waters. Companion receiving water assessments were also conducted across the three pilot study locations, with collection of water column samples when WWOs were spilling from the ERSs at these respective study locations.

4.5.1 Toxicity studies at the Vineyard location

Objectives

The overall aim of this study was to explore the relationship between WWOs, rainfall, and subsequent impacts on aquatic life. Specifically, we focused on the Vineyard Creek sewer, where we assessed the effects of sewer influent during two distinct rain events (intermediate wet-weather and wet-weather) and under dry weather. Intermediate wet-weather is represented by sampling inadvertently triggered by an ERS situated on another sewer main (as outlined under study objective (v) in Section 4.4).

This study conducted direct toxicity assessment (DTA) with acute and chronic bioassays, incorporation of toxicity identification evaluation (TIE) manipulations identified the contaminants responsible for toxicity. Three influent samples were collected representing dry weather, intermediate wet-weather and wet-weather. In addition, receiving water samples were also collected during wet weather. Three standard ecotoxicological model species, a microalga, *Chlorella vulgaris*, the water flea, *Ceriodaphnia dubia* and the midge larva, *Chironomus tepperi* were used.

Acute toxicity testing of the water flea (*C. dubia*) and the midge larva (*C. tepperi*) had a two-day (48 h) duration and were used to determine the concentration of influent that causes death during this relatively short-term exposure. While chronic toxicity tests were conducted over three days (72 h) for the microalga (*C. vulgaris*), seven days (168 h) for the midge larva and eight days for the water flea. These chronic tests estimated the influent concentration that interferes with growth and in the case of the midge larva survival was also assessed. Detailed descriptions of these toxicity testing laboratory protocols are provided in Kumar et al. (submitted). A graphical outline of this sub-study is provided in Figure 4-6. The autosamplers used under the contaminants of concern

study of Section 4.4 were used in the current study for collection of influent water column and receiving water samples (Figure 4-7).

Sample event ranking within documented ERS spill events

The lack of infiltration into the sewerage system in the 13 days preceding the collection of dry-weather influent (January 21, 2021) enabled measurement of relatively concentrated influent. The intermediate wet-weather event of May 6, 2021 had a 25th percentile rank when assessed against WWO volume from amongst the 54 WWO spill events documented between October 1, 2018, and October 31, 2022 (Supplementary Table S1 of Kumar et al., submitted). The WWO event collected on October 9, 2022, had an 83rd percentile rank, suggesting that influent concentrations were much more diluted within the sewer main from infiltration (Table S1 of Kumar et al., submitted). Corresponding rainfall totals in the preceding 72 h were 37 and 55 mm, respectively.

Acute ecotoxicological assessment outcomes

Under the two-day (48 h) acute toxicity tests the wet-weather influent and downstream samples were non-toxic to both the water flea (*C. dubia*) and midge larva (*C. tepperi*). Whereas acute toxicity was observed in sewer influent samples collected under dry weather and intermediate wet-weather conditions, although toxicity of influent was reduced under the intermediate wet-weather conditions from rain ingress into the sewerage network.

Chronic ecotoxicological assessment outcomes

The eight-day (192 h) chronic toxicity tests with the water flea (*C. dubia*) exhibited the greatest toxicity for the dry-weather influent samples, with diminishing toxicity for the intermediate wet-weather samples and toxicity further reduced in the wet-weather sewer influent and companion receiving water samples. Whereas the seven-day (168 h) chronic toxicity testing of the midge larva (*C. tepperi*) returned the same pattern as observed under acute testing with wet-weather influent and downstream samples being non-toxic.

The chronic toxicity tests conducted over three days (72 h) for the microalga (*C. vulgaris*) indicated toxicity for the dry-weather influent sample. Assessment of intermediate wet-weather samples with chronic toxicity testing did not detect a consistent trend of decreasing algal growth with increasing influent concentration, which precluded defining toxicity values, and informed the decision to not use this test to assess wet-weather influent or downstream samples.

Text, graphic and citations in this section were drawn from the peer-reviewed journal publication: Kumar, A., Batley, G.E., Adams, M., Nguyen, T.V., Nidumolu, B., Nguyen, H., Gregg, A., Cassidy, M., Besley, C.H. submitted. Ecotoxicological assessment of sanitary sewer overflows and rainfall dynamics offers insights into conditions for potential adverse ecological outcomes.

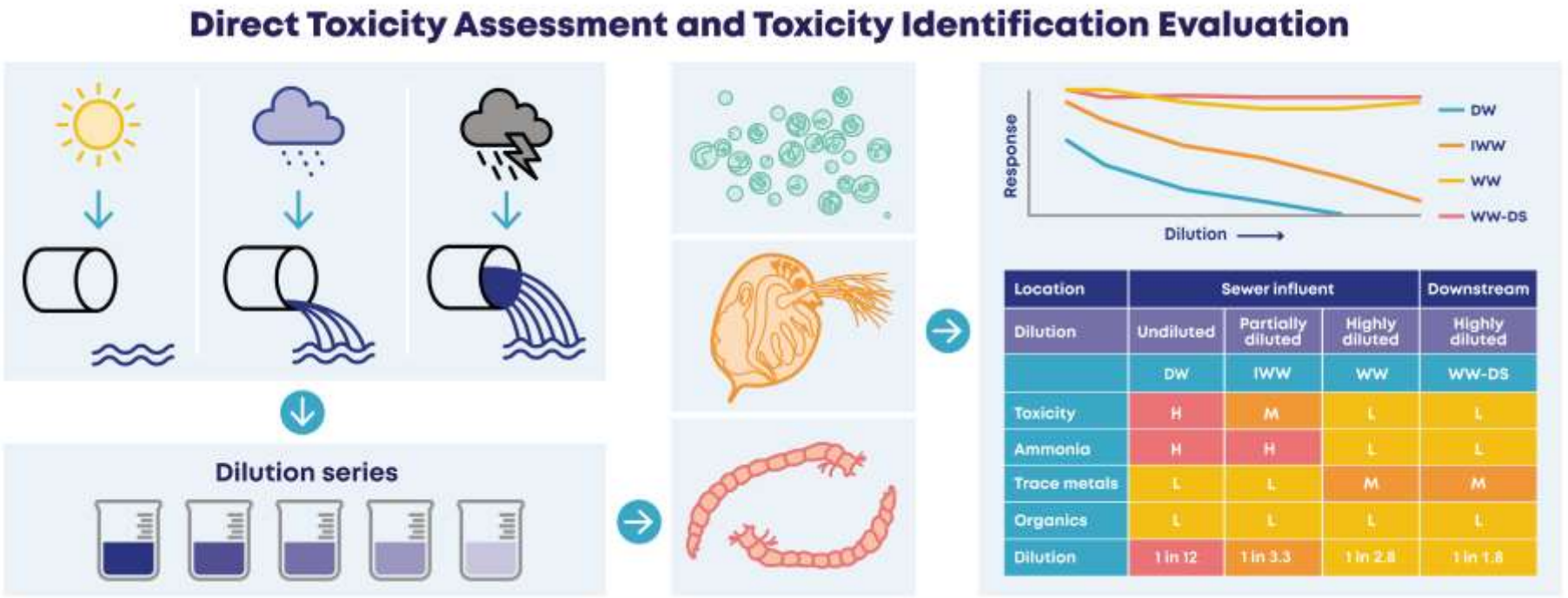


Figure 4-6: Graphical abstract for the ecotoxicological assessment study of sewer influent across various weather conditions

Weather conditions sampled: DW = dry weather; IWW = intermediate wet-weather; WW = wet-weather; WWDS = wet-weather downstream receiving waters. Overall risk assessment categories based upon hazard quotients: H = high; M = medium; L = low (Table 3 and associated text of Kumar et al., submitted)

Figure 4-7: Overall location map and detail of Vineyard Creek (VC)

Red circles represent sewer autosampler, blue circles represent downstream, receiving water autosampler, black diamonds indicate the position in sewer catchments of gauged ERSs (sewer overflow points), and brown lines represent sewer trunk and reticulation pipes of these urban study locations. The photograph at the bottom of the figure was taken when ERS AGN 8723124, SO0014 (at left of the photograph) was in a siphonic discharge state, while on the right-hand side of the photograph gravity fed ERS AGN 1396981, SO0016 and AGN 1374307, SO0017 (at far right) were also spilling.

Toxicity identification and evaluation to assess sources of toxicity

Potential toxicants, with concentrations high enough to cause toxicity to the water flea (*C. dubia*) were trace metals and ammonia. To confirm this, acute (two-day) TIE testing was applied to dry weather and intermediate wet-weather influent samples, with toxicity changes assessed after addition of ethylenediaminetetraacetic acid (EDTA) for metals, and adjustment of the influent pH to assess pH-dependent ammonia toxicity.

TIE manipulation involving pH adjustment and aeration of the dry weather and intermediate wet-weather samples significantly reduced the baseline toxicity to midge and water flea during two-day exposures. This assessment confirms ammonia is a component of toxicity in these influent samples.

Metals present in influent were also identified as chemicals of concern. Addition of EDTA reduced the toxicity but did not completely remove the toxicity in dry weather and intermediate wet-weather influent samples. Copper and zinc were identified as toxicants of concern from the tested metals.

Diminishing ammonia concentration within influent as rainfall ingress into the sewer increased

Ammonia levels were 79× higher in dry weather compared to the ANZG (2018) guideline value for 95% species protection. However, in intermediate wet-weather, ammonia toxicity substantially decreased to 19×, and further dropped to 1.4× in the wet-weather sampling event. Besley et al. (2023) measured ammonia in the receiving waters for both the Buffalo and Darling Mills freshwater creeks to be at least 4× less than the guideline value. Outcomes of these water column studies by Besley et al. (2023) indicate that the risk of adverse ecological effect from influent ammonia concentrations was reduced through dilution in the receiving waters, although this was not the case in the receiving waters of Vineyard Creek.

Increasing dissolved oxygen concentration within influent as rainfall ingress into the sewer increases

Dissolved oxygen was lowest in dry-weather influent and increased in response to increasing rainwater ingress into the sewer (Table 4-2). In contrast, biological oxygen demand concentrations were observed to diminish under increasing rainwater ingress within the sewer (Table 4-2).

Influent and receiving water samples collected during wet-weather conditions had both dissolved oxygen concentrations and percent saturation above ANZECC/ARMCANZ (2000) recommended thresholds of > 6 mg/L or > 80-90% saturation (Table 4-2). Conversely, results from dry-weather and intermediate wet-weather were well below these thresholds (Table 4-2).

Table 4-2: Measurements of dissolved oxygen and biological oxygen demand in water samples collected for toxicity testing events

	Dry-weather	Intermediate wet-weather	Wet-weather					
	January 21, 2021	May 6, 2021	October 9, 2022		March 1, 2022		February 22, 2022	
	VC	VC	VC-influent	VC-ds	BC-influent	BC-ds	DMC-influent	DMC-ds
Dissolved oxygen saturation (%)	18	20	90	92	90	99	96	99
Field dissolved oxygen (mg/L)	1.6	1.9	9.5	8.3	7.8	8.7	8.5	8.8
$CBOD_5$ (mg/L)	330	99	8	9	10	3	11	2
BOD_5 (mg/L)	400	110	11	9	15	5	22	3

VC = Vineyard sewer carrier; VC-ds = Vineyard Creek downstream receiving waters
BC = Buffalo sewer carrier; BC-ds = Buffalo Creek downstream receiving waters
DMC = Darling Mills sewer carrier; DMC-ds = Darling Mills Creek system downstream receiving waters

Increasing metals concentration within influent as rainfall ingress into the sewer increased

A trend of increasing metal (such as copper, lead, iron) concentrations was apparent across influent samples collected from dry-, intermediate wet-weather and wet-weather. This likely reflects the settling of particulate contaminants in the sewer during dry-weather days causes siltation within the network of sewer pipes (Shi et al., 2018). During wet-weather events, large quantities of contaminants that have settled and accumulated in sewer pipes can be resuspended, leading to pulses of contaminants discharged to receiving waters (Wang et al., 2011). Botturi et al. (2021), in a study of a CSO, suggested that particulate-bound metals (for example lead and copper) were mostly caused by in-sewer deposit erosion in the influent discharged. A further description from the literature of bedload accumulation and erosion is contained within Kumar et al. (submitted).

An illustration of change in total metal concentrations from bedload erosion is provided by a WWO event tracked across three consecutive days from July 26-28, 2020, along with context afforded by the preceding single-day WWO event of July 14, 2020. This event was tracked from a trunk sewer main (SWSOOS at Mill Stream) of the Sydney region. Total metal concentrations of copper, zinc and lead measured on that earlier July 14 event were similar to those documented for July 26, the first day of the multi-day sampling event (see Table S7 of Kumar et al., submitted). A diminishing trend in total metal concentrations was observed across the following two consecutive days (July 27-28, 2020). A complementary diminishing trend was recorded across these three days from suspended solids and field turbidity lines of evidence.

A further illustration of change in total metal concentrations from bedload erosion is provided in the preceding influent water column study (Besley et al., 2023) with both total and dissolved metals of copper, zinc, lead and iron analysed from three days of the wet-weather period October 5-9, 2022 (Table S8 of Kumar et al., submitted) from the Vineyard trunk sewer main. Collection of influent samples on each of two days (October 5 and 6, 2022) coincided with an active WWO spill. Influent samples were collected on a third day of October 9, 2022, occurred after overflow discharges had recommenced on October 8, 2022. Increased dissolved concentrations of copper, lead and iron were evident for the third collection on October 9, 2022, when samples were collected for toxicity testing. These elevated metal concentrations of the toxicity testing water samples collected on October 9, 2022, are depicted in the summary of risk contained in Table 3 of Kumar et al. (submitted; and reproduced in the table of the graphical abstract Figure 4-6).

In this assessment of the Vineyard locations across the various weather conditions, the trace element concentrations were converted into Hazard Quotients (HQs). This is a measure used in ecological risk assessments to evaluate the potential risk of exposure of an organism to a particular substance, such as metals. HQs were calculated for each trace metal in terms of the 95% species protection default guideline values (Table 2 of Kumar et al., submitted). Hazard quotients for total metals will always be high as these are comparing the total concentration against a dissolved metal guideline value. However, the bioavailability of the particulate metal contribution is unknown, so the HQ only indicates the presence of high particulate metals where the value is high.

The contribution to toxicity of bioavailable concentrations of copper and zinc may be minimal. Kumar et al. (submitted) stated ‘During long rainfall events, both the dissolved and total metals

decrease as a function of time. This is largely due to a first flush effect where a largely stagnant system flows as a function of an incipient flow. Often the peak in particulate metal input occurs slightly after the dissolved metal peak. This less distinct change in dissolved metal concentrations may reflect that metals in bedload sediments are likely to be present as insoluble sulfides and after release into the influent water column where metals undergo oxidation, but oxidised iron sulfide (FeS) quickly forms colloidal or precipitated iron (III) oxide-hydroxide ($Fe(OH)_3$) which can sequester metals again from the influent water column.' Besley et al. (2023) indicated that the bioavailable concentration of copper and zinc will be significantly reduced by the high concentration of dissolved organic carbon in wet-weather samples observed in influent of Vineyard, Buffalo and Darling Mills. While iron concentrations in influent samples also exceeded the guideline value, there is evidence that it is likely to be complexed by natural organic ligands and not necessarily bioavailable or toxic (Nagai et al. 2007).

Overall hazard assessment

In the present study, sewer influent during dry- and intermediate wet-weather showed acute toxicity (under two-day tests). However, samples of influent collected during wet-weather and from the receiving waters were non-toxic. Conversely, influent and receiving environment samples collected during wet weather, when assessed with chronic toxicity tests, exhibited low toxicity from seven- and eight-day tests. Through acute (48 h) TIE testing ammonia and metals were determined to be contributors to the toxicity.

Dilutions required to remove toxicity to the sensitive test species were determined for the intermediate wet-weather and wet-weather influent samples and wet-weather receiving water samples were calculated based on the eight-day chronic toxicity of the water flea (*C. dubia*). Those dilutions for the water flea were higher than that returned for the midge larva (*C. tepperi*) for both acute (two-day) and chronic (seven-day) tests (Table 1 of Kumar et al., submitted). The dilution for dry-weather was taken from chronic (72 h) microalga testing (*C. vulgaris*) as it was marginally higher (1 in 12) than that returned for the water flea chronic test (1 in 10: Table with Fig. 4. of Kumar et al., submitted) and much higher than for the midge larvae (1 in 3.3: Table 1 of Kumar et al., submitted) chronic test outcome.

Overall, the ecotoxicological methods and approaches used in this study have allowed evaluation of the change in influent concentrations within the sewer during rainfall ingress at this high frequency and volume WWO (with the atypical arrangement of three adjacent ERSs) at Vineyard. A similar diminishing trend was provided in a ANZG (2018) assessment of ammonia concentrations from influent collected under dry-weather, intermediate wet-weather and wet-weather (Besley et al., 2023) was also documented under the current toxicity testing pilot study. This diminishing trend was not evident from the hazard quotient assessment of metals. In contrast, assessment of metals detected remobilization of metals from sewer sediment bedload under prolonged wet-weather events. The contribution to toxicity of the metals copper and zinc may be ameliorated to some extent as outlined above. These study results suggest that metals are of secondary concern with a possible latent effect while the primary driver of adverse risk of ecological effects is from ammonia.

This study documented low toxicity in the receiving waters 140 m downstream of these three adjacent ERSs and contrasts with a previous study in the Sydney region by Bickford et al. (1999)

that employed bioassays on sewer influent and receiving water samples of similar or higher modelled volume and frequency WWOs. They were unable to draw conclusions on the impact of overflows on the toxicity of the downstream receiving waters, although they documented a similar outcome that sewage influent was toxic. A companion risk assessment in that study highlighted ammonia as a chemical of concern in WWOs, which has been supported by the WWOM studies of Besley et al. (2023) and Kumar et al. (submitted). For Bickford et al. (1999), their risk assessment identified another 12 contaminants of concern, with 85% of the load of those chemicals contained in stormwater (described further in Section 4.1). Among those 12 contaminants were three metals that included copper and zinc (along with silver). That work suggested secondary contributions were from WWOs, and our current WWOM study outcomes are potentially supportive of that earlier finding. In the WWOM studies, silver was returned at detection limits except for a few samples that were just above the detection limit possibly due to the demise of photographic laboratories.

Multiple day length of toxicity testing methods and relevance to episodic WWO spills durations

The existing direct toxicity assessment methods, designed for continuous discharges like wastewater treatment plants, might not be suitable for episodic (pulsed) discharges such as WWOs. Assessing the environmental impacts of these events demands toxicity test designs that are both environmentally relevant and scientifically sound. Current standardised chronic testing protocols, over several days, do not adequately represent short-duration episodic discharges (of 24 hours or less duration representing 69% of the 54 WWO events recorded at Vineyard Creek) (Table 4-3). Further support for almost all WWO spill durations occurring within 24 h or less across another 84 gauged ERSs of the WWOM is contained in Section 4.5.2.

To better represent short-duration episodic WWO spills, an alternative approach was explored under toxicity testing of Buffalo and Darling Mills locations (Section 4.5.3). This involved pulsed exposures of WWO rain-diluted influent and receiving-water samples collected under wet-weather conditions with transfer of test organisms to clean water for the rest of the test duration, thereby offering a more representative evaluation of biological effects.

Summary and recommendation

Based on chronic testing toxicity results, the highest dilutions of sewer influent (untreated raw wastewater from the sanitary sewer system) required to remove toxicity to sensitive biota ranged from 1 in 12 in dry weather to 1 in 2.8 in wet weather when inflow and infiltration was at high levels into the sewer, Dilutions diminished to 1 in 1.8 in receiving waters measured at 140 m from the WWO in wet-weather conditions. In contrast, acute toxicity testing results indicated wet-weather influent samples were non-toxic and lower dilutions were returned for dry-weather and intermediate weather influent samples to remove toxicity. These outcomes indicate chronic toxicity testing conducted over multiple days provides conservative estimates of the dilutions required to remove toxicity. However, these multiple day tests do not adequately represent the 70% of the 54 WWO events recorded at the Vineyard study location that had spill durations of 24 h or less. That is, those chronic toxicity tests did not effectively mimic the short pulse durations of the most frequent WWOs recorded as spilling into receiving waters of Vineyard Creek.

Toxicity identification and evaluation testing determined ammonia and the metals of copper and zinc to be contaminants of potential concern. Ammonia concentration results evaluated against the ANZG (2018) guideline value for protection of 95% of species best represented toxicity testing outcomes with risk of adverse ecological effect reduced under increased dilution from inflow and infiltration of rainwater ingress into the sewer system. Copper and zinc appeared to be of secondary concern as hazard quotient assessment depicted sewer bedload mobilization under wet-weather rain ingress conditions but did not mirror either the overall chronic or acute toxicity testing outcomes.

Recommendation

Given that WWOs frequently spill for less than 24 h, while our chronic toxicity tests extend for 8 days, it may be prudent to consider testing pulsed exposures and using time-averaged concentrations to more correctly evaluate the risks over extended periods. It is recommended to better represent short-duration episodic WWO spills, an alternative approach be explored under toxicity testing of Buffalo and Darling Mills locations (Section 4.5.3). That approach would evaluate pulsed exposures of WWO spills of rainwater-ingress diluted influent and receiving water samples collected under wet-weather conditions with transfer of test organisms to clean water for the rest of the test duration. Both six- and 24 h durations are suggested to be evaluated as these represent the majority of short WWO durations (as determined from assessment in Section 4.5.2).

4.5.2 Durations of WWO spills

The durations of WWO spills were evaluated in response to the considerations raised in Kumar et al. (submitted) for overflow spill duration versus the duration of chronic toxicity testing protocols run over three, seven or eight days.

A summary of the percentage of WWO event spill durations across five periods (Table 4-3) and associated rainfall for each WWO event has been drawn from the seven ERSs studied across Besley and Cassidy (2022 Supplementary Table S1), Kumar et al. (submitted, Supplementary Table S1) and Kumar et al. (submitted 2, Supplementary Table S1). These five periods represent toxicity testing measurement endpoints. While these seven ERSs represent a cross section of low, medium, and high volume and frequency WWOs (based upon hydraulic modelling results listed in those papers).

Tabulation of WWO event spill duration (from gauging data collected as part of the WWOM) against five toxicity testing measurement endpoints for each of these seven ERSs (Table 4-3, Figure 4-8) indicated that WWO spill durations of 24 h or less comprised from 68% to 100% of WWO events across these seven ERS locations, while spill durations of less than 6 h ranged from 43% to 69% of WWO events with the exception of the Vineyard location WWO (Table 4-3, Figure 4-8).

Table 4-3: Percent of WWO event spill durations falling within five periods

ERS location (no. of gauged overflows)	< 6 h	6 to 24 h	> 24 to 48 h	> 48 to 96 h	> 96 to 168 h
Vineyard Creek* (54)	28%	41%	13%	9%	9%
Darling Mills Creek* (31)	45%	39%	10%	6%	0%
Buffalo Creek* (48)	69%	17%	6%	6%	2%
Salt Pan Creek (30)**	63%	30%	0%	7%	0%
The Ponds Creek (16)**	43%	25%	13%	13%	6%
Hunts Creek (12)**	92%	0%	8%	0%	0%
Mill Creek (2)**	100%	0%	0%	0%	0%

* gauged overflow events between October 1, 2018 and October 31, 2022
** gauged overflow events between August 1, 2018 and March 31, 2021

Companion plots of spill durations of these WWO events by ERS location illustrated within these plots (Figures 4.9 to 4.15) were the dominance of overflows with 24-h or less duration including the overflow events sampled for the Darling Mills (Figure 4-10) and Buffalo (Figure 4-11) wet-weather toxicity testing events. While the overflow duration of the Vineyard wet-weather toxicity testing event was between 24 to 48 h (Figure 4-9). The rainfall overlay on these plots illustrated that longer duration overflow events typically occurred during storm events with relatively higher rainfall totals (Figure 4-9 to Figure 4-15).

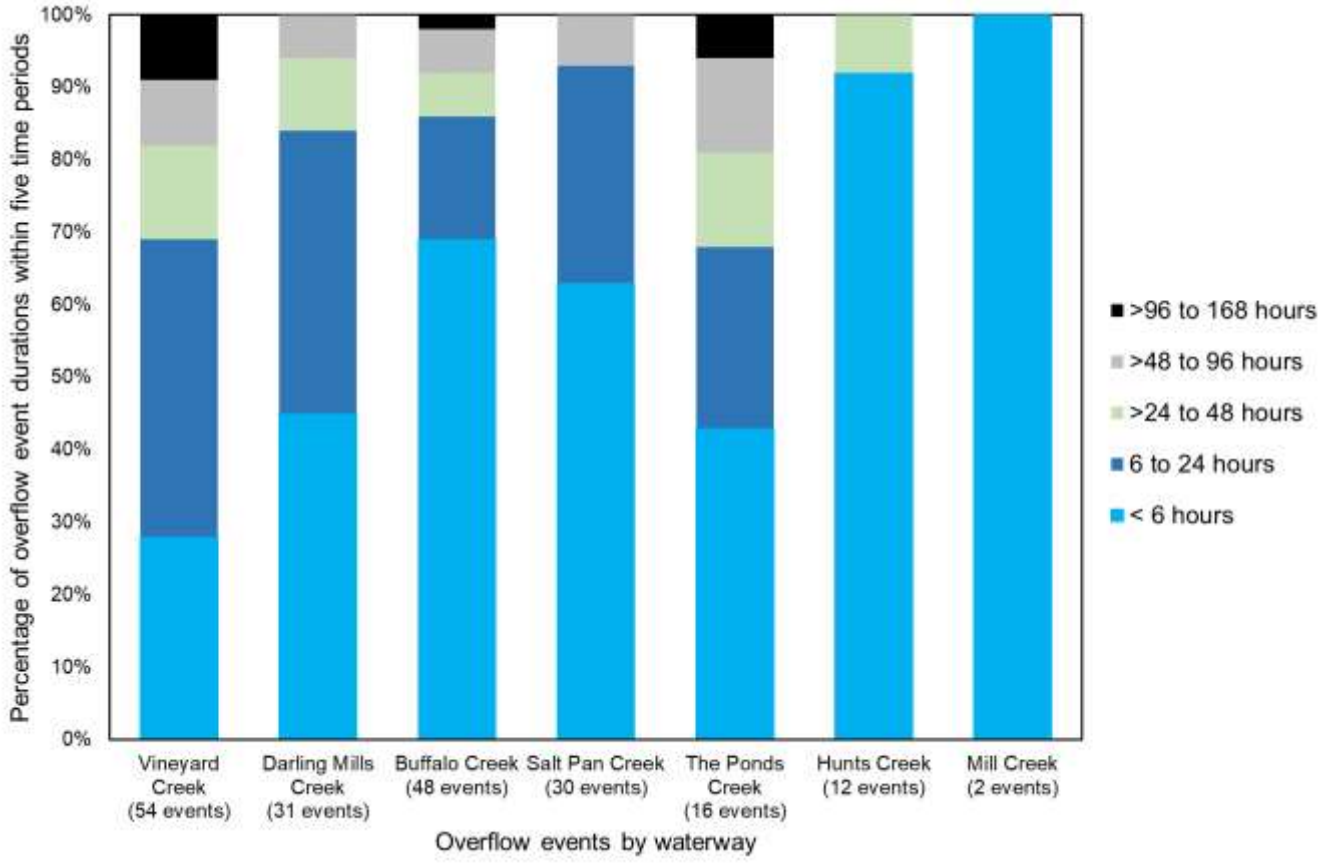


Figure 4-8: Percentage of WWO event spill durations falling within five periods across seven ERSs

Analysis of WWO event spill durations from another 31 gauged ERSs that spilt into freshwater streams between October 31, 2018 and March 31, 2021 indicated that WWO spill durations of 24 h or less comprised 100% of WWO events. While WWO spill durations of less than 6 h comprised 63% (364 of 579) of WWO events. Similar results were returned from analysis of another 47 ERSs that spill into estuarine waterways with WWO spill durations of 24 h or less comprised 97% (509 of 522) of WWO events, while WWO spill durations of less than 6 h comprised 69% (361 of 522) of WWO events.

This suggests that toxicity testing outcomes measured at 24 h are potentially the most relevant toxicity testing measurement time-point for water column samples. Longer duration overflow events were typically associated with higher rainfall totals that potentially contributed to greater rainwater ingress and dilution of influent within the sanitary sewer pipe network before overflowing

to a receiving waterway. As these more dilute conditions, particularly for the receiving waters, were not mimicked by toxicity testing, the measurement endpoints beyond 24 h potentially afford less realistic assessments to the recorded performance of the sewerage system of the Sydney region.

The ERS that spills into Mill Creek is an example of a low spilling volume and frequency WWO (Figure 4-15) that was documented as only occurring during major storm events under east coast low conditions with over 300 mm of rainfall (Besley and Cassidy, 2022). Other examples of low volume and frequency WWOs were revealed during an intense brief rainfall event on February 6, 2023, when all seven gauged ERSs (Figure 1, Besley et al., 2023) in the Gymea Bay sewer catchment spilled for 2.7 h in response to an intense storm event that deposited 30 mm of rainfall in 1 h. The only other occasion when these seven ERSs spilled at the same time within the study period occurred during the February 2020 east coast low event. More typically ERS spill activity during rainfall events in the Gymea Bay sewer catchment were illustrated by two other WWO spill events documented in Besley et al. (2023). An event on October 6, 2022, recorded three spatially separated ERSs as spilling. These ERSs commenced spilling within 15 min of each other and continued for 1 h to 2.5 h. Another wet-weather overflow event was sampled on January 6, 2023, after a single ERS spilt for 4.25 h. The above examples illustrate that under protracted or very intense rainfall events these deliberately placed ERSs across the sewerage system perform their designed function to prevent inundation of properties and protect human health by spilling to receiving waters. While under rainfall events that fall between these two extremes, a reduced cohort of ERSs perform their designed function.

The WWOM project strategy to gauge a range of ERSs to represent a cross section of low, medium, and high volume and frequency WWOs (based upon hydraulic modelling) has provided a real-world insight into functioning of the Sydney sewerage system that was not available from the Bickford et al. (1999) review.

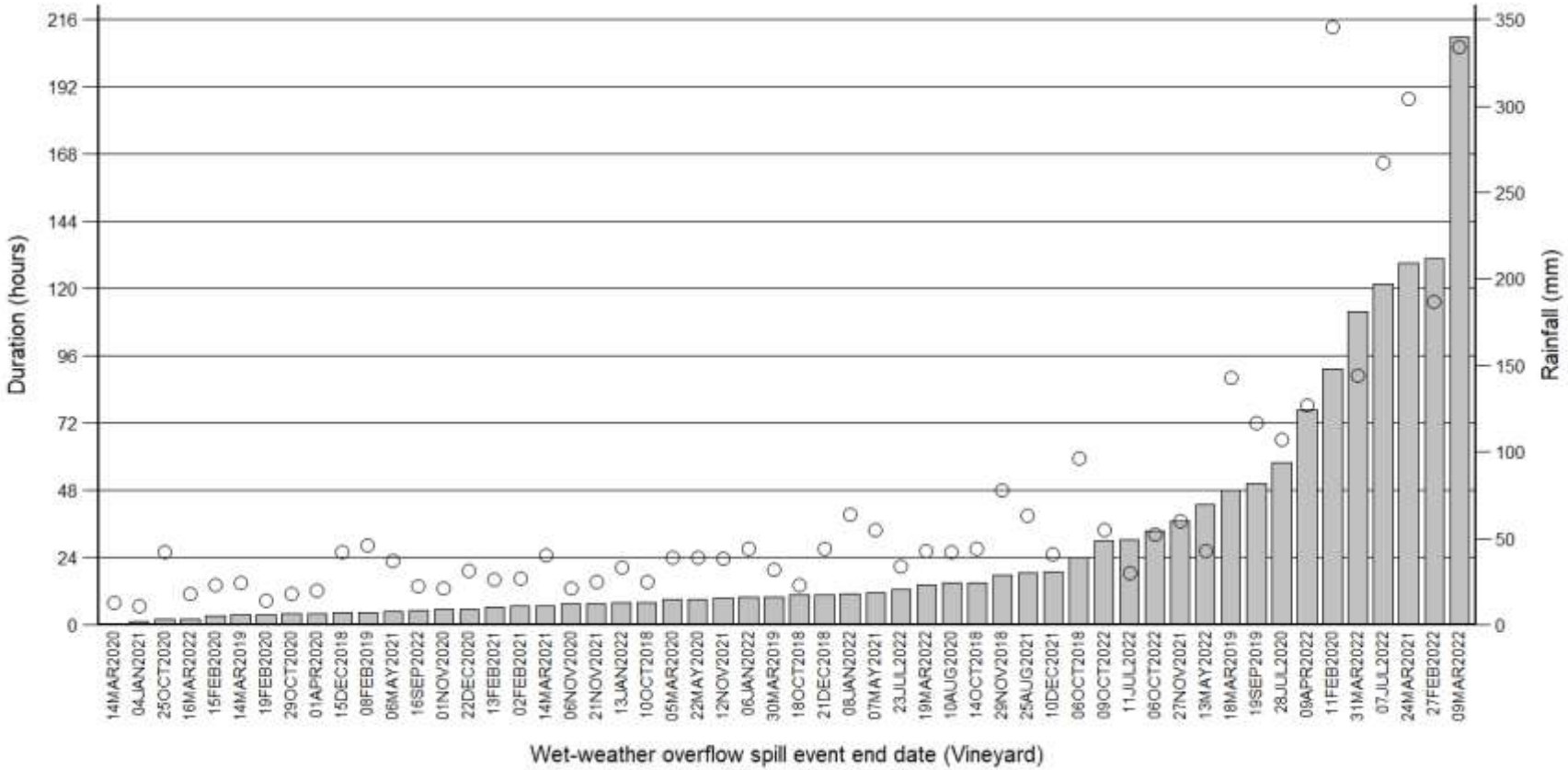


Figure 4-9: Duration of each measured (gauged) overflow spill event at the Vineyard study location between October 2018 and October 2022.

Corresponding rainfall displayed by open circles for the preceding 72 h to the end of each spill event, except for seven events on right-hand side of plot that spilled in excess of 72 h with the corresponding rainfall durations listed in Kumar et al. (submitted, Supplementary Table S1). The longest duration measured by gauging across the three ERSs is displayed for each event.

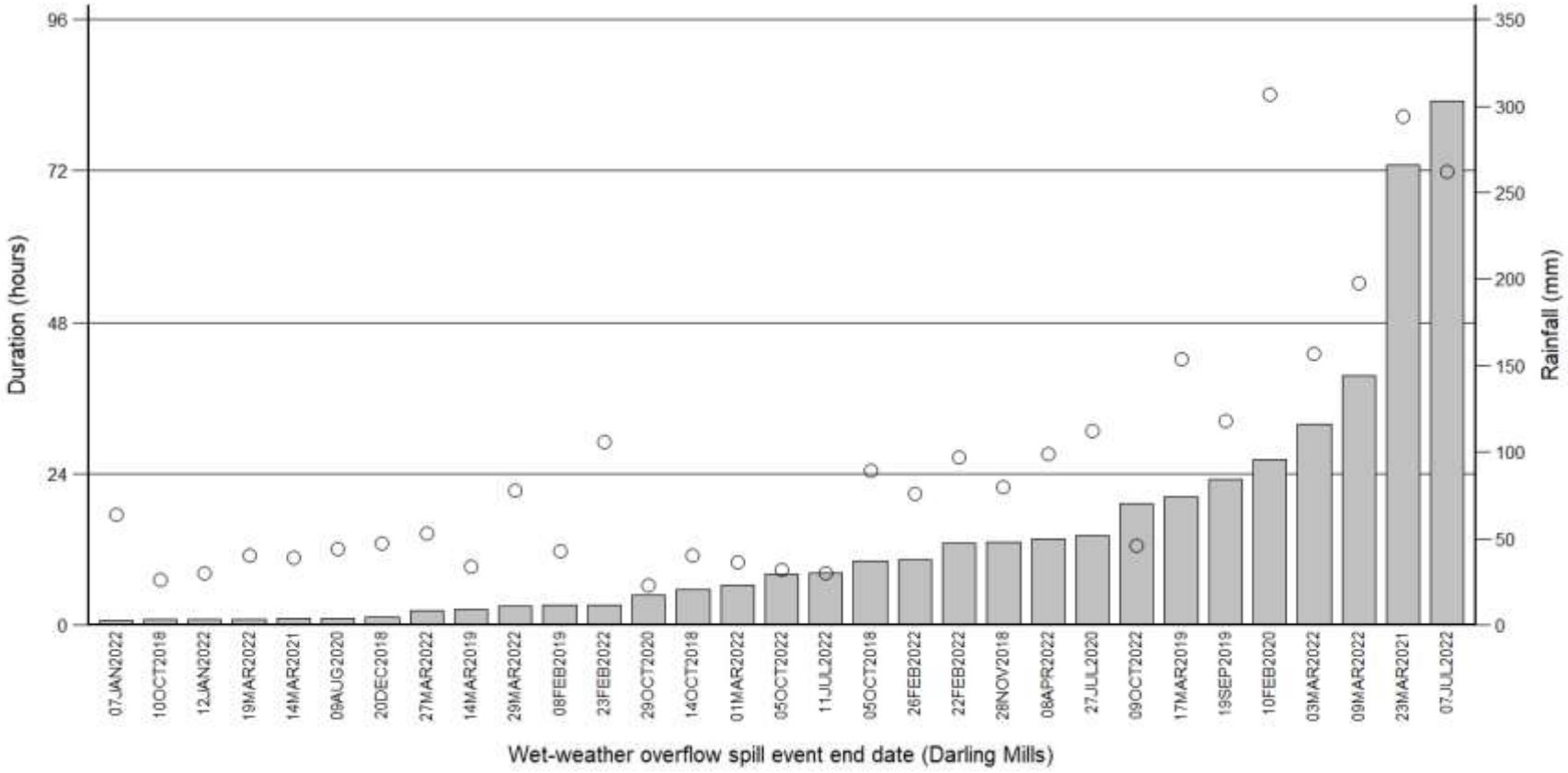


Figure 4-10: Duration of each measured (gauged) overflow spill event at the Darling Mills Creek study location between October 2018 and October 2022

Corresponding rainfall displayed by open circles for the preceding 72 h to the end of each spill event, except for two events on right-hand side of plot that spilled in excess of 72 h with the corresponding rainfall durations listed in Kumar et al. (in prep, Supplementary Table S1). The longest duration measured by gauging across the two ERSs is displayed for each event.

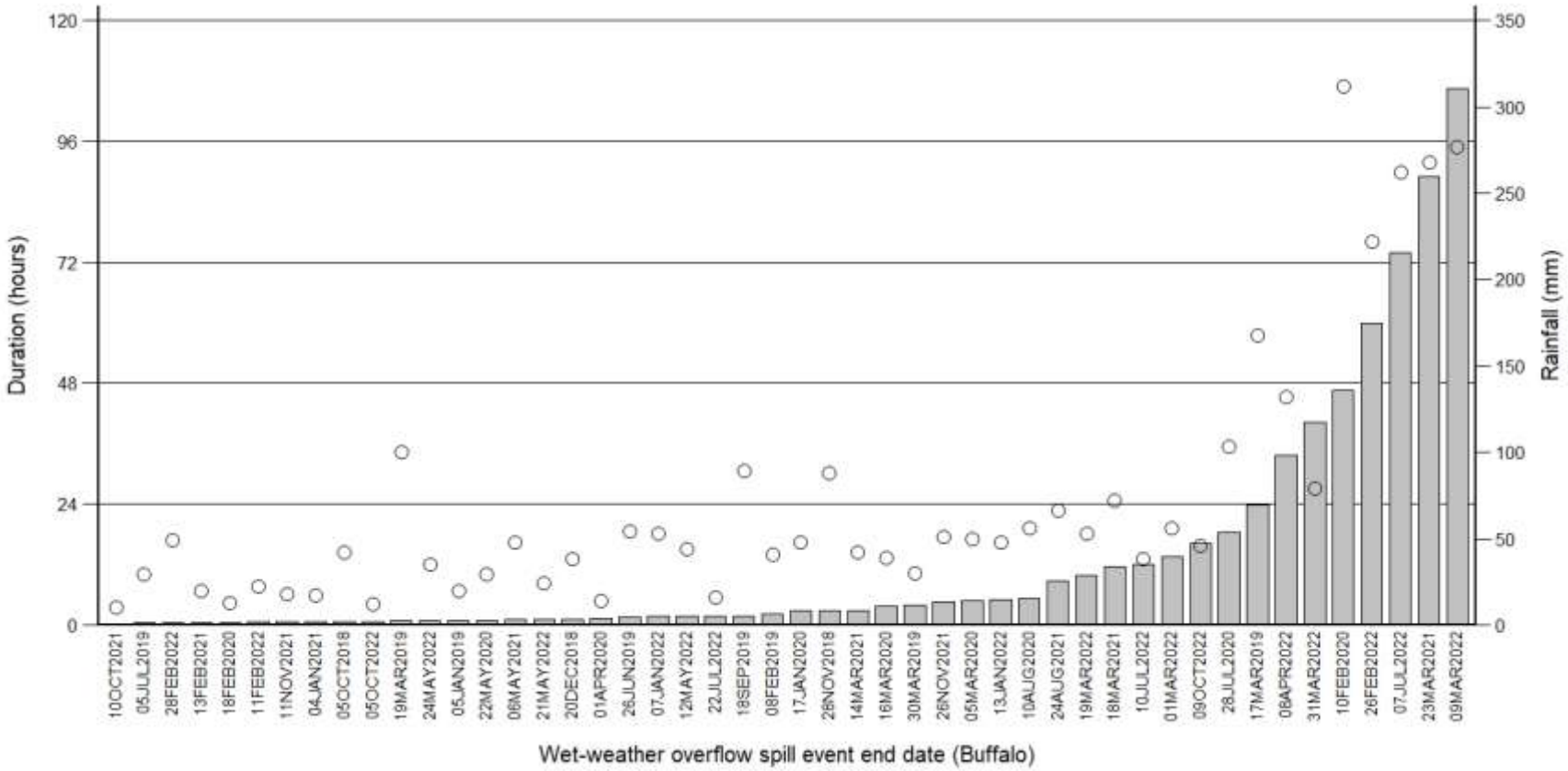


Figure 4-11: Duration of each measured (gauged) overflow spill event at the Buffalo Creek study location between October 2018 and October 2022

Corresponding rainfall displayed by open circles for the preceding 72 h to the end of each spill event, except for three events on right-hand side of plot that spilled in excess of 72 h with the corresponding rainfall durations listed in Kumar et al. (in prep, Supplementary Table S2). The longest duration measured by gauging across the two ERSs is displayed for each event.

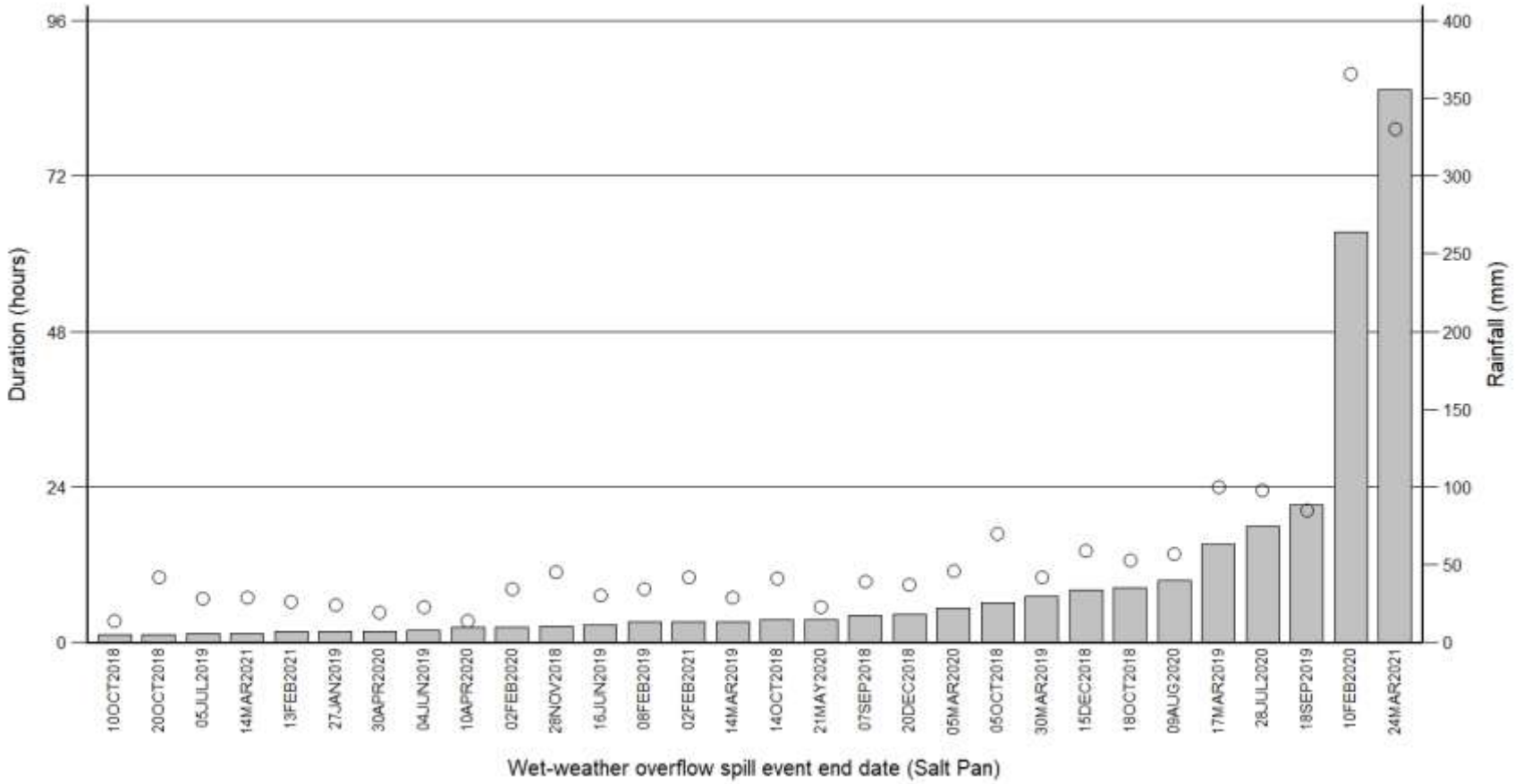


Figure 4-12: Duration of each measured (gauged) overflow spill event at the Salt Pan Creek study location between August 1, 2018 and March 31, 2021

Corresponding rainfall displayed by open circles for the preceding 72 h to the end of each spill event, except for the event on right-hand side of plot that spilled in excess of 72 h with the corresponding rainfall durations listed in Supplementary Table S1 of Besley and Cassidy (2022). Corresponding overflow volumes and spill rates are detailed in Besley and Cassidy (2022).

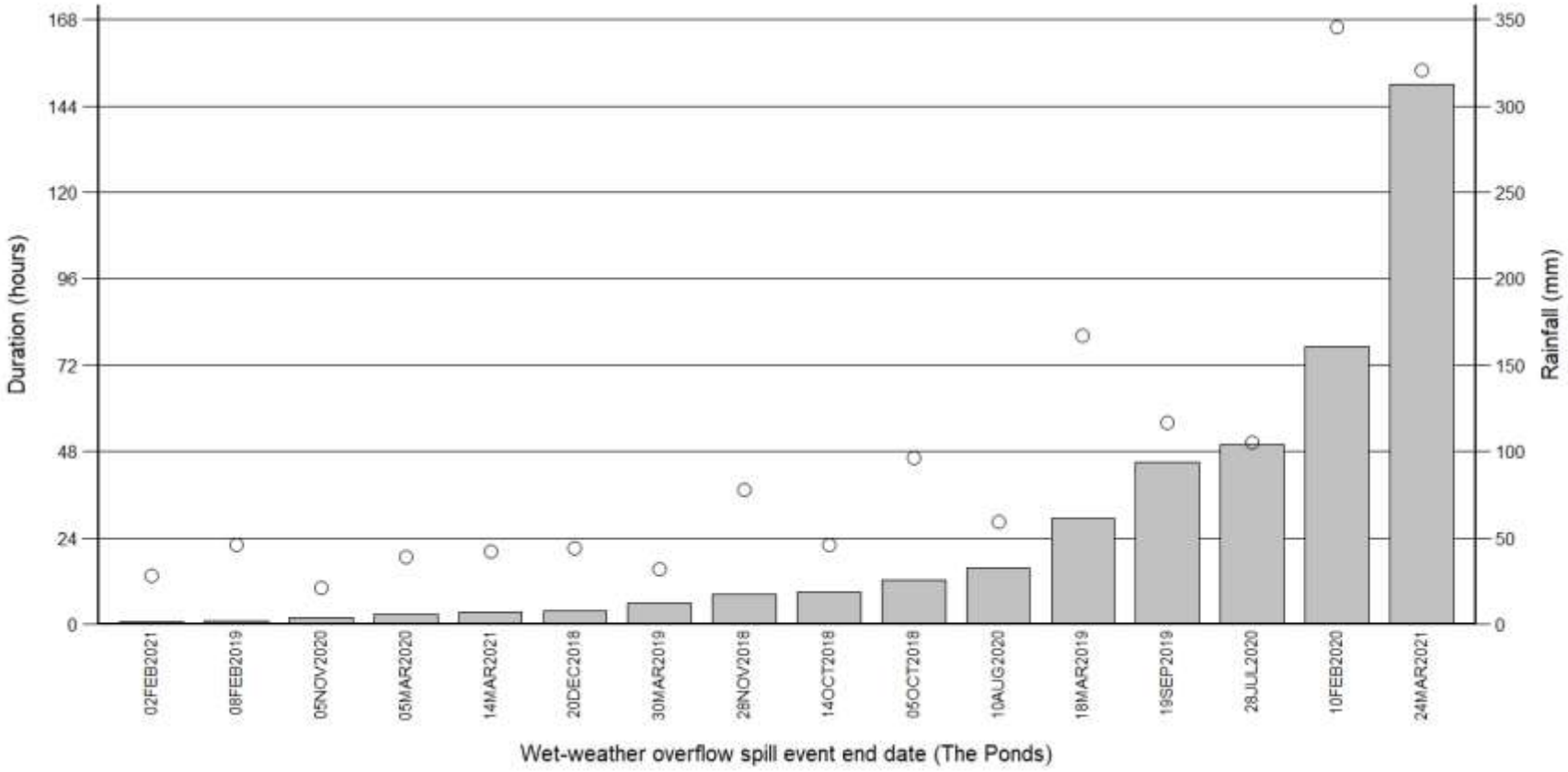


Figure 4-13: Duration of each measured (gauged) overflow spill event at The Ponds Creek study location between August 1, 2018 and March 31, 2021

Corresponding rainfall displayed by open circles for the preceding 72 h to the end of each spill event, except for two events on right-hand side of plot that spilled in excess of 72 h with those corresponding rainfall durations listed in Supplementary Table S1 of Besley and Cassidy (2022). Corresponding overflow volumes and spill rates are detailed in Besley and Cassidy (2022).

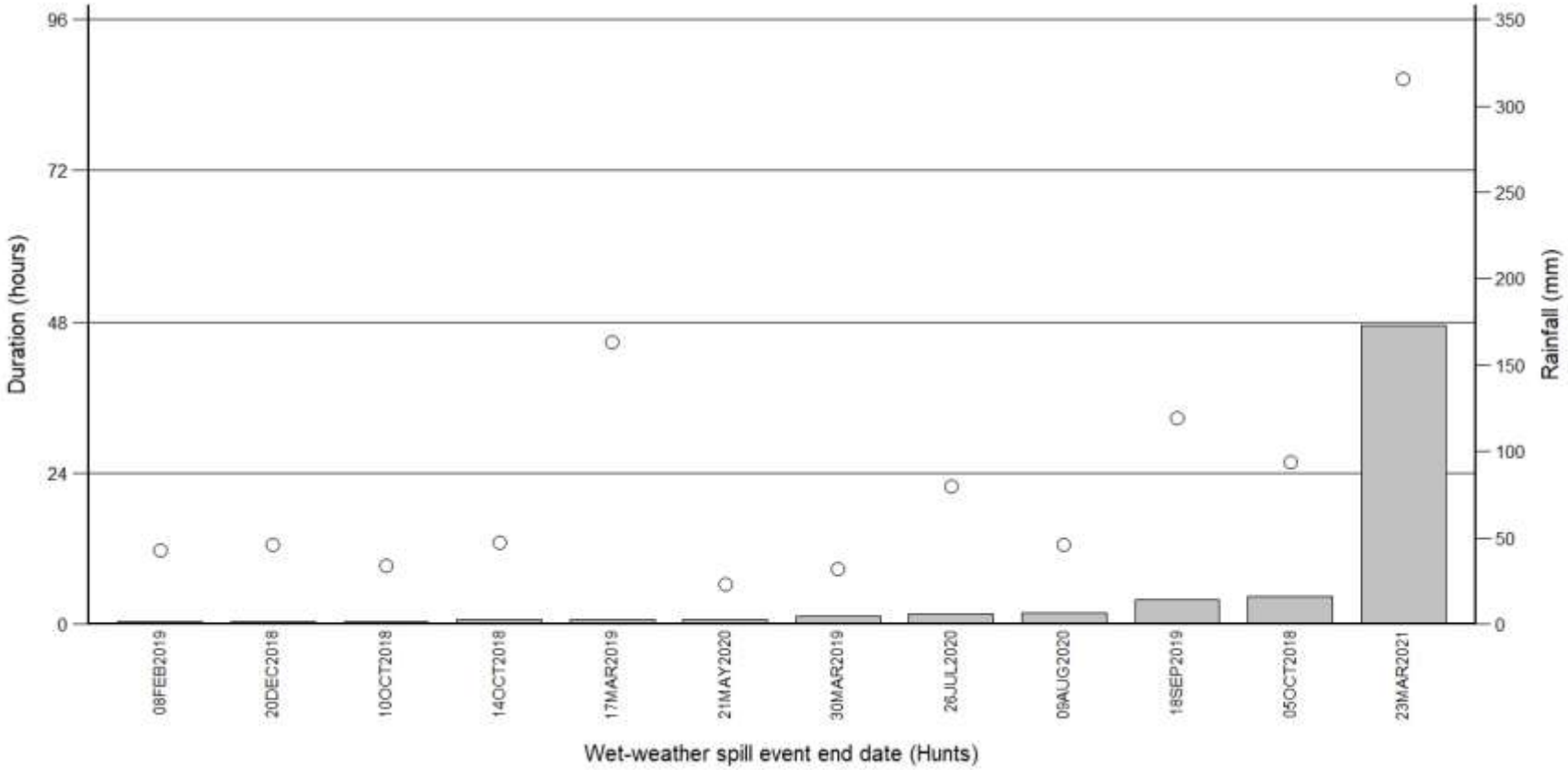


Figure 4-14: Duration of each measured (gauged) overflow spill event at Hunts Creek study location between August 1, 2018 and March 31, 2021

Corresponding rainfall displayed by open circles for the preceding 72 h to the end of each spill event. Corresponding overflow volumes and spill rates are detailed in Besley and Cassidy (2022).

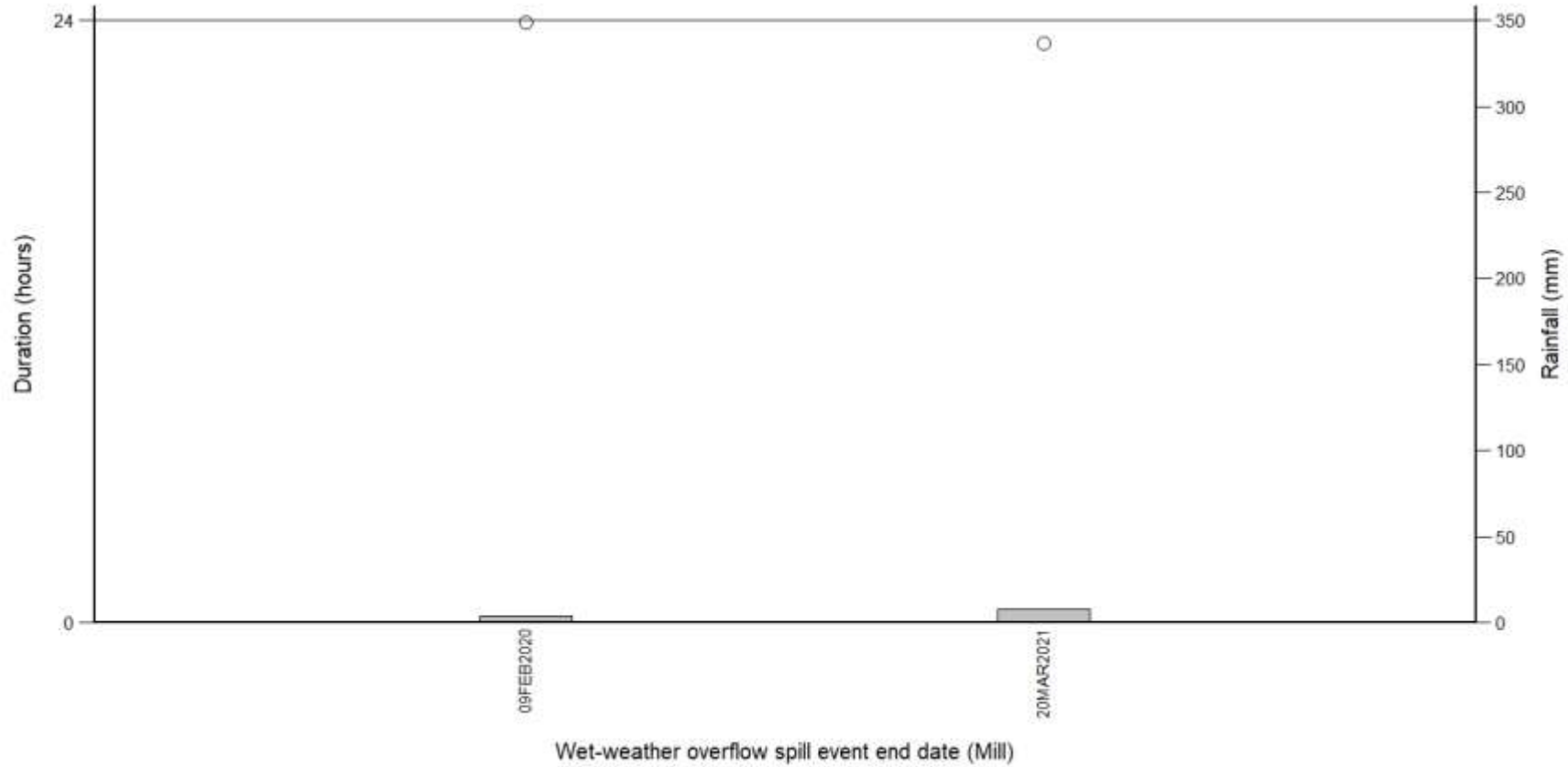


Figure 4-15: Duration of each measured (gauged) overflow spill event at Mill Creek study location between August 1, 2018 and March 31, 2021

Corresponding rainfall displayed by open circles for the preceding 72 h to the end of each spill event. Corresponding overflow volumes and spill rates are detailed in Besley and Cassidy (2022).

4.5.3 Toxicity studies of Darling Mills and Buffalo locations

Analysis undertaken in Section 4.5.2 on durations of WWO spills reflects a spill duration of six hours encompassing about 60% of WWOs, while WWO spill duration of 24 h or less accounts for almost all other WWOs. The few events with spill durations longer than 24 h were associated with protracted high rainfall total storm events occurring across a few to multiple days. The inclusion of 6 h and 24 h pulsed exposures into toxicity testing of the Buffalo and Darling Mills locations under wet-weather conditions was based upon the above analysis, and this inclusion investigates the recommendation arising from toxicity testing of the Vineyard location documented in Section 4.5.1 (Kumar et al., submitted).

Objectives

This study assessed the chronic toxicity to *Ceriodaphnia dubia* of wet-weather rain ingress diluted sewer influent along with water samples from the downstream receiving streams of Darling Mills Creek and Buffalo Creek (Figure 4-17), under scenarios using continuous and pulsed exposures. To assist with interpretation of toxicity results, concentrations of copper, zinc and ammonia were evaluated with a time-weighted average concentration approach as these chemicals were expected to cause toxicity to aquatic organisms based on earlier WWOM ecotoxicological investigations (Section 4.4, Besley et al., 2023; Section 4.5.1, Kumar et al., submitted).

A notable outcome of the preceding Vineyard toxicity study was the lack of acute toxicity in samples of wet-weather rain ingress diluted sewer influent and downstream receiving water samples collected during WWO spills (Section 4.5.1). Hence, the focus under the current investigation on chronic toxicity.

A graphical outline of this sub-study is provided in Figure 4-16. The water flea, *C. dubia* was used for conducting continuous and pulse chronic bioassays. These chronic tests with *C. dubia* estimated the influent concentration that interferes with growth. Continuous exposure of *C. dubia* to wet-weather influent and downstream receiving water samples was conducted over 192 h (8 days). Tests of pulsed exposure of *C. dubia* to wet-weather influent and to downstream receiving water samples were conducted for two periods of 6 h or 24 h. After each exposure period, *C. dubia* specimens were transferred to water that did not contain diluted influent. To account for potential handling stress, organisms in the control group were also moved to new control beakers following the same schedule. A fuller description of this process is detailed within Kumar et al. (submitted).

Text and citations in this section were drawn from the peer-reviewed publication: Kumar, A., Batley, G.E., Nguyen, T.V., Nguyen, T., Cassidy, M., Besley, C.H. submitted. Pulsed versus continuous exposures to evaluate the toxicity of sanitary sewer wet-weather overflows.

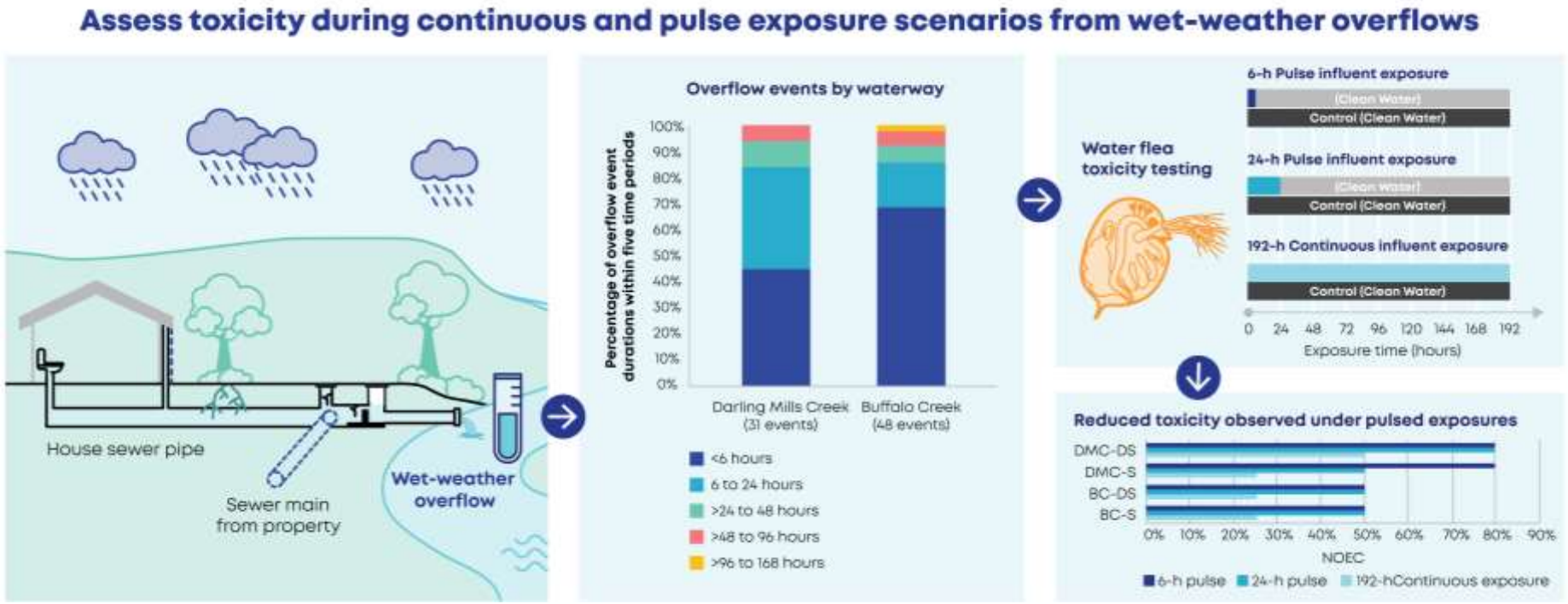


Figure 4-16: Graphical abstract for the toxicity testing that evaluated pulsed and continuous exposure of water fleas to wet-weather rain diluted sewer influent and to water samples from downstream receiving streams

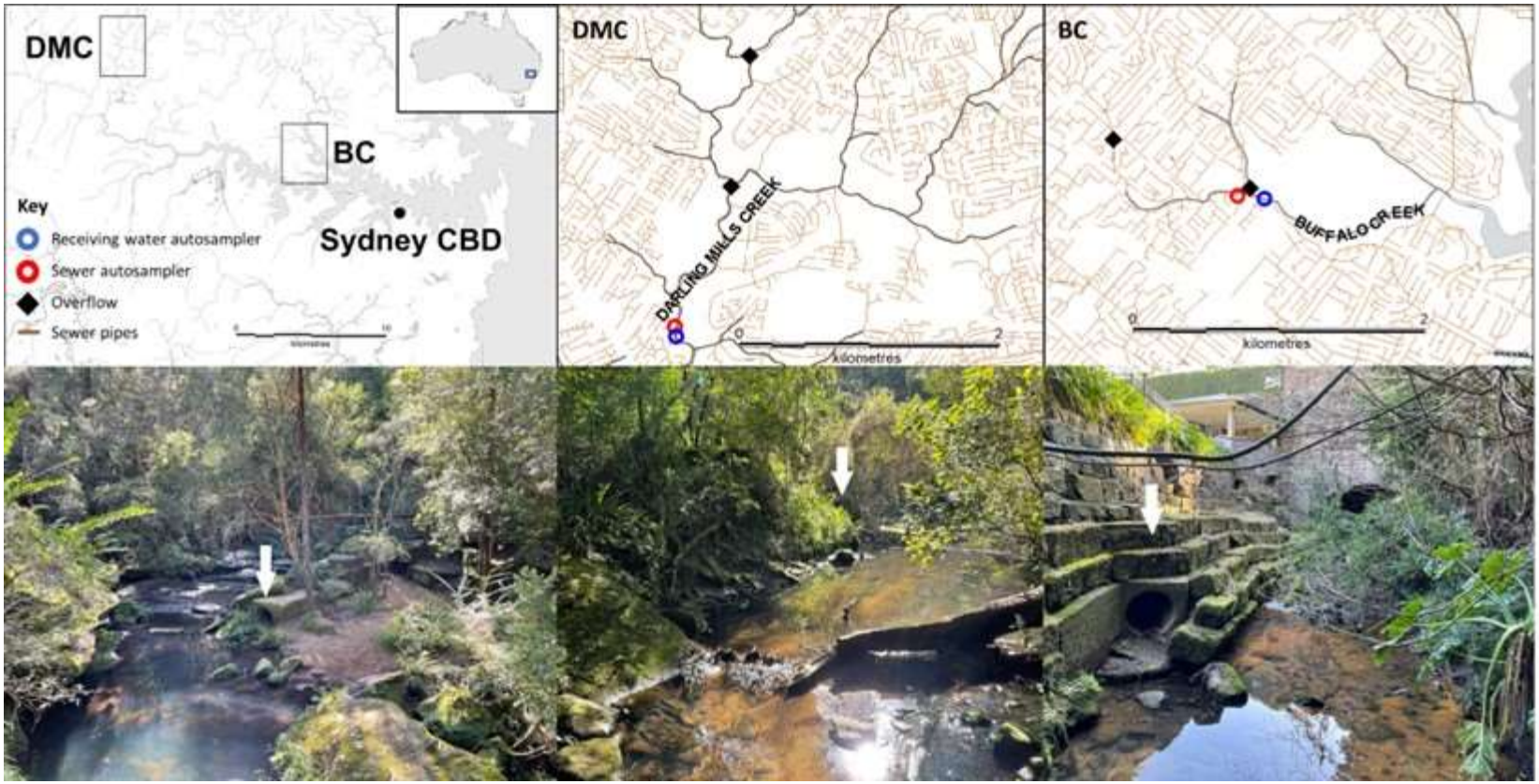


Figure 4-17: Overall location map and detail of Darling Mills Creek (DMC) and Buffalo Creek (BC) study locations

Red circles represent sewer autosamplers, blue circles represent downstream receiving water autosamplers, black diamonds indicate the position in sewer catchments of gauged ERS (sewer overflow points), and brown lines represent the sewer trunk and reticulation pipes of these urban study locations. Photograph at bottom of figure was taken under dry weather: left hand photograph of DMC ERS 2 km upstream of autosampler; middle photograph of DMC ERS 1 km upstream of autosampler; and right-hand photograph of gravity fed BC ERS 50 m upstream of autosampler.

Continuous exposure versus pulsed exposure toxicity test outcomes and dilutions required to remove toxicity

The chronic toxicity test of the water flea (*C. dubia*) under 192 h continuous exposure to wet-weather influent and downstream samples exhibited greatest toxicity no-observable-effect concentrations (NOEC) when contrasted with the 24 h and 6 h pulse exposures (Table 4-4). The NOEC represents the highest test concentration at which there is no statistically significant reduction in growth.

Dilutions for the continuous or pulse exposures necessary to remove toxicity were calculated based on the overall toxicity to *C. dubia.* The dilutions required to remove toxicity during continuous exposure to the sewer influent was 1 in 4 for sewer influent samples from both the Buffalo and Darling Mills Creek locations. In contrast, the 24 h and 6 h pulse exposures resulted in reduction of the required dilutions for greater than 1 in 2 for *C. dubia* (Table 4-4).

Table 4-4: The calculated no-observed effect concentrations (NOEC) and dilutions required to remove toxicity were based on *Ceriodaphnia dubia* chronic tests of continuous or pulse exposures

Sample	192 h continuous exposure		24 h pulse exposure		6 h pulse exposure	
	NOEC	Dilution	NOEC	Dilution	NOEC	Dilution
BC influent	25%	1 in 4	50%	1 in 2	50%	1 in 2
BC downstream	25%	1 in 4	50%	1 in 2	50%	1 in 2
DMC influent	25%	1 in 4	50%	1 in 2	80%	1 in 1.25
DMC downstream	50%	1 in 2	80%	1 in 1.25	80%	1 in 1.25

BC = Buffalo Creek; DMC = Darling Mills Creek; Dilutions required to remove toxicity is 100/NOEC
A higher NOEC value indicates less toxic effect

Context of continuous exposure dilutions required to remove toxicity

The percentile rankings of overflow volume for each of the three wet-weather events were relatively similar over the period October 1, 2018 to October 31, 2022. The wet-weather event collected from the Vineyard Creek location on October 9, 2022, had an 83rd percentile rank (Kumar et al., in preparation), with a 79th percentile rank calculated for the event sampled March 1, 2022, at the Buffalo location, and a 70th percentile rank was determined for the Darling Mills location sampled on February 22, 2022. This suggests that toxicity testing across these three wet-weather events had assessed reasonably similar rainwater ingress diluted influent conditions across these three different sections of the sewerage network.

In contrast, the May 6, 2021, event collected from the Vineyard location under intermediate wet-weather conditions had a 25th percentile ranking with notably less dilute chemical concentrations present in the influent. For example, ammonia was measured at 15 mg total ammonia N/L for the May 6, 2021, event whereas ammonia was measured at 1.1, 1.9 and 0.8 mg total ammonia N/L for

the three higher percentile ranked wet-weather events collected from Vineyard (October 9, 2022), Darling Mills (February 22, 2022) and Buffalo (March 1, 2022), respectively (Besley et al., 2023). This suggests the comparison of dilution outcomes from these three locations is a potentially valid line of evidence as the sampled wet-weather events considering the similar percentile ranks outlined above.

The previous study results from Kumar et al. (submitted, Section 4.5.1) defined the dilutions required to remove toxicity for the continuous exposures of *C. dubia* to Vineyard sewer influent from samples collected under wet-weather rain ingress diluted conditions of 1 in 2.8, and the dilution required reduced to 1 in 1.8 in downstream (140 m) receiving water samples. As *C. dubia* was assessed in both studies for the same continuous exposure duration of 192 h, these similar dilutions required to remove toxicity documented in the Vineyard study are supportive of the continuous exposure dilutions from the current study. This in turn suggests that the dilutions required to remove toxicity obtained from the 6 h and 24 h pulse exposures provide realistic scenarios as part of informing continuous improvement of the Sydney Water risk prioritisation tool in management of WWOs.

Time-weighted average concentration (TAC)

To further evaluate *C. dubia* pulse data, time-averaged concentrations (TACs) were raised for chemical data (ammonia, copper and zinc). In its simplest form this means converting the toxic response measure to an average concentration over the 8-day test duration. This would mean that for pulses of 6 and 24 h, the TACs would be respectively 3 and 12.5% of the continuous 8-day chronic endpoint (NOEC, EC10).

Ammonia TACs were higher at 192 h, exceeding the ANZG (2018) default guideline value (DGV) only for the two sewer sites. The ammonia levels at downstream sites were below DGVs at 192 h, 6 h, and 24 h pulse exposures (Table 4-5).

Table 4-5: Time-averaged concentrations of ammonia during continuous or pulse exposures

Locations	192 h continuous exposure	24 h pulse exposure	6 h pulse exposure
Buffalo Creek sewer	**0.82** ±0.41	0.10	0.03
Buffalo Creek downstream	0.15±0.04	0.019	0.005
Darling Mills Creek sewer	**2.79** ±0.44	0.35	0.09
Darling Mills Creek downstream	0.03±0.01	0.003	0.001

Bold font represents concentration exceeds DGV 0.79 mg N/L

When copper and zinc pulse exposure concentrations were represented using the time-weighted average concentration (TAC), the concentrations in pulse exposures were lower than observed in continuous exposures to the same metal (Table 4-6). The sewer sites exhibited higher concentrations than their respective downstream sites. Additionally, TACs of both copper and zinc

at 192 h exceeded Australian and New Zealand default guideline values (DGVs) (ANZG, 2018), with the Buffalo Creek sewer influent sample reporting the highest concentrations. The TACs of both copper and zinc were below DGVs during 6 h and 24 h pulse exposures, with the exception of Darling Mills sewer influent sample reporting zinc at 12.8 µg/L, which was 1.5× higher than the DGV, and at the Buffalo sewer copper was at the DGV concentration for the 24 h pulse (Table 4-6).

Exhibited in Figure 4-16 and Table 4-6 are lower ammonia, copper and zinc concentrations for 24 h and 6 h pulse exposures. These samples had reduced toxic effects, as illustrated by the corresponding NOEC values ranging between 50 - 80% (Table 4-4).

A discussion within Kumar et al. (submitted) details other studies employing TACs, however, no study of WWOs was identified within that literature. The most relevant study investigated stormwater runoff with *C. dubia* for pulse exposures (Karic et al., 2022). Although, only on acute responses of 6, 12 and 24 h which were compared with continuous 96 h acute exposures. That study showed that whole effluent testing after 96 h over-estimated toxicity. That premise was also the premise underlying the WWOM toxicity studies conducted on samples of the Buffalo and Darling Mills locations. Hence, Kumar et al. (submitted) represents the first study of WWOs with pulse exposure toxicity testing.

Table 4-6: Time-averaged concentrations of copper and zinc during continuous or pulse exposures

Locations	Copper (µg/L)			Zinc (µg/L)		
	192 h*	24 h	6 h	192 h*	24 h	6 h
Buffalo Creek sewer	**11.1** ± 1.4	**1.4**	0.3	**23.6** ±2.9	2.9	0.7
Buffalo Creek downstream	**6.3** ± 1.2	0.8	0.2	**14.8** ±3.3	1.9	0.5
Darling Mills Creek sewer	**4.1** ±0.7	0.5	0.1	**102** ±23.5	**12.8**	3.2
Darling Mills Creek downstream	**2.8** ±0.6	0.4	0.1	**32.8** ±6.4	4.1	1.03

Concentrations as an average of 5-7 samples. **Bold font** values are above DGVs; Cu 1.4 µg/L; Zn 8 µg/L

Overall hazard assessment

The hazard quotients (HQ) for copper and zinc based on TACs for both 6 h and 24 h pulse scenarios were below 1, except for zinc in the Darling Mills sewer influent sample for the 24 h pulse scenario (with a HQ of 1.6). A HQ below 1 is perceived as low. In contrast all chronic exposure HQs were well above 1 (ranging from 1.8 to 12.8).

When HQs from these two metals were combined into a hazard quotient index (HQI) both sewer influent sample scenarios for 24 h pulses exceed a HQI of 1 (with 1.4 for Buffalo influent and 2.0 for Darling Mills influent), whereas the 6 h pulse scenarios were all below 1.

The outcome of the overall hazard assessment based on TACs suggests that WWO spill durations of 24 h or less into receiving waterways where adequate dilution occurs was evident at these two

study locations (Section 4.4), are scenarios under which adverse ecological effects are least likely to manifest.

Revisiting loading of metals in stormwater versus loading from WWOs

In an earlier study in the Sydney region, Bickford et al. (1999), identified a chemical loading of 85% in stormwater for contaminants of concern that included copper, zinc and silver (described further in Section 4.1). A subsequent investigation of highly urbanised sub-catchments of the Sydney estuary (Davis and Birch, 2009) showed that the contaminants copper, lead and zinc were predominantly (79-87%) derived from diffuse sources (residential properties and roads). While Birch (2024) recently identified road-derived metals as the chief contributor of metals to stormwater from a review and critical assessment of over three decades of research supplemented by global studies. Birch (2024) states an important implication is that road-derived material is highly transportable by fluvial and aeolian mechanisms active on street margins and road-derived metal is easily and directly dispersed to the stormwater system via gutters, gully pots and drains. These studies suggest that the dominant metal loading is delivered by stormwater inflows with minor secondary contributions delivered from WWO spills. Support for a lesser contribution of metals from WWOs is provided by the two former largest (by volume) WWOs that spilt into Sydney Harbour from an ERS at Long Bay (3689 ML/year) and from another ERS in the Lane Cove River (2740 ML/year), where an absence of observable enrichment in heavy metals adjacent to these ERSs indicated that WWOs do not contribute significant metallic contaminants to this estuarine system (described in Section 4.1). Since 2001, under all but extreme wet-weather conditions, spills from these two ERSs are now captured by the Northside Storage Tunnel (Figure 3-2). The smaller contribution from WWOs of metals is also potentially supported by this current study (Kumar et al., submitted), which has informed our understanding of the pulsed and relatively temporally brief periods of WWO spills. During a WWO, dissolved metals contained within the water column along with metals in the bedload sediment within the sewerage network at that point in time are released. Whereas tyre wear within urban catchments is constant and accumulated roadside material containing road-dervied metals is delivered to the receiving waters by storm events. This constant accumlation of other catchment sources may also occur such as degradation of metal roofs as they age.

Given the above, there appears to be no apparent need to conduct a study to simultaneously assess metal loading from catchment sources of urban runoff and from WWOs in an urbanised stream catchment of Sydney. Although, a study such as this may have some merit to help inform management of trunk stormwater assets managed by Sydney Water.

Summary

The estimated dilutions required to remove toxicity of 1 in 4 were determined from continuous exposure toxicity testing of wet-weather rain ingress diluted sewer influent samples from both the Buffalo and Darling Mills Creek locations. Testing under 24 h and 6 h pulse exposures of wet-weather influent returned reduced the dilutions required to remove toxicity to greater than 1 in 2.

The time-weighted average concentration (TAC) line of evidence to assist interpretation of toxicity results documented lower ammonia, copper and zinc concentrations for 24 h and 6 h pulse exposures compared to continuous 192 h exposures. Corresponding no-observable-effect concentrations (NOEC) values from pulsed toxicity testing illustrated reduced toxic effects when compared to continuous exposure toxicity testing over 192 h.

The overall hazard assessment of copper and zinc indicated a low hazard risk for all four 6 h pulse scenarios and for the downstream receiving water scenarios for 24 h pulses in both streams.

As stated for the preceding toxicity study of the Vineyard location, continuous exposure dilution values required to remove toxicity are considered conservative estimates. Whereas pulse exposure scenarios presented in the current study increased realism to better represent WWO spill durations of 24 h or less documented across the 83 ERSs studied that spill into receiving waters under the WWOM project (as outlined in Section 4.5.2).

4.6 Key findings from contaminants of concern studies

An assessment of the toxicity against scientific literature for the 18 organic chemicals (pharmaceutical and personal-care products) of concern tracked in the study outlined in Section 4.4 (Besley et al., 2023), which indicated that none appeared to pose concerns for ecosystem health before wet-weather ingress dilution, and this was even less likely after dilution in the receiving waters (Section 4.4, Besley et al. 2023). Companion laboratory measurements of this suite of 18 organics from water samples collected as part of laboratory toxicity testing recorded similar low concentrations (as documented in Section 4.4, Besley et al., 2023).

The study identified influent concentrations of ammonia was 79× higher in dry weather than the ANZG (2018) guideline value for 95% species protection along, with subsequent decrease of ammonia documented in intermediate wet-weather (19×), and a further concentration decline (1.4×) occurred in wet-weather samples. This established that dry-weather overflow spills would present a substantial risk of adverse ecological effect at Vineyard compared to an increasingly diminished risk as wet-weather severity increased.

Exceedance of the ANZG (2018) guideline ammonia value was established for the receiving waters of Vineyard, whereas at the receiving water locations of Buffalo, Darling Mills and Gymea Bay, measured concentrations were below the ANZG (2018) guideline value for 95% species protection. At these four locations with spatially separated ERSs, there appeared to be sufficient dilution from rainwater ingress within the sewer carrier or by further dilution upon mixing with receiving waters (Section 4.4, Besley et al., 2023). Hence, tracking of ammonia in the water column affords a line of evidence to evaluate locations where ammonia is a concern.

Outcomes of toxicity testing of influent collected under rainfall ingress diluted conditions along with companion receiving water samples determined a risk of adverse ecological effect for samples collected from the Vineyard location that had an atypical convergence of trunk sewers and spatially adjacent ERS spilling into a small urban stream. This provided a supportive line of evidence to the relatively cost-effective assessment afforded by direct ammonia measurement and comparison to ANZG (2018) guideline value.

Interrogation of WWO spill durations from assembled ERS gauging records clearly indicated WWOs were of short durations with the vast majority below a 24 h duration with durations of 6 h or less typically comprised over 60% of WWO spill events (Section 4.5.2). This insight into WWO spill durations suggested the multiday chronic toxicity test outcomes potentially over-represent the risk of adverse ecological effect. As such, the determined required dilutions to remove toxicity are conservative (Section 4.5.1: Kumar et al., submitted). The pulsed toxicity tests subsequently applied under evaluation of WWOs sampled at the Buffalo and Darling Mills locations (Section 4.5.2, Kumar et al., submitted) better mimic the vast majority of ERS spill event durations under all but the most extreme storm events where rainfall was at the highest amounts documented (Figures 4.9 to 4.15), suggesting very high dilutions in the receiving waters. Dilutions required to remove toxicity from testing under 24 h and 6 h pulse exposure scenarios of wet-weather influent from Buffalo and Darling Mills locations were greater than 1 in 2.

To assist with interpretation of toxicity results, concentrations of copper, zinc and ammonia were evaluated with a time-weighted average concentration (TAC). This line of evidence documented

lower ammonia, copper and zinc concentrations for 24 h and 6 h pulse exposures compared to continuous 192 h exposures. Corresponding no-observable-effect concentrations (NOEC) values from pulsed toxicity testing illustrated reduced toxic effects when compared to continuous exposure toxicity testing over 192 h. These outcomes were also supported by the overall hazard assessment of copper and zinc, which indicated a low risk for all four 6 h pulse scenarios and for the downstream receiving water scenarios for 24 h pulses in both streams.

Toxicity testing established that the metals copper and zinc were contaminants of concern, although influence on ecosystem health from these two metals is potentially ameliorated by relatively high dissolved organic carbon concentrations within influent. The observed remobilisation of metals from sewer sediment bedload under prolonged wet-weather events tracked across multiple days when assessed under hazard quotient criteria suggested a contribution to toxicity under wet-weather conditions, but that analysis did not reflect the overall hazard assessment, which reflected a diminishing trend from direct toxicity assessment under wet-weather conditions as was also observed for ammonia (Section 4.5.1; Kumar et al., submitted). Bedload sediments are likely to be present as insoluble sulfides and after release into the influent water column where metals undergo oxidation, oxidised iron sulfide (FeS) quickly forms colloidal or precipitated $Fe(OH)_3$ which can sequester metals again from the influent water column. This may contribute to minimising the extent of biological effect from metals. Hence, copper and zinc can be considered as secondary contaminants of concern after ammonia.

Copper, lead and zinc are also commonly studied road-derived metals. Road-derived metals have been identified as the chief contributor of metals to stormwater from a review and critical assessment of over three decades of research supplemented by global studies (Birch, 2024). Copper and zinc were amongst those 12 contaminants identified by Bickford et al. (1999) risk assessment of the Sydney region, with 85% of the load of those chemicals contained in stormwater (described further in Section 4.1). Investigations of highly urbanised sub-catchments of the Sydney estuary (Davis and Birch, 2009) showed that the contaminants copper, lead and zinc were predominantly (79-87%) derived from diffuse sources (residential properties and roads). This raises the question, does monitoring copper and zinc in the receiving water column provide a meaningful evaluation of the risk posed for adverse ecological risk by WWOs given the potential confounding from stormwater contributions? Hence, the occurrence of copper, lead and zinc in both stormwater and in sewer influent inhibits a potentially meaningful line of evidence in assessment of WWOs. Therefore, assessment of metals is not recommended to be a line of evidence for the WWOA program.

Comparison of contaminants of concern studies with the 1999 risk assessment of Sydney region

The identification of ammonia as a contaminant of potential concern in both influent and in the receiving waters of Vineyard Creek 140 m downstream of the atypical agglomeration (cluster) of ERSs (Section 4.4, Figure 4-5, Besley et al., 2023), together with the toxicity identified through the application of ecotoxicological approaches (Section 4.5.1, Kumar et al., submitted), provides evidence in support of the findings of Bickford et al (1999) that ammonia is a chemical of concern in influent from their risk assessment of 11 major ERSs of the Sydney region. Although they were unable to draw conclusions on toxicity in downstream receiving waters (as outlined in Section 4.2). In contrast, the WWOM pilot studies have enabled comment beyond those of Bickford et al.

(1999), that ecological effects may be manifest at some receiving water locations where dilution of WWO spills is inadequate. The question arises, at what percentage of locations is there a risk of potential adverse ecological effects?

4.7 Locations where adverse ecological effects may occur

The above pilot studies (Besley et al. 2023: Kumar et al. submitted, submitted2) when evaluated in concert indicate an adverse ecological risk is likely to be observed in urban streams where the volume of spilt influent from WWOs is greater than the amount of dilution possible in the freshwater streams under stormflows, as appears to be the case for Vineyard Creek. A confirmed ecological effect was documented under companion studies of morphometrically identified and enumerated macroinvertebrates collected from Vineyard Creek (Section 5.3.1). This adverse ecological effect manifested as a discernible ongoing (press) disturbance in biotic index scores and was also evident in assemblage structure data (Section 5.3.1).

Under the morphometric macroinvertebrate line of evidence, another 22 urban streams (including Buffalo and Darling Mills) were studied. The potential adverse ecological effect indicated in Section 4.4 and Section 4.5.1 above was observed with differences in biotic index scores and assemblage structure data patterns in Sections 5.3.1. This represented an adverse ecological effect over and above the background stormwater disturbance from the urban landscape. Discernible ongoing disturbances were also apparent at Kittys (Section 5.3.5, Figure 4-19), Frenchs (Section 5.3.6, Figure 4-20) and Girraween (Section 5.3.4, Figure 4-21) creeks.

These discernible ongoing disturbances at Kittys, Frenchs and Girraween creeks appear to relate to insufficient dilution of WWOs in these receiving waters. This was observed after visual inspection of the receiving water environments of Girraween, Kittys, and Frenchs creeks along with context of ERSs size and spatial placement (Figure 4-18). Descriptions of each of these creeks along with ERS context are provided below.

Kittys Creek, East Ryde, is situated in a narrow bushland corridor with adjacent housing. Kittys Creek is very narrow and mostly ponded with virtually no flow under dry weather conditions (Figure 4-19A). Kittys Creek had little flow (Figure 4-19B and C) at the end of a brief 5 mm, 30 min rain event, and the flow that was observed to Kittys Creek was efficiently delivered by urban stormwater drains, an example of which is illustrated in Figure 4-19D. On Kittys Creek, a gravity fed ERS (AGN 1286788, SO0153) with an opening of 500 mm was observed. Documented across the duration of WWOM, this ERS (AGN 1286788, SO0153) had an average spill rate of 13 L/s (maximum 124 L/s).

Frenchs Creek, Belrose, is situated in bushland below housing. Pooled sections of Frenchs Creek are wider than Kittys Creek but still relatively shallow. Reaches adjoining those pooled sections were just flowing after 5 mm of rainfall the night before visiting (Figure 4-20B & C), suggesting dilution of WWOs in this receiving waterway were relatively low. The gauged gravity fed ERS (AGN 1395935, SO0131) had an opening of 600 mm in diameter (Figure 4-20D). On the screen bars, prohibiting entry into this ERS, were accumulated wet wipes (Figure 4-20D). Wet wipes on screen bars of the ERS suggests a relatively higher volume WWO and a higher average spill rate for a gravity-fed ERS that is comparable to the ERS that spills into The Ponds Creek (219 L/s)

documented in Besley and Cassidy (2022, Table 2, Figure 3 and reproduced below Figure 4-23). Gauging of the French's Creek ERS (AGN 1395935, SO0131) from January 2020 to March 2021 recorded an average spill rate of 264 L/s (maximum 420 L/s). An adjoining stream upstream of the Frenchs Creek downstream site had an ungauged ERS (400 mm) with the discharge race scoured clean of moss suggesting WWO activity (Figure 4-20E & F). At the time of the site visit, virtually no flow was evident in this adjoining stream despite rainfall the night before. This further suggests insufficient dilution of WWOs may be occurring in Frenchs Creek, resulting in the documented adverse ecological effect.

Girraween Creek, Toongabbie, drains a more urbanised landscape with a riparian zone comprised of a thin line of planted trees for much of this urban stream reach between the upstream and downstream sites (Figure 4-21). Girraween Creek is also a relatively narrow stream and had low water depth observed on the day of the catchment walk despite 5 mm of rainfall the night before (Figure 4-21). Along this creek, four (highest risk prioritised) ERSs spill directly into Girraween Creek (Figure 4-21E and F) within a 2 km stream reach between the upstream and downstream sites. Another four ERSs with low modelled volumes (2.2 ML/year) discharge into stormwater that drains the surrounding urban landscape before entering between these sites. Two ERSs (both 600 mm) fitted with duckbill check valves were present immediately above the downstream site (Figure 4-21E and F), as was a third upper most of these four ERS that discharged directly to Girraween Creek. The fourth outlet was 300 mm in diameter (AGN 1298935, SO0081) and nearest to the downstream site. The ERS had a gauged volume of 36 ML between October 2018 and April 2021. The combined modelled volume of these four ERS discharging directly to Girraween Creek is 56 ML/year. During the catchment walk, 15 stormwater outlets of 150 to 200 mm diameter were observed within a 0.5 km reach of the stream (Figure 4-21D), along with a 750 mm stormwater outlet draining road runoff (Figure 4-21) just above the downstream site within this 2 km stream reach ground-truthed. This suggests that the volumes of WWO spills delivered to Girraween Creek are highly unlikely to be insufficiently diluted by stormwater inflows or by the starting dry weather volume of Girraween Creek.

A fifth ongoing (press) disturbance was identified in Avondale Creek, Turramurra (Section 5.2), however, this appears to relate to a partly flooded ERS (AGN 1336048, SO0188) (Figure 4-22), with potential two-way flows occurring into the sewer and from the sewer into this receiving waterway. A strong sewer odour at this location was detectable up to 100 m away from the ERS. Hence, this adverse ecological effect appears to be related to poor sewerage system maintenance (a sewer network fault), as opposed to the stream size dilution capacity described above for Kittys, Frenchs and Girraween creeks.

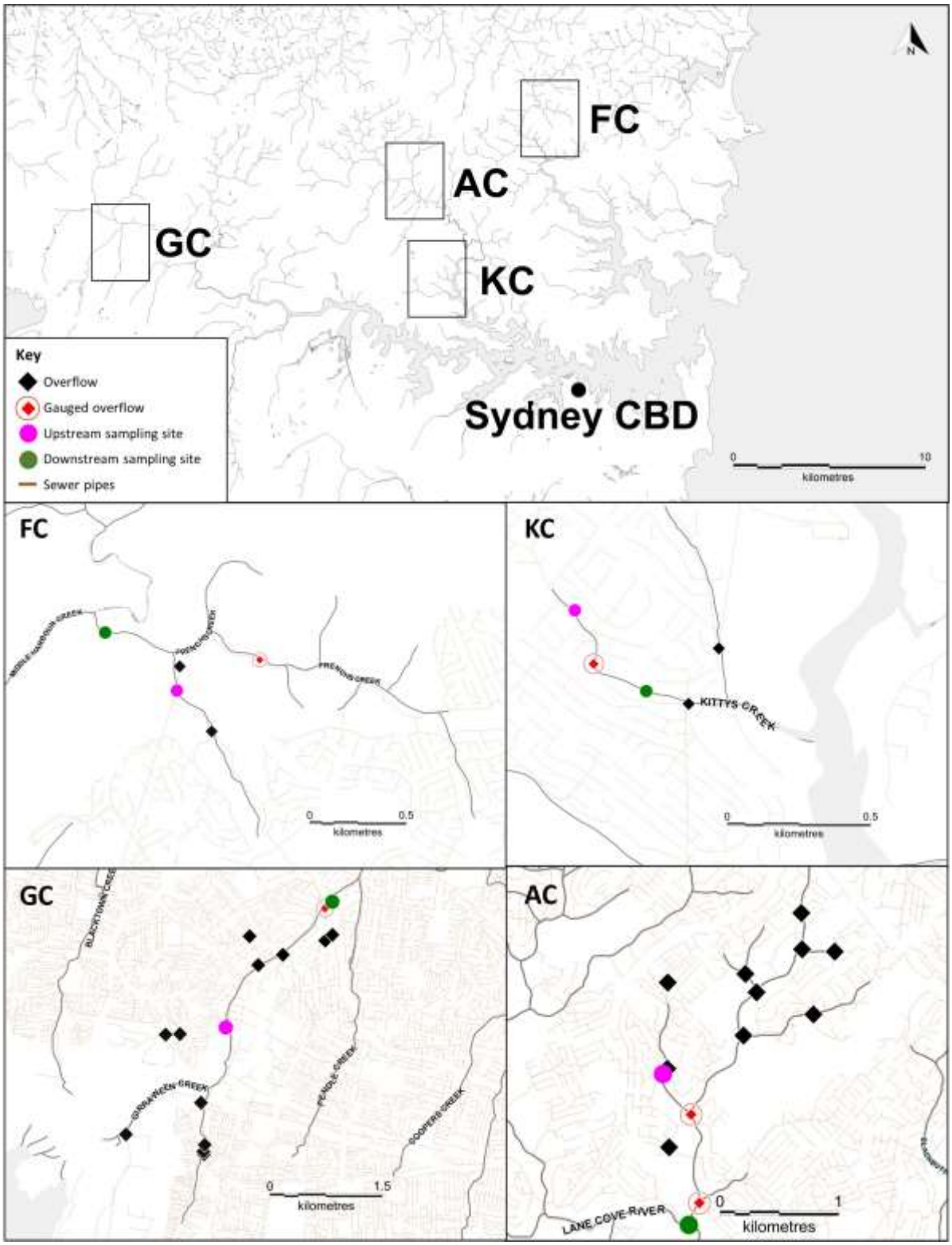


Figure 4-18: Overall location map and detailed maps of Frenchs Creek (FC), Kittys Creek (KC), Girraween Creek (GC) and Avondale Creek (AC) study locations

Figure 4-19: Images of Kittys Creek under dry and under a minor storm event

Kittys Creek: (A) during an extensive period of dry weather virtually no flow; (B) flow in creek at end of a 5-mm in 30-min rain event, the ERS did not discharge under this rainfall; 500 mm ERS visible in both images; (C) in stream reach downstream of ERS showing flow after the rainfall event; (D) showing stormwater delivered from an urban roadway stormwater drain 100 m upstream of the ERS.

Figure 4-20: Images of Frenchs Creek on a day after 5 mm of rainfall the night before visit

Frenchs Creek (A) and (B) looking upstream; (C) looking downstream, with Nathan next to end of ERS outlet; (D) ERS outlet with wet wipes obvious on-screen bars suggesting recent activity; (E) and (F) lower two images of adjoining branch of Frenchs Creek that is upstream of downstream site with another ungauged ERS with concrete race suggesting activity. Visit was on a day after 5-mm of rainfall overnight, both branches of Frenchs Creek had virtually no flow.

Figure 4-21: Images of Girraween Creek on a day after 5 mm of rainfall the night before the visit
Girraween Creek: (A) upstream site; (B) 500 m below upstream; (C) 750 m from upstream site; (D) about 1 km from upstream site (E) 600 mm duckbill ERS mid-way toward downstream site; (F) 600 mm duckbill ERS just above downstream site; (G) 750 mm stormwater outlet draining road runoff; (H) looking upstream from downstream site; (I) ponded downstream site; and (J) zoomed in image of that ponded water. The visit was after 5 mm of rainfall overnight, and there was virtually no flow.

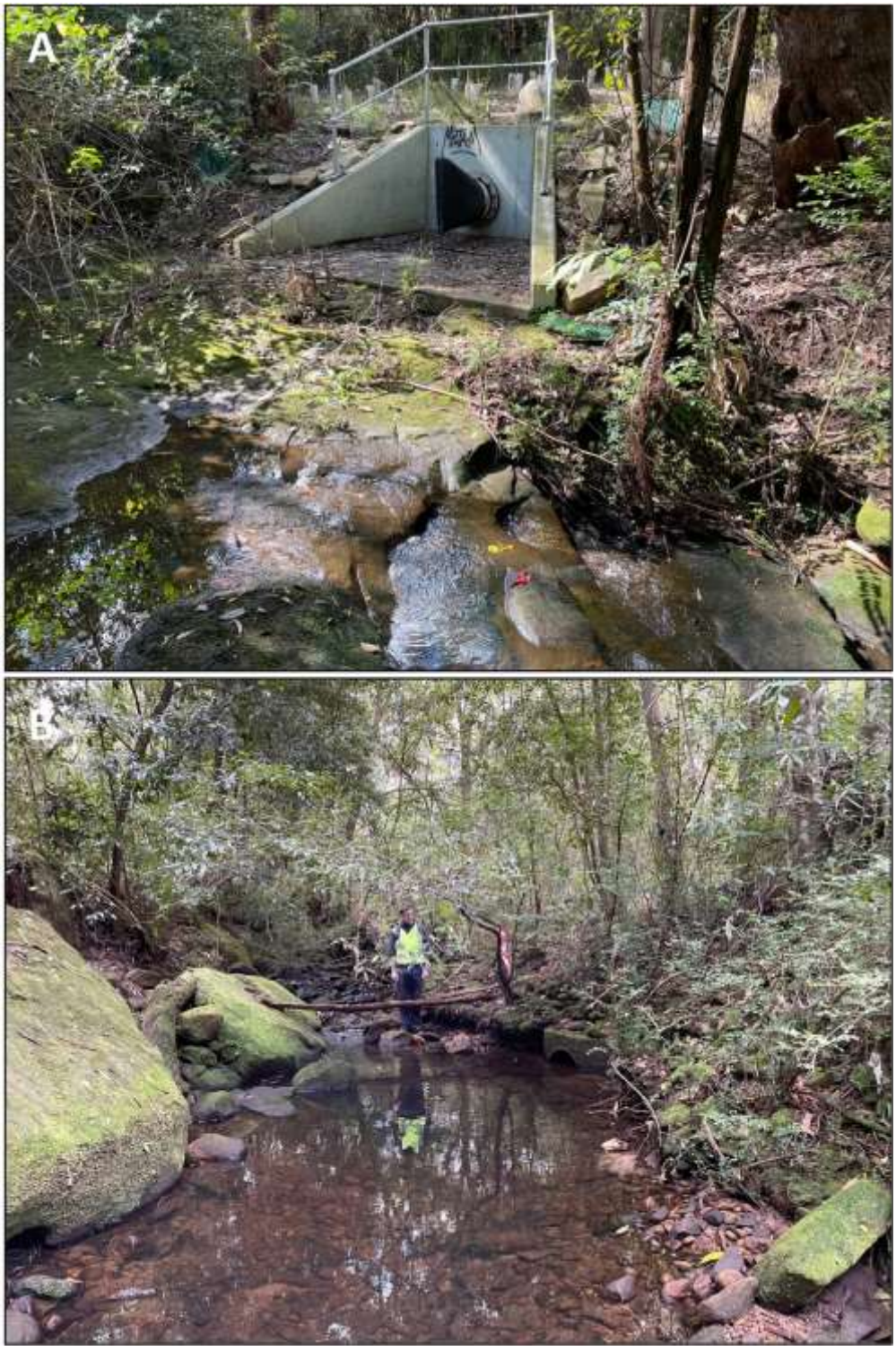


Figure 4-22: Images of Avondale Creek on a day after 5 mm of rainfall the night before the visit
Avondale Creek: (A) showing upstream site and new duckbill ERS, ecology data suggests this ERS is not active; (B) showing downstream site looking toward study ERS mostly submerged. A very strong sewer odour was apparent up to 100 m as approaching this downstream site. Ecology data suggest an ongoing adverse ecological effect exists at this downstream site.

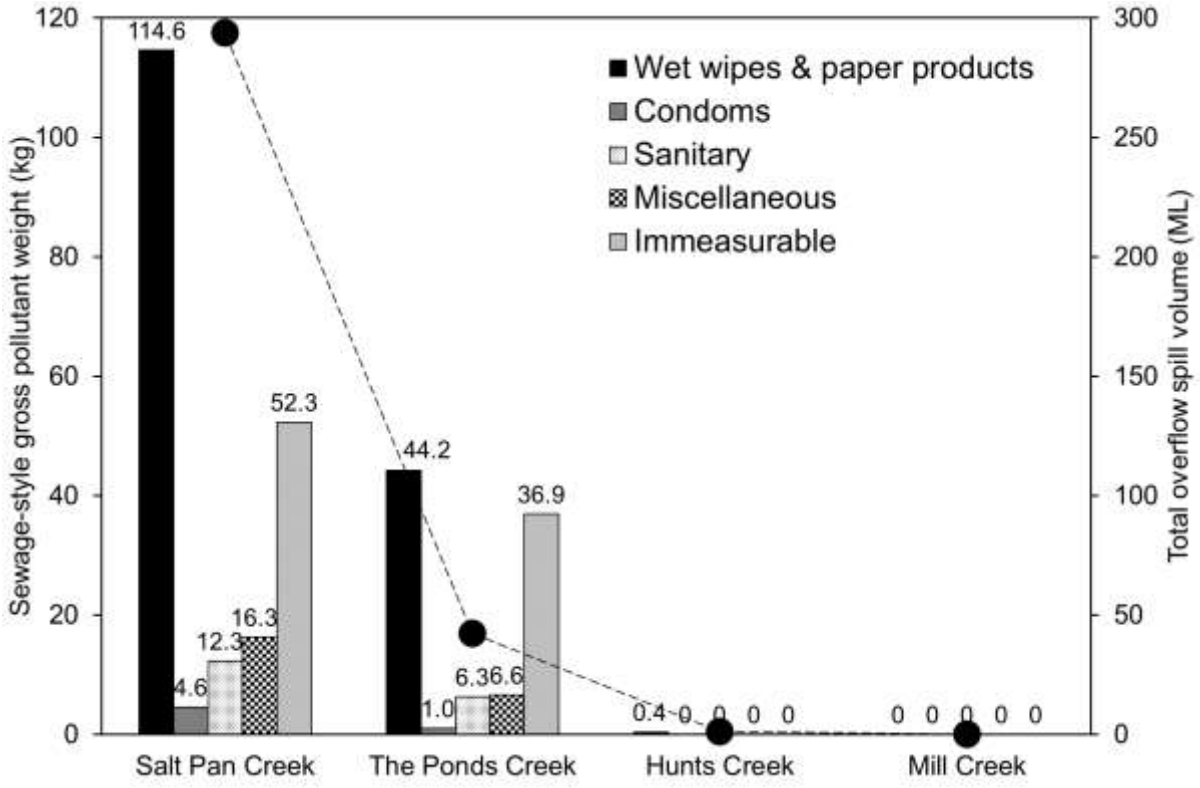


Figure 4-23: Summary of sewage-derived gross pollutants by category collected from four ERSs that each had a net attached to collect these pollutants during WWOs

Individual category weights (kg) by location are shown at top of each bar. Black dot indicates total overflow spill volume (ML) from Besley and Cassidy (2022) WWOM study of gross pollutants from WWOs.

The adverse ecological effects outlined above are described further in Section 5.3.

4.8 Application of learnings from contaminants of concern pilot studies in future management of the wet-weather overflow abatement program

Inputs to Sydney Water WWOA program risk assessment prioritisation tool

In determining where adverse ecological effects are likely to be observed, it is recommended that for each stream, the number of ERSs and their spatial proximity along the stream is evaluated in conjunction with the stream size, particularly width and depth, as inputs into the risk assessment prioritisation tool. The size of each individual ERS outlet pipe should also be evaluated against stream size.

This recommendation would help determine risks where multiple ERSs spill into a very small stream and for situations where an oversized ERS spills into a very small stream.

Evaluations of both ERS size, ERS number and stream size would allow identification of locations where dilution in receiving waters is likely to be inadequate and ammonia may occur as a contaminant of potential concern.

It is envisaged that evaluations would require an initial desk top assessment for ERS characteristics followed by a catchment walk to verify the desktop assessment and to categorise stream size attributes.

Chemical tracers of the presence of sewage

The suite of eight chemical markers (acetaminophen, ibuprofen, metformin, sucralose, theobromine and three benzotriazoles) identified from the tracking contaminants study (Section 4.4, Besley et al., 2023) may be advantageous in locating and confirming sources of sewage contamination. This suite can be used under:

- Wet-weather conditions provided sampling is conducted within two days of a suspected ERS spill to enable detection of sewage contamination
- Dry-weather conditions, where the detection of these markers would indicate an existing sewage source and ongoing sewage contamination, as these chemical markers have short half-lives (Section 4.4).

Ammonia as an indicator of adverse ecological risk

The study at Vineyard Creek, where ammonia presented as a contaminant of potential concern (Sections 4.4 and 4.5.1), established ammonia as an assessment indicator for potential adverse ecological risk.

Outline of proposed future application of ammonia in the WWOA risk prioritisation methodology of the WWOA

In determining where the risk of adverse ecological effects is likely to be observed, identifying urban streams with a diminished capacity for dilution of WWO spilled volumes would be prudent. **As a first pass** in assessing this risk:

I. it is recommended that for each stream, the number of ERSs and their spatial proximity along the stream is evaluated in conjunction with the stream size, particularly width and depth under dry weather base flow for an understanding of potential receiving water dilution capacity. Catchment walks maybe required to undertake this evaluation
II. the size of each individual ERS outlet pipe should also be evaluated against stream size
III. an understanding of WWO exposure expressed as the modelled volume (in the absence of gauged data) from all ERSs spilling to a stream or stream reach is required

Consideration of these three measures in concert will identify streams that may have a risk of adverse ecological effect. Assessing these measures would help determine situations where multiple ERS spill into a very small stream and for situations where an oversized ERS spills into a very small stream. Situations where atypical agglomerations of ERS discharge in close proximity, such as in a stream like Vineyard Creek, should be assessed in the future program.

Outcomes of morphometric ecology line of evidence identified 22% of urban streams assessed had apparent ongoing disturbances attributed to WWO spills (Section 5.3). This in turn perhaps suggests up to 660 of 3000 ERSs may require further assessment as recommended below. The earlier risk assessment conducted by Bickford et al. (1999) indicated 200 of the 3000 ERSs operate in smaller rain events. This perhaps informs the level of effort that may be required in the next phase of assessment to define a revised input for the WWOA risk prioritisation methodology.

Second phase assessment. Once candidate ERSs have been identified to pose a risk of adverse ecological effect from diminished capacity for dilution out of the first pass assessment, instrument arrays would be deployed for periods of time to assess ammonia in receiving waters before during and post WWOs.

I. Deployment of instrument arrays that can continuously measure and provide total ammonia. For example, from one potential supplier, their ammonium sensor will automatically derive ammonia and total ammonia when used in conjunction with the pH/ORP and conductivity sensors
II. Ideally instrument arrays are deployed to record data across a few differing magnitude wet-weather events to effectively ground truth the desktop and field walk assessment from phase 1 outlined above
III. Gathered total ammonia data would then be evaluated against the ANZG (2018) guideline value for protection of 95% of species. The 95% guideline value is considered prudent given the stormwater delivered metal contamination has excluded metal intolerant species from urban streams
IV. Ammonia concentration departures above the ANZG (2018) guideline value for each respective ERS would then be an input into the WWOA risk prioritisation methodology to represent risk of adverse ecological effect from WWOs from an ERS

V. ERS not taken forward into the second phase could be assigned ammonia concentrations equal to half the guideline value to allow inclusion in overall ranking across the 3000 ERSs

Deployment of instrument arrays avoids the cost associated with:

- setup of auto-samplers, collection of samples, laboratory processing of samples
- limited human resourcing to respond after wet-weather events, and
- also avoids the cost of buying and installing and maintaining sewer gauges into ERS

Downloading of data from retrieved instrument arrays and interpreting against Hydstra rainfall data would represent post processing to identify ammonia maximums in wet-weather to rank ERSs as an input into the risk prioritisation methodology and in the case of dry-weather ammonia detections to further explore undetected leakage issues.

It is envisaged after an ERS or stream reach with a number of ERS are assessed that instrument array would be redeployed to another location. This could be progressively performed to continuously improve the WWOA risk prioritisation methodology.

5 Evaluating adverse ecological effects

The studied waterways documented within this section of the report were evaluated for the potential of adverse ecological effects from WWOs in the receiving waters of freshwater streams and in urban estuarine waterways of the Sydney region.

The overarching question in this evaluation was:

> Can adverse ecological effects from WWOs be separated from those of urban stormwater? If so, do these adverse ecological effects manifest as an ongoing (press) or episodic-intermittent (pulse) response in taxonomic assemblages (communities)?

In a study of urban stormwater and combined sewer overflow discharges, Grapentine et al. (2004) stated that 'while sediments and some biota at sites exposed to wet-weather discharges were often contaminated with metals and PAHs and enriched with nutrients, significant biological degradation measured by sediment toxicity or depauperated benthic communities was not evident. Exposure to stormwater discharges at sites below outfalls could alter composition of benthic communities, but these effects were not strongly related to contaminant concentrations in sediment or invertebrate tissue.' They went onto state that effects of wet-weather discharges on benthic communities at the urban stream sites studied appeared to be small, and their detection was limited by several inherent conditions. These conditions include natural heterogeneity in distribution of benthic invertebrates, episodic (intermittent) exposure to discharges and contaminant flues allowing some recovery, background levels of disturbance, poorly delineated changes in communities caused by physical effects, such as flow and sediment transport, and community dynamic responses (Grapentine et al., 2004). Hence, there are potential challenges in the separation of effects from WWOs and stormwater.

In addition to the above challenges, the following natural influences on variation in taxonomic assemblage structures of urban streams (Section 5.2) also need to be considered in interpretation of statistical analyses. The influence of the weather conditions experienced in the Sydney region during the WWOM study from less frequent rainfall (drying conditions) into drought and back to quite frequent rainfall has also presented a challenge in this evaluation, this is discussed further in Section 5.2.1. Another consideration in the Sydney region is the change in subsurface geology between sandstone and shale. This influence is illustrated in Section 5.2.2. Natural heterogeneity from substrate differences within the same mesohabitat type between sites are also a known influence on community structure data within the Sydney region, which has been documented offshore (Besley and Birch 2019) and in freshwater streams (Besley and Chessman, 2008) of the Sydney region. This consideration in assessing freshwater stream data patterns is discussed in Section 5.2.3.

An overview of the ecological assemblages assessed in urban streams and estuarine receiving waters is provided in Section 5.1. Assessment of urban stream taxonomic assemblages for adverse ecological effects is presented in Section 5.3. To assess potential predictive capacity of

the companion metadata variables, multivariate regression modelling was undertaken against morphometric macroinvertebrate assemblage structure (Section 5.4).

Within Section 5.5, two sub-sections explore raising biotic index (SIGNAL-SG) scores from molecular metabarcoding derived taxonomic assemblage structure data from community-DNA (extracted from hand-picked macroinvertebrate specimens, Section 5.5.1) and from environmental DNA extracted from sediment samples with taxonomy subset to macroinvertebrate sized taxa (Section 5.5.3). The resultant biotic index scores were compared against those obtained from samples of morphometric macroinvertebrate assemblage structure data. A known limitation of gaps in taxonomy within molecular DNA databases is discussed in Section 5.5.2. While Section 5.5.4 provides a recommendation for a future pilot study to refine the molecular recovery of macroinvertebrate size class taxa.

Within Sections 5.6 and 5.7, environmental DNA assemblages in urban streams and estuarine receiving waters were analysed and modelled with complex statistical methods. These methods explored the relationships between metabarcoded derived communities and environmental variables using Canonical Correspondence Analysis (CCA). In addition to the CCA approach, dominant gradients in species composition were explored using Generalised Linear Latent Variable Models (GLLVMs). Two machine learning approaches (neural network and random forest) also examined relationships between the metabarcoded data and companion metadata, and also to biotic index scores (SIGNAL-SG) raised from the morphometric dataset. Text within Sections 5.6 and 5.7 was provided by the Macquarie University project team.

Section 5.8 summarises the outcomes of ecological analyses documented in the above-mentioned sections while Section 5.9 documents recommendations to refine application of molecular methods for paired-site assessments.

Reference

Grapentine, L., Rochfort, Q., Marsalek, J. 2004. Benthic responses to wet-weather discharges in urban streams in southern Ontario. Water Qual. Res. J. Canada. 39, 374-391.

5.1 Overview of ecological assemblages evaluated

Freshwater urban streams

Sydney Water Laboratory Services, a NATA-accredited facility, has long-term capability (1990 onward) in assessing freshwater (benthic) macroinvertebrates based upon morphometric (external physical) characteristics forming one line of evidence. Peer panel advice, when indicators were explored, suggested this existing high capability with macroinvertebrate assessments based on morphometric identification should be included as a line of evidence for assessing urban freshwater streams. This line of evidence was assessed on a taxonomic assemblage (community structure) basis with multivariate statistical approaches, or with a biotic index as an evaluation of stream health at each study site. Appendix 1 contains a graphical summary of the morphometrically identified macroinvertebrate data for all studied streams.

Macroinvertebrates are small animals without a backbone that can be seen without a microscope. They live on the surface or in the sediments of water bodies. Inhabited surfaces include

submerged riparian plant roots; submerged rocks; on and within leaf packs falling from overhanging vegetation; and on and within fallen tree branches. Macroinvertebrates include many insect larvae, for example mosquitoes, dragonflies and caddisflies. Other examples of common macroinvertebrates include crustaceans (such as crayfish), snails, worms and leeches. Macroinvertebrates can populate ponds or streams in large numbers, some of them up to thousands in a square metre.

Biotic indices used in other parts of the world include the ASPT index in Britain (Hawkes, 1997), the ASPT index of the South African Scoring System (SASS: Dickens and Graham, 2002), the Spanish average Biological Monitoring Water Quality (a-BMWQ) score (Camargo, 1993), the New Zealand Macroinvertebrate Community Index (MCI) and its quantitative and semi-quantitative equivalents (Stark 1998, Stark and Maxted 2007), and the North Carolina Biotic Index (Lenat, 1993). The conceptual basis underlying all of these indices is that in the presence of stressors, such as sewage influent, taxa that are sensitive to the stressors tend to be eliminated, or greatly reduced in abundance. Conversely, tolerant taxa persist, and may multiply as a result of less competition, or predation, or because their food supply is increased by organic or nutrient enrichment. Consequently, stress results in a decline in the average sensitivity value of the taxa collected. Index scores therefore act as indicators of the presence and intensity of those stressors to which the index is attuned (Besley and Chessman 2008).

Sydney Water has assessed 'stream health' with the Stream Invertebrate Grade Number Average Level (SIGNAL-SG) biotic index tool. 'S' indicates Sydney region version and 'G' indicates taxonomy is at the genus taxonomic level. This tool provides sensitivity grades for macroinvertebrate taxa and these grades can range from 1 to 10, where a grade of 1 is extremely insensitive to pollution effects and a grade of 10 is extremely sensitive. Sample scores combined the grades of the taxa present along their square root abundance of the taxa counts. The latest version of SIGNAL-SG has determined sensitivity grades of 367 genera over the greater Sydney region according to increasing disturbance from pollution and takes into account stream type (order) and altitude (Chessman et al., 2007). Where higher scores (>6.5) indicate more healthy conditions, scores between 6.5 and 5.1 indicate low-moderate degradation, and under 3.7 indicate severe degradation.

The Sydney region specific version of SIGNAL-SG (Chessman et al., 2007) has benefited from development and testing since the original version of SIGNAL (Chessman, 1995). This testing included the response of SIGNAL to natural and human influenced (anthropogenic) environmental factors (Growns et al., 1995), variations in sampling and sample processing methods (Growns et al., 1997; Metzeling et al., 2003) and most importantly setting sensitivity grades of the taxa objectively (Chessman et al., 1997; Chessman 2003). The SIGNAL-SG biotic index has been demonstrated as an easily communicated measure of wastewater impacts on macroinvertebrates in Blue Mountain streams (Besley and Chessman, 2008).

The emerging environmental DNA (eDNA) approach was also employed to assess freshwater streams based upon samples of benthic sediment. Collection of sediment samples and extraction of eDNA was undertaken by Sydney Water Laboratory Services. Extraction of eDNA was performed by the same team of highly skilled staff that conducted similar molecular work for the human health microbial source tracking pilot studies outlined above in Section 3.

This emerging eDNA approach enabled evaluation of bacteria, diatoms and metazoans (invertebrates) from laboratory processing of six compatible primer pairs (amplicons). Post laboratory analysis (sequencing of eDNA) bioinformatics enabled outputting of data with both taxonomic information and operational taxonomic units. The outputted data of this line of evidence were then assessed with multivariate statistical approaches and under a machine learning approach. This is further discussed in Section 5.6.

An overview of WWOM freshwater study site placements across Sydney is provided in Figure 5-1, in light green. The same sites were assessed under both morphometric and eDNA approaches.

Estuarine waters

As outlined in Section 1.2.1, Sydney Water's previous environmental monitoring of estuarine waters for ecological assemblages focused upon the intertidal zones with epibenthos on rock platform rather than benthic sediments (Section 1.2.1). Under the WWOM study of estuarine waters, sediment samples were collected to form a line of evidence for evaluation of WWOs into urban estuarine locations based upon eDNA with another combination of six primer pairs to allow interrogation of bacteria, diatoms and metazoans (invertebrates). This line of evidence was assessed with multivariate statistical approaches and under a machine learning approach. This is further discussed in Section 5.7. An overview of WWOM estuarine study site placements across Sydney is provided in Figure 5-1, in dark green.

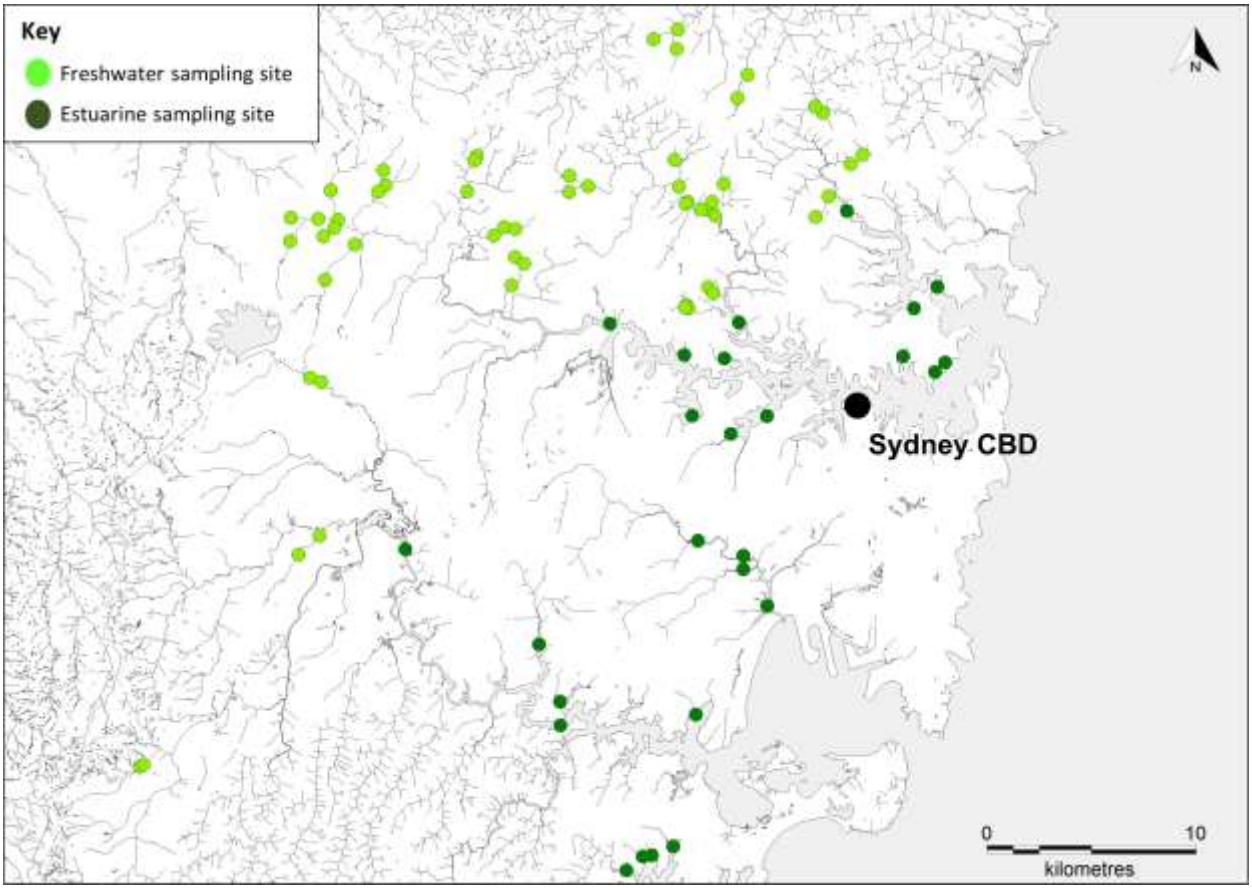


Figure 5-1: Sampling sites for collection of macroinvertebrates, eDNA sediments and deployment of passive samplers

References

Besley, C.H., Chessman B.C. 2008. Rapid biological assessment charts the recovery of stream macroinvertebrate assemblages after sewage discharges cease. Ecol. Indic. 8, 625-638.

Camargo, J.A. 1993. Macrobenthic surveys as a valuable tool for assessing freshwater quality in the Iberian Peninsula. Environ. Monit. Assess. 24, 71-90.

Chessman, B.C., Williams, S.A., Besley, C.H. 2007. Bioassessment of streams with macroinvertebrates: effect of sampled habitat and taxonomic resolution. J. North Am. Benthol. Soc. 26, 546-565.

Chessman, B.C. 2003. New Sensitivity grades for Australian river macroinvertebrates. Mar. Freshwater. Res. 54, 95-103.

Chessman, B.C., Growns, J.E., Kotlash, A.R. 1997. Objective derivation of macroinvertebrate family sensitivity grade numbers for the SIGNAL biotic index: application to the Hunter River system, New South Wales. Mar. Freshwater. Res. 48, 159-172

Dickens, C.W.S., Graham, P.M. 2002. The South African Scoring System (SASS) version 5 rapid bioassessment method for rivers. Afr. J. Aquat. Sci. 27, 1-10.

Growns, J.E, Chessman, B.C., McEvoy, P.K., Wright, I.A. 1995. Rapid assessment of rivers using macroinvertebrates: Case studies in the Nepean River and Blue Mountains, NSW. Aust. J. Ecol. 20, 130-141.

Growns, J.E., Chessman, B.C., Jackson, J.E., Ross, D.G. 1997. Rapid assessment of Australian rivers using macroinvertebrates: cost efficiency of 6 methods of sample processing. J. North Am. Benthol. Soc. 16, 682-693.

Hawkes, H.A. 1997. Origin and development of the Biological Monitoring Working Party Score System. Water Res. 32, 964-968.

Lenat, D.R. 1993. A biotic index for the southeastern United States: derivation and list of tolerance values, with criteria for assigning water-quality ratings. J. North Am. Benthol. Soc. 12, 279-290.

Metzeling, L., Chessman, B. Hardwick, R. Wong, V., 2003. Rapid assessment of rivers using macroinvertebrates: the role of experience, and comparisons with quantitative methods. Hydrobiologia, 510, 39-52

Stark, J. 1998. SQMCI: a biotic index for freshwater macroinvertebrate coded-abundance data. N. Z. J. Mar. Freshwater Res.32, 55-66.

Stark, J. Maxted J.R., 2007. A biotic index for New Zealand's soft-bottomed streams. N. Z. J. Mar. Freshwater Res. 41, 43-61.

5.2 Urban streams

5.2.1 Urban stormwater influence and weather conditions during study period

Urban stream syndrome

The primary degrading process to urban streams is suggested to be 'effective imperviousness' (Walsh et al. 2005a), provided sewer overflows, wastewater treatment plant (WWTP) discharges, or long-lived pollutants from earlier land uses are not operable as these can obscure stormwater impacts (Walsh et al. 2005b). Walsh et al. (2005a) defined 'effective imperviousness' as the proportion of a catchment covered by impervious surfaces that is directly connected to a stream by stormwater pipes. Walsh (2004) determined macroinvertebrate community composition was strongly explained by the gradient of urban density and that most sensitive taxa were absent from urban sites with greater than 20% connection of impervious surfaces to streams by pipes. The direct connection of impervious surfaces, such as roofs, gutters, roads, paths and car parks to a stream allows small rainfall events to produce surface runoff that cause frequent disturbance to the stream through flashy hydrology and delivery of stormwater containing catchment contaminants (Walsh et al. 2005a). It is against this background that we assessed potential changes in community structure or stream health from adverse ecological effects of WWO spills of rainfall diluted influent into urban receiving waters.

As outlined in Section 4.1, road-derived metals including copper and zinc are stressors within stormwater. The literature summarised in that section suggested roads generate a considerable metal load while loading of metals from WWOs was less than in stormwater. Copper and zinc were also identified from as being toxic from testing of influent (Section 4.5). The frequency of stormwater runoff is higher than of WWOs that typically occur in response to higher rainfall. An illustration of this is provided by a visit to Kittys Creek immediately after 5 mm of rainfall where a stormwater drain was observed to be discharging while the ERS was not (Figure 4-19B and D). This may suggest that the more frequent delivery of road-derived metals in stormwater to urban streams and associated toxicity from these delivered metals has potentially influenced ecological assemblages to a greater degree than the additional smaller contributions of metals contained in WWO spills. A strong correlation between the toxic pressure of dissolved metals and invertebrate species was observed by Liess et al. (2017) from reviewing a wide geographical range of Australian streams that were contaminated with heavy metals (mainly copper and zinc). They determined that heavy metal toxicity was positively related to the proportion of predators within the invertebrate assemblage and that taxa richness was negatively affected. A relevant study of metal contaminated streams that also received organic contamination was conducted on Japanese streams by Iwasaki et al. (2018). They predicted that total zinc concentrations of 60 µg/L, twice the Japanese environmental quality standard, do not lead to significant reductions in richness or abundance of macroinvertebrates in organic-contaminated rivers (BOD > 3mg/L). They found at a regional and local scale very few species were present, and that metal-sensitive mayflies were absent. From this study, Iwasaki et al. (2018) stated that an important implication was that macroinvertebrate taxa that are susceptible to metal pollution should be sparse or absent in organic contaminated rivers, so the impacts of metals, such as zinc may be limited as the communities are already species poor. Our pilot study streams meet the Iwasaki et al. (2018) criterion (BOD > 3mg/L) as organic-contaminated, as our companion BOD measurements taken

when water samples were collected for toxicity testing from the downstream receiving waters of Vineyard, Darling Mills and Buffalo creeks were 9, 3 and 5 mg/L, respectively.

Inspection of shade plots for the occurrence of mayfly larvae in the urban streams of the WWOM study indicated sporadic collection of a single mayfly taxa Caenidae *Tasmanocoenis* in Girraween Creek in collection period 13 (Figure 5-25). A single mayfly Baetidae *Cloeon* larva was observed in Frenchs Creek (collection period 10) (Figure 5-33). In Blacktown Creek, both these two mayfly genera were encountered (Figure 5-44). While no mayfly larvae were observed in Avondale Creek (Figure 5-4), Vineyard Creek (Figure 5-13), Buffalo Creek (Figure 5-17), Kittys Creek (Figure 5-29), or in the Darling Mills Creek system (Figure 5-38). In contrast, in near-pristine streams an average of 3 genera (standard deviation of 1) and 5 genera (standard deviation of 2) of mayfly larvae were collected in samples. These near-pristine freshwater stream sites were McCarrs Creek at McCarrs Creek Bridge Road in Garigal National Park and at McKell Avenue at the Hacking River in the Royal National Park. These sites were sampled 27 times each spring and autumn between 2008 and 2021 from the edge habitat of stream pools under the SWAM (Sydney Water Aquatic Monitoring) program.

Hence the smaller metal loading in WWOs is unlikely to contribute further to the pre-existing adverse ecological effects of the dominant metal loading delivered by stormwater inflows, as metal sensitive taxa are already sparse or absent in the macroinvertebrate assemblage in urban streams as illustrated in the above paragraph discussing aquatic mayfly larvae. This in turn suggests that where ammonia from WWOs is at concentrations as a '*contaminant of potential concern*' after mixing in urban streams (Section 4.4), ammonia may contribute the primary separation of adverse ecological effects of WWOs from that of the background adverse ecological effects of stormwater.

References

Iwasaki, Y., Kagaya, T., Matsuda, H. 2018. Comparing macroinvertebrate assemblages at organic-contaminated river sites with different zinc concentrations: Metal-sensitive taxa may already be absent. Environ. Pollut. 241, 272-278.

Liess, M., Gerner, N.V., Kefford, B. 2017. Metal toxicity affects predatory stream invertebrates less than other functional feeding groups. Environ. Pollut. 227, 505-512.

Walsh, C.J. 2004. Protection of in-stream biota from urban impacts: minimise catchment imperviousness or improve drainage design? Mar. Freshwater. Res. 55, 317-326.

Walsh, C.J., Fletcher, T.D., Ladson, A.R. 2005a. Stream restoration in urban catchments through re-designing stormwater systems: looking to the catchment to save the stream. J. North Am. Benthol. Soc. 24, 690-705.

Walsh, C.J, Roy, A.H., Feminella, J.W., Cottingham, P.D., Groffman, P.M., Morgan, R.P. 2005b. The urban stream syndrome: current knowledge and the search for a cure. J. North Am. Benthol. Soc. 24, 706-723.

Influence of weather conditions experienced in the Sydney region during the WWOM

The influence of the weather conditions experienced in the Sydney region during the WWOM study from less frequent rainfall (drying conditions) into drought (Figure 5-2) and return to quite frequent rainfall did induce changes in base flow and community structure.

Figure 5-2: Blacktown Creek: (A) sampling location in 2018; and (B) approximately 200 m upstream of (A) during the severe drought in 2019

To illustrate the influence of broader overarching weather conditions, natural variation can be demonstrated from the WWOM study of morphometric macroinvertebrates in Avondale and Rudder creeks. Up to 17 collection events were conducted at eight weekly intervals from October 2018 to April 2021. In the case of Rudder Creek, a delay in the installation of equipment resulted in no samples being collected from the first period. Collection periods 4 to 6 were sampled during severe drought weather conditions between April 2019 to early September 2019. In mid-September 2019, rainfall influenced inflow and infiltration to cause overflows in some other catchments, but conditions remained relatively dry in Rudder Creek until the start of 2020.

The change in flow conditions is illustrated by the change between presence and absence of three macroinvertebrate animals observed in Avondale Creek across the collection periods. The mosquito larvae Culicidae *Culex* (Figure 5-3A) was collected from the upstream site in collection periods of October 2018 to May 2020 (1 to 11), but was absent in periods July 2020 to April 2021 (12 to 17), when higher base-flow conditions had reestablished (Figure 5-4). This taxon prefers still water ponded conditions. In contrast, the water penny beetle larvae Psephenidae *Sclerocyphon* (Figure 5-3C) prefer flowing water and was regularly collected across collection periods of May 2020 to April 2021 (11 to 17). They were infrequently collected through collection periods of October 2018 to March 2020 (1 to 10) at the upstream site (Figure 5-4). At the downstream site snails of the family Tateidae (*Posticobia* [Figure 5-3B] and *Potamopyrgus*) were routinely collected in periods 1 and 2 and then infrequently collected until period 7 (Figure 5-4).

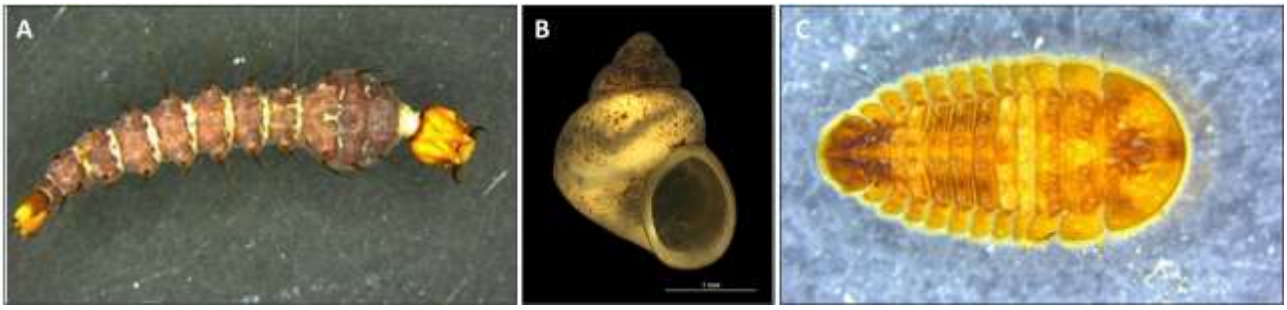


Figure 5-3: (A) Mosquito larvae Culicidae *Culex*; (B) snail Tateidae *Posticobia*; (C) water penny beetle larvae Psephenidae *Sclerocyphon*

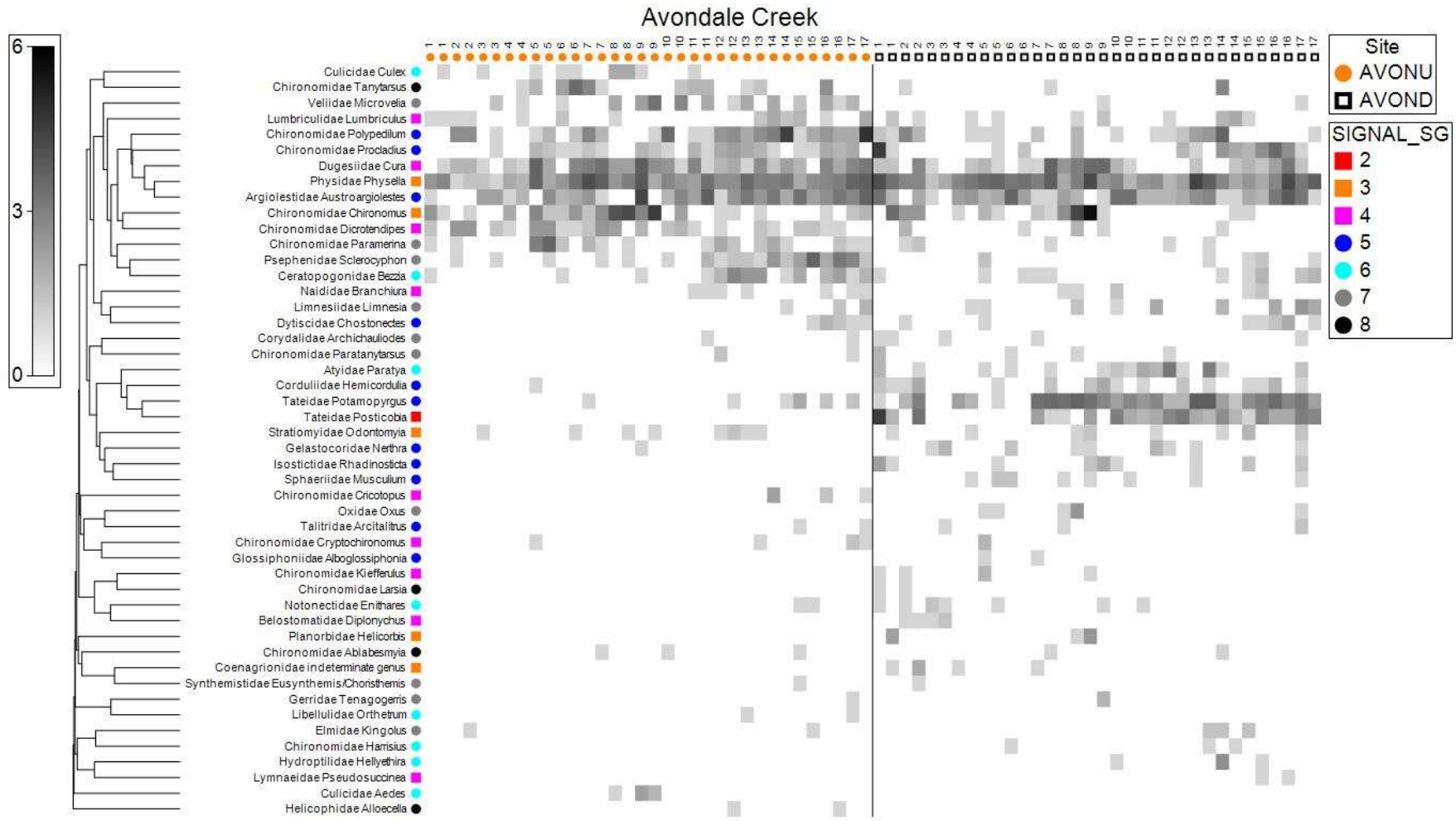


Figure 5-4: Shade plot of community structure data of Avondale Creek by collection period and site.
Upstream site = AVONU; downstream site = AVOND

The ridge line of the Rudder Creek catchment is highly urbanised with extensive bushland flanking this stream in the mid to lower catchment (Figure 5-5). The upstream site is nearest to the urbanised area, while the downstream site is situated about 0.4 km below the upstream site, deeper into the flanking bushland (Figure 5-5). No designed ERS (overflow points) occur above the upstream site. The downstream site on Rudder Creek was inadvertently situated above two ERSs situated on a sewer trunk main that crosses low down on Rudder Creek (represented by a diamond in Figure 5-5). The ERS (AGN 1395908) on the lefthand side of Rudder Creek was gauged. That ERS has only discharged briefly twice, and this was during east-coast-low massive rainfall events (February 2020 and March 2021) of over 300 mm, suggesting enormously high dilution of these spills would have occurred. Both ERSs (AGN 1395908 and 1334893) were modelled to have zero volume and no overflows (in ten years).

Figure 5-5: Aerial view of the Rudder Creek catchment with position of upstream (blue star) and downstream (purple star) sites

The expected adverse ecological effects pattern of higher SIGNAL-SG scores at the upstream site and lower scores recorded at the downstream site was not illustrated in the control chart plot for Rudder Creek (Figure 5-6). Rather, the control chart plot displayed poorer stream health (lower SIGNAL-SG scores) for the upstream site in collection periods before and after the drought. Notably during the drought when stormwater was greatly minimised, the upstream SIGNAL-SG scores increased and became similar to those documented for the downstream site (Figure 5-6). Relatively more stable stream health was observed across the pre-drought period (drying),

drought, and post drought period with return to more frequent rainfall for the downstream site (Figure 5-6).

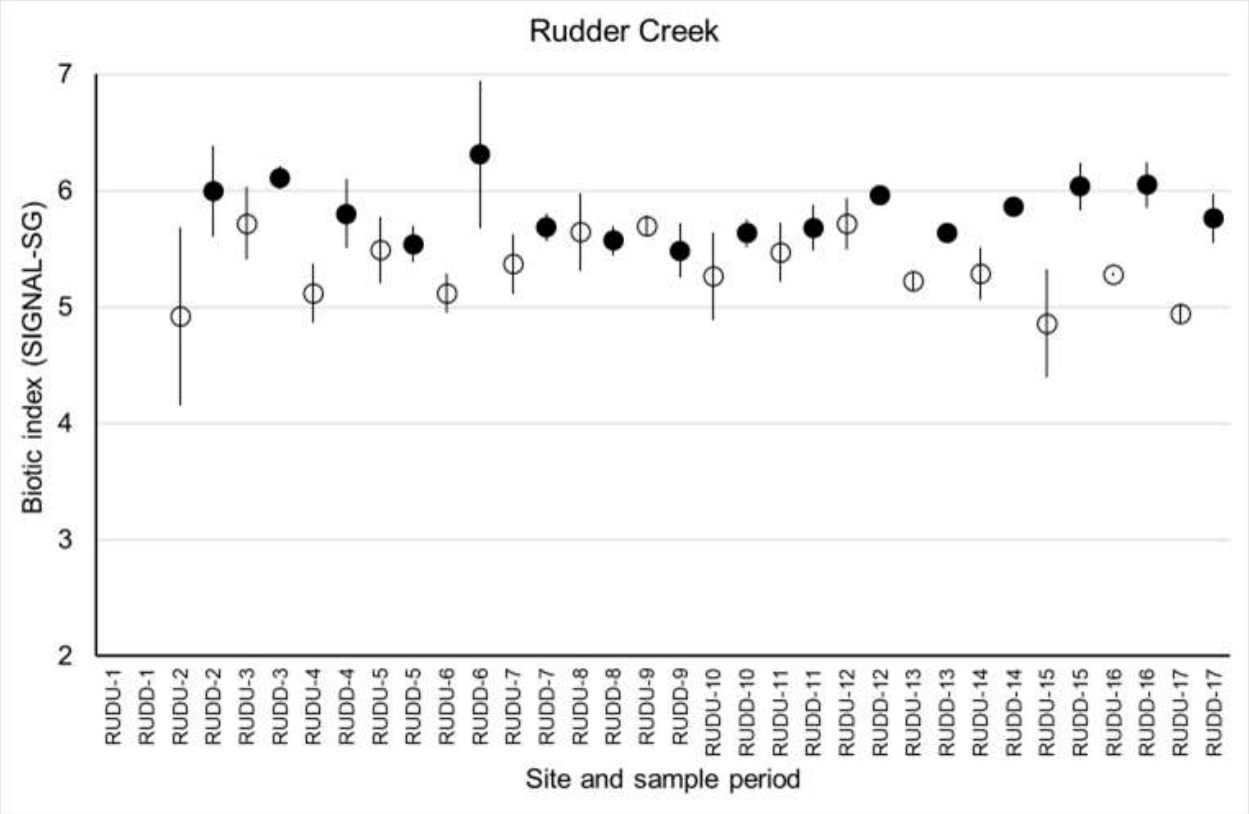


Figure 5-6: Stream health control chart of Rudder Creek upstream (RUDU) and downstream (RUDD) sites across 16 collection periods

This relative stability at the downstream site was also reflected in underlying community structure from Analysis of Similarities (ANOSIM) (Table 5-1). Pairwise comparisons of these three broader temporal periods for the downstream site returned low-range R-values that indicated minor differences in community structure (0.18, 0.19 and 0.22). This contrasted to ANOSIM pairwise comparison results of the upstream site with R-values of 0.29, 0.40 and 0.81 (that was nearest the maximum R-value of 1). These three broader periods were assembled from the 17 sample collection periods:

- Drying = collection periods 2 and 3 (December 2018 and February 2019)
- Drought = collection periods 4 to 9 (April 2019 to January 2020)
- Wetter = collection periods 10 to 17 (March 2020 to April 2021)

Together these ANOSIM pairwise comparison results, along with the patterns, returned in graphical analysis of ecosystem health with the SIGNAL-SG biotic index suggest stormwater inflows from urbanisation above the upstream site influence greater departures in stream health during wetter climatic conditions than under drier conditions, and that this stormwater influence

appears to be ameliorated further downstream where additional stormwater inputs into the Rudder Creek catchment were absent.

Table 5-1: ANOSIM pairwise test comparisons of Rudder Creek community structure across three climatic periods for each site

Pairwise test	R value	P value
Downstream site comparisons		
drying versus drought	0.19	0.055
drying versus wetter	0.22	0.044
drought versus wetter	0.18	0.012
Upstream site comparisons		
drying versus drought	0.40	0.002
drying versus wetter	0.81	0.0001
drought versus wetter	0.29	0.0009

Downstream site pairwise test global R = 0.18, p = 0.008
Upstream site pairwise test global R = 0.46, p = 0.0001
Pairwise comparison climatic periods: Drying = collection periods 2 to 3; Drought = collection periods 4 to 9; Wetter = collection periods 10 to 17

To minimise the influence of this natural variation in taxonomic assemblage structure induced from broader rainfall patterns through the October 2018 to April 2021 WWOM study period, collection periods 10 to 17 (March 2020 to April 2021) were selected for statistical analysis of the morphometric macroinvertebrate dataset. These eight collection periods represented more frequent occurrences of rainfall along with associated widespread WWO activity across study locations.

Extracts of eDNA were sequenced for six of these eight collection periods, as a consequence of constraints arising from human resourcing restraints of the sequencing service provider. These sequenced sample collection periods were:

- period 10 occurred in March 2020 should encapsulate any changes from major east-coast low weather event of 8-10 February 2020.
- period 11 was sampled in May 2020,
- while period 13 was sampled in September 2020 after a moderate rainfall event of 26-28 July 2020.
- sample collections for periods 14 and 15 were conducted in November 2020 and January 2021.
- period 17 saw sample collection in April 2021 after major east-coast low weather event of 18-23 March 2021.

5.2.2 Subsurface geology

Environmental waters of streams flowing through sandstone areas are generally nutrient-poor compared to streams flowing through shale areas. This aspect can be reflected in the richness and abundance (how many) of invertebrates are supported in the stream (carrying capacity). Under the Sydney Water Aquatic Monitoring (previously STSIMP) program samples were collected from both a shale (Hacking River at McKell Avenue in Royal National Park) and a sandstone (McCarrs Creek at McCarrs Creek Bridge Road in Garigal National Park) near-pristine streams. These reference sites have been sampled 27 times each spring and autumn between spring 2008 and autumn 2021 from the edge habitat of stream pools. Over this period, a lower average (22) number of genera (with a standard deviation of 5) were collected from McCarrs Creek compared with collections from the Hacking River with an average of 40 genera, standard deviation of 11. The same trend was seen in collected specimens under a one-hour collection duration, with an average of 77 specimens (standard deviation of 24) from McCarrs Creek across this period compared with an average of 164 specimens (standard deviation of 11) from the Hacking River. This illustrates a higher carrying capacity for the shale reference stream of the Hacking River. Average SIGNAL-SG scores in these near-pristine streams over the spring 2018 to autumn 2021 period were 6.24 (SD ± 0.40) for the Hacking River and 7.45 (SD ± 0.15) for McCarrs Creek.

Another consideration in ecological analyses was the influence of sediment from subsurface geology. Streams with sandstone subsurface geology generally have coarser (metal-poor) sediments compared to streams with shale subsurface geology that have finer particle sized sediments. Finer-sized particles provide potentially more binding sites for contaminants such as metals (Forstner, 1982).

The potential influence of this subsurface geology may be reflected in the early split sites in the cluster analysis (dendrogram – tree diagram) plot of community structure (based on square-root transformed data with the Bray-Curtis resemblance measure with distances among centroids calculated between urban stream site samples) in Figure 5-7. In general, sites with shale subsurface geology had lower SIGNAL-SG scores (Figure 5-7).

To minimise these apparent sources of natural variation, overall modelling was conducted on two groups of sites separated by subsurface geology.

Reference

Forstner, U. 1982. Accumulative phases for heavy metals in limnic sediments. Hydrobiologia, 91, 269-284.

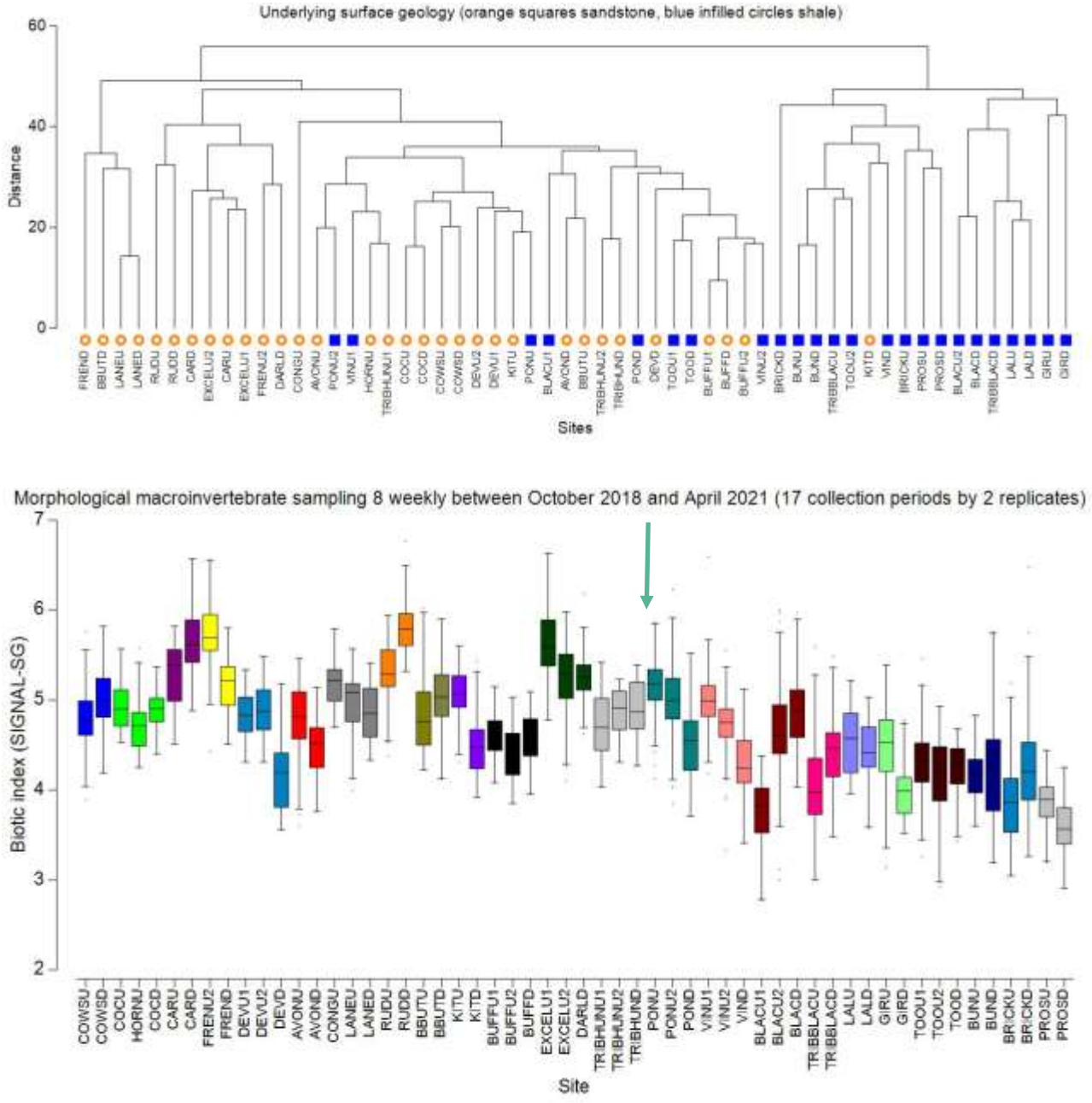


Figure 5-7: Upper pane: Cluster analysis plot based on distance among centroids of community structure data. Lower pane: Box plot of SIGNAL-SG biotic index scores by site across all 17 collection periods

Upper pane: Open circles represent sites with sandstone subsurface geology, while filled-in squares represent sites with shale subsurface geology. Lower pane: Sites to left of annotated arrow in areas with sandstone subsurface geology. Sites to right of arrow occur in areas with surface subsurface geology of shale.

5.2.3 Site-specific assemblage structure influenced by mesohabitat differences

Across the urban study area of the WWOM, the dominant and generally solely present stream habitat was the pool-edge water. Field inspection of the urban stream sites chosen for WWOM revealed the presence of pool-edge habitat in 93% of potential sites (70). Only seven sites had submerged aquatic plant habitat in addition to the edge habitat. A riffle habitat (a water flow over a rock stream bed) was only present at three sites, where the edge habitat was also present. To enable consistency between sample collections, pool-edge waters were sampled at sites over the October 2018 to April 2021 period. This approach reduced the known natural influence of differing taxonomic assemblages between mesohabitats as illustrated in Figure 5-8 (shade plot).

Natural heterogeneity from substrate differences within the same mesohabitat type between sites are also a known influence on community structure data within the Sydney region. Besley and Chessman (2008) documented site-specific macroinvertebrate assemblages of near-pristine streams situated in National Parks or bushland reserves of the Sydney region. They found assemblages for the pool-edge water mesohabitat significantly differed between the Georges River and O'Hares Creek for both freshwater streams ($R = 0.44$, $p = 0.001$) that were 4 km apart on the same stream system sampled each six months between 1997 and 2003. On O'Hares Creek, another four sites at three different distances from source of stream were sampled on 4 to 6 occasions between 1995 and 1997 with significant ANOSIM R-values returned between 0.38 and 0.58 (13 and 25 km apart) (Besley and Chessman, 2008). A non-significant ANOSIM R-value was returned for the pair of sites 0.3 km apart.

A similar non-significant outcome was observed for a pair of sites (0.2 km apart) on Erskine Creek sampled on 4 occasions between 1995 and 1997, although for the riffle habitat ($R = 0.46$) a significant outcome was observed between 0.2 km apart between a boulder section versus a cobble section (Besley and Chessman, 2008). These ANOSIM test R-value outcomes from pairwise tests of near-pristine stream sites across varying spatial differences provide context to assess R-values against from pairwise comparisons of urban stream macroinvertebrate assemblages assessed under the WWOM pilot study. This context avoids falsely declaring adverse ecological effects when in fact community assemblage differences are due to differences in natural heterogeneity in substrate between sites.

References

Besley, C.H., Chessman B.C. 2008. Rapid biological assessment charts the recovery of stream macroinvertebrate assemblages after sewage discharges cease. Ecol. Indic. 8, 625-638.

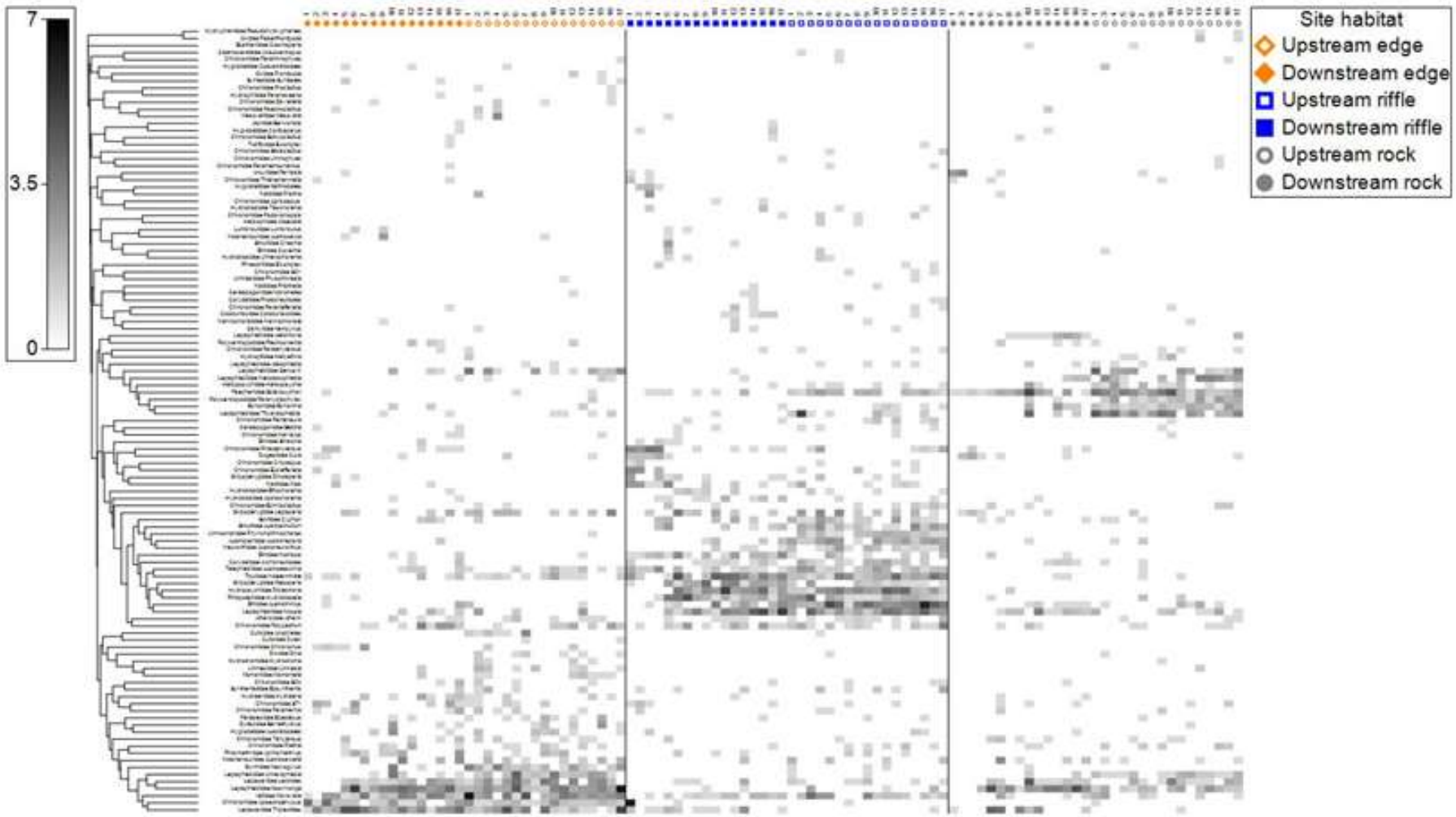


Figure 5-8: Shade plot of three (pool-edge water, riffle and pool rock) mesohabitats of Blue Mountains Creek illustrating differences between dominant taxa assemblages of each mesohabitat

Darker shading represents higher (square root transformed) abundance in a sample

5.3 Assessment of morphometric macroinvertebrate assemblages in urban streams for adverse ecological effects

As outlined in Section 4.7, the pilot studies of Besley et al. (2023) and Kumar et al. (submitted; submitted2) when evaluated in concert indicate an adverse ecological risk is likely in an urban stream where the volume of spilt influent from WWOs is unable to be adequately diluted once spilt into the freshwater stream under stormflow conditions. This appears to be the case for Vineyard Creek where the dilutions required to remove toxicity were not achieved. The companion study based upon morphometrically identified and enumerated macroinvertebrates collected from Vineyard Creek illustrated this adverse ecological effect manifests as a discernible ongoing (press) disturbance in biotic index scores that was also evident in community structure data (Section 5.3.1). Across the other 22 urban streams (that included Buffalo and Darling Mills) assessed with the morphometric macroinvertebrate line of evidence, discernible ongoing disturbances appeared to occur at the downstream study sites of Kittys (Figure 4-19), Frenchs (Figure 4-20) and Girraween (Figure 4-21) creeks and at the upstream site 1 of Blacktown Creek (Figure 5-41). In contrast, discernible disturbances attributable to adverse ecological effects from WWOs in macroinvertebrate assemblage structure were not apparent at downstream sites of Buffalo Creek and Darling Mills Creek. These creeks were identified by data exploration detailed in Section 10.1, further data exploration is detailed below.

The assessment approach to explore data patterns and statistical differences between upstream and downstream morphometric macroinvertebrate assemblages involved the following lines of evidence:

- univariate representation of assemblage data with biotic index SIGNAL-SG scores plotted in an ecosystem health process control chart, followed by scores between sites assessed with a SNK multiple means comparison test
- multivariate assessment of stream assemblages plotted in an nMDS ordination plot, dendrogram (tree diagram) from cluster (classification) analysis, ANOSIM testing for differences between the factor ‘site’ and PERMDISP (homogeneity of dispersions) testing of factor ‘site’

Further detail of these assessment approaches is provided below.

Primary assessment of scores calculated from the SIGNAL-SG biotic index was done visually using plots along the lines of a process control chart for ecological monitoring presented by Burgman et al. (2012) to display information in a simple, practical and scientifically credible way. This style of control chart illustrates temporal trends and allows interpretation of data against background natural disturbance and variation of the respective streams. In these control chart plots, the range of each site period has the mean plotted together with error bars of ± one standard deviation (1SD) of the mean, as recommended by ANZECC/ARMCANZ (2000) for basing ecological decisions. These ± 1SD of the mean formed ranges of stream health for period displayed. These charts were plotted on a collection period basis.

These comparisons had three possible outcomes:

1) Mean downstream stream health was within the range recorded for the upstream site of the same collection period suggesting no impairment from inflow sources between sites
2) Mean downstream stream health was lower than that of the upstream site suggesting impairment from inflows between sites
3) Mean upstream stream health was lower than that of the downstream site suggesting impairment from inflows above the upstream subsequently ameliorated over the distance between the sites

Analysis of variance (ANOVA) was then undertaken based on SIGNAL-SG scores for the factor 'site'. If results from ANOVA were significant Student Newman Keuls (SNK) multiple mean comparison tests were performed to assess site differences. SAS version 9.4 was employed in this univariate testing.

Multivariate data analyses were performed using statistical routines of the PRIMER Version 7.0.23 software package (Clarke et al. 2014) and the add-on module PERMANOVA+ (Anderson et al. 2008).

As specimens were live-picked with generally no more than 20 of the same taxon (as best as could be determined with the naked eye) collected per occasion, this approach partly ranked the data. Laboratory identified and counted data were transformed with a square root transformation to avoid over-transforming the data matrix and squeezing out too much of the quantitative information from mid- to low-abundance genera.

An association matrix was then constructed based upon the Bray-Curtis resemblance measure. This measure was used as the basis for classification, ordination and hypothesis testing of site sample data. The Bray-Curtis resemblance measure is focused on compositional changes in taxa identities (Anderson and Walsh 2013). As such, this is an appropriate choice since we understand that measurable adverse ecological effects were recorded at former aged Blue Mountains WWTPs caused a change in the composition of the freshwater macroinvertebrate assemblage (Besley and Chessman 2008).

The group average classification technique was used to place the sampling sites into groups, each of which had a characteristic invertebrate community based on relative similarity of their attributes. The group average classification technique initially forms pairs of samples with the most similar taxa and gradually fuses the pairs into larger groups (clusters) with increasing internal variability.

Classification techniques will form groups even if the dataset actually forms a continuum. In order to determine whether the groups were 'real', the samples were ordinated using the non-metric multidimensional scaling (nMDS) technique. Ordination produces a plot of sites on two or three axes such that sites with similar taxa lie close together and sites with a differing taxon composition lie farther apart. Output from classification analysis was then checked against sample groupings on the ordination plot to see if site (a-priori) groups of samples occurred which would indicate a response from WWO spills.

An unconstrained ordination procedure such as nMDS usually introduces distortion when trying to represent the similarities between large numbers of samples in only two or three dimensions. The

success of the procedure is measured by a stress value, which indicates the degree of distortion imposed. In the PRIMER software package, a stress value of below 0.2 indicates an acceptable representation of the original data although lower values are desirable.

Shade plots provide a visual display in the form of the data matrix with a rectangle display for each sample. White represents zero counts, while black rectangles represent maximum abundance after dispersion weighting and square root transformation. Increasing grey shading represents increasing abundance. Thus, shade plots represent the patterns of dominant and less abundant genera collected in each sample. To improve visualisation of data patterns in shade plots, genera were serially reordered based on classification of genera. Classification on genera was based on square-root-transformed data that were standardised by total followed by construction of a data matrix based on Whittaker's (1952) Index of Association resemblance measure. SIGNAL-SG grades of each genus level taxon were also annotated onto these plots. These grades provided an indication of sensitivity to increasing pollution that each taxon had which in turn aided interpretation of data patterns.

To statistically test for multivariate dispersion the PERMDISP routine of PERMANOVA+ was run on the factor 'site'. If PERMDISP analysis returned a non-significant result, that indicated a similar pattern of dispersion (spacing between same site samples) for the two sites of the habitat samples being analysed. A non-significant outcome would suggest the variability in taxonomic make-up of samples collected over time was at similar levels for both sites through the period tested. This result then also implies subsequent results of ANOSIM tests are focused on spatial differences in the taxonomic assemblages between sites. In contrast, if dispersion was significant, then subsequent results of ANOSIM tests are describing both the variability in taxonomic make-up of samples collected over time as well as community structure differences between sites.

ANOSIM provides an absolute measure of how separated groups of samples are on a scale of - 1 to 1 (Clarke, 1993). As the R-value approaches 1, this indicates all temporal samples from a site were more similar to each other than they were to samples from another site; that is, groups are clearly different. When the R-value approaches 0, temporal samples within and between sites are equally similar; that is, no differences between groups. If the R-value approaches –1, then pairs consisting of one temporal sample from each site are more similar to each other than pairs of temporal samples from the same site (Clarke, 1993).

5.3.1 Vineyard Creek

Two upstream sites and a downstream site were sampled in Vineyard Creek. The upstream sites were situated on each of two arms of Vineyard Creek about 1.5 km above the downstream site (Figure 5-9).

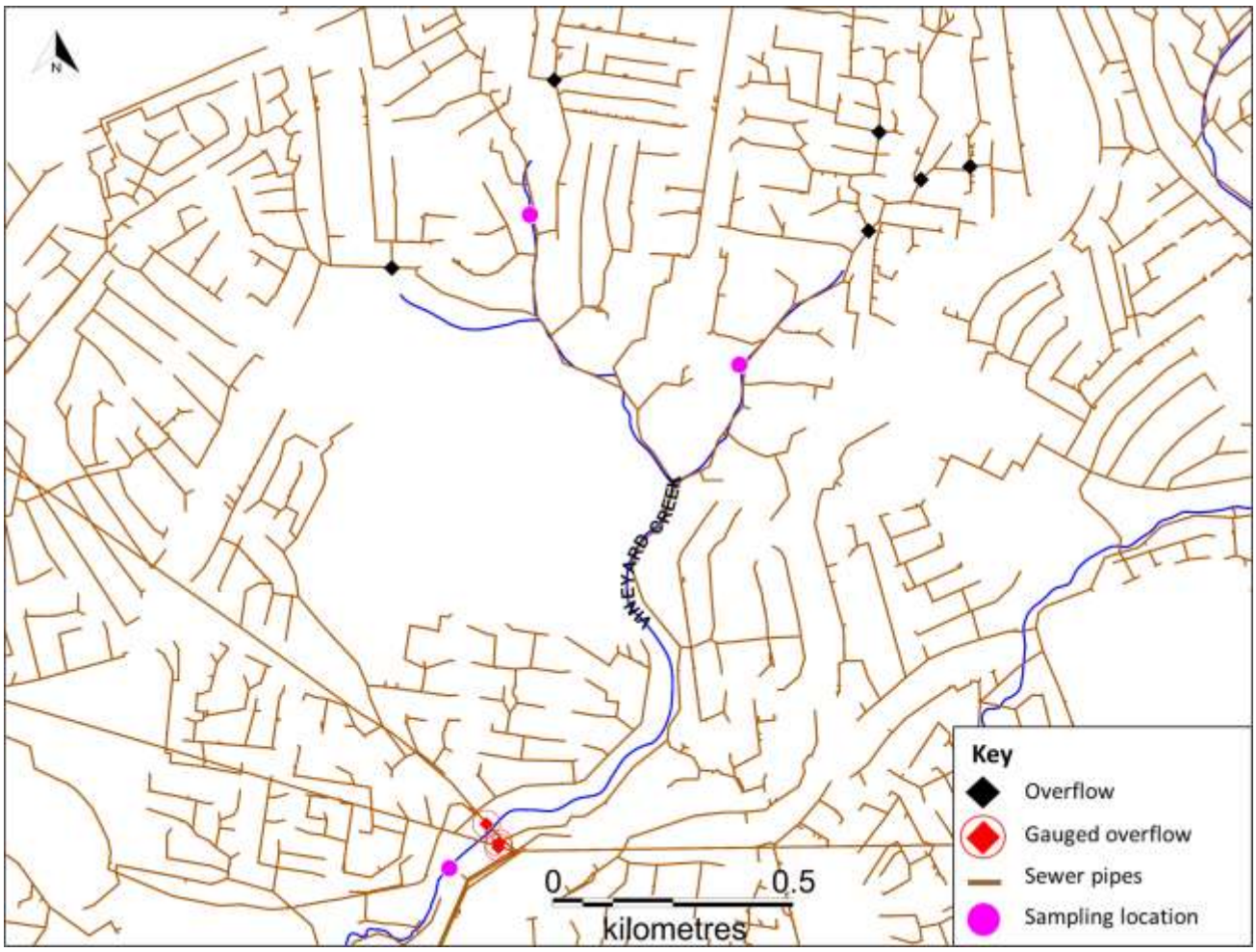


Figure 5-9: Location map of three Vineyard Creek sites

The ecological control chart plot of SIGNAL-SG scores visualised stream health of Vineyard Creek through the WWOM study across October 2018–April 2021 by each 8-weekly collection (periods 1 to 17). This visual comparison illustrated downstream stream health was consistently lower than that of both upstream sites in the post drought La Niña period represented by collection periods 10 to 17 (March 2020 to April 2021) (Figure 5-10).

Statistical testing was conducted on data from collection periods 10 to 17. A comparison of the upstream-downstream SIGNAL-SG site scores from a SNK multiple mean comparison test indicated that the downstream site had a statistically lower mean score (VIND = 4.24 SD 0.40) than that of both the upstream sites (VINU1 = 5.03 SD 0.23, VINU2 = 4.96 SD 0.27), which supported the visual trends in the control chart plot (Figure 5-10). The accompanying ANOVA test of the factor 'site' was significant (P < 0.0001) and Brown and Forsythe's test for homogeneity of variance was non-significant (P =0.2354).

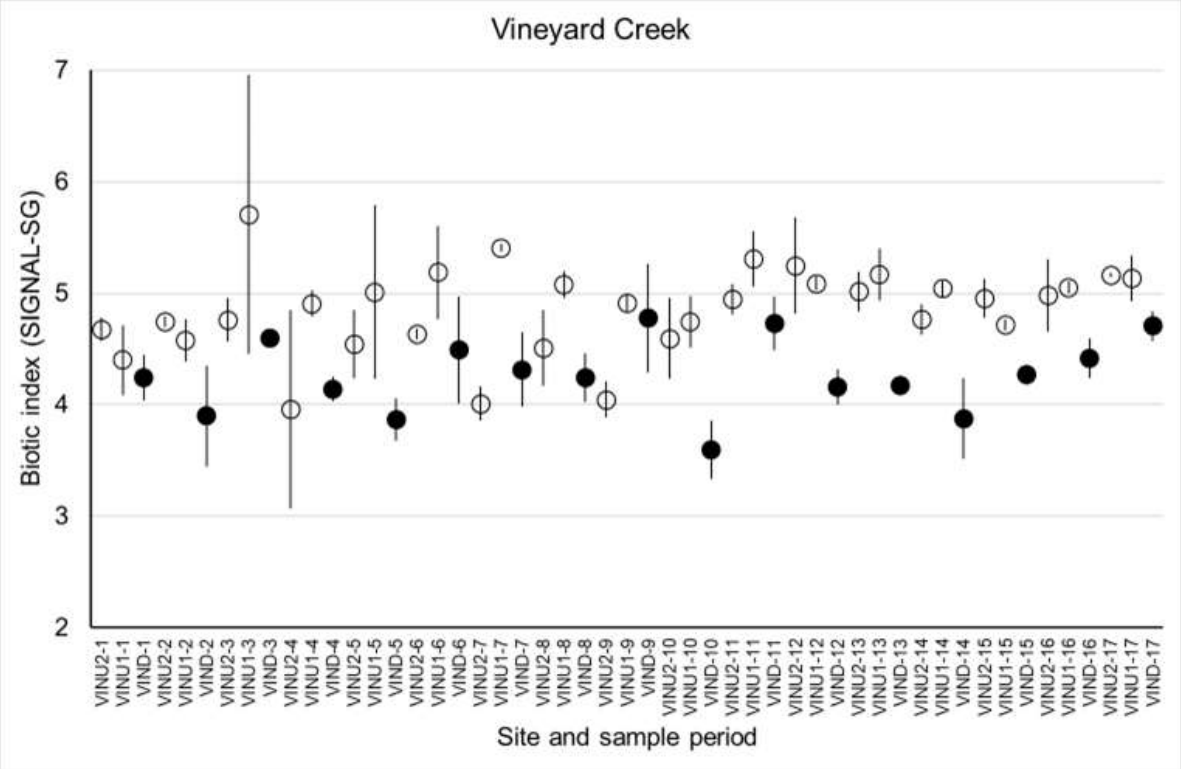


Figure 5-10: Vineyard Creek macroinvertebrate pool-edge water samples across 17 collection periods from three study sites

In a two-dimensional nMDS ordination plot of Vineyard Creek samples from across collection periods 10-17, a distinct group of downstream site samples was well separated from a cluster of upstream sites samples (Figure 5-11). The ordination pattern was confirmed in the corresponding tree diagram (dendrogram) from classification analysis as early divisions separated downstream site samples from upstream samples of both sites (Figure 5-12). This initial separation also occurred at a quite low similarity of 23% (Figure 5-12).

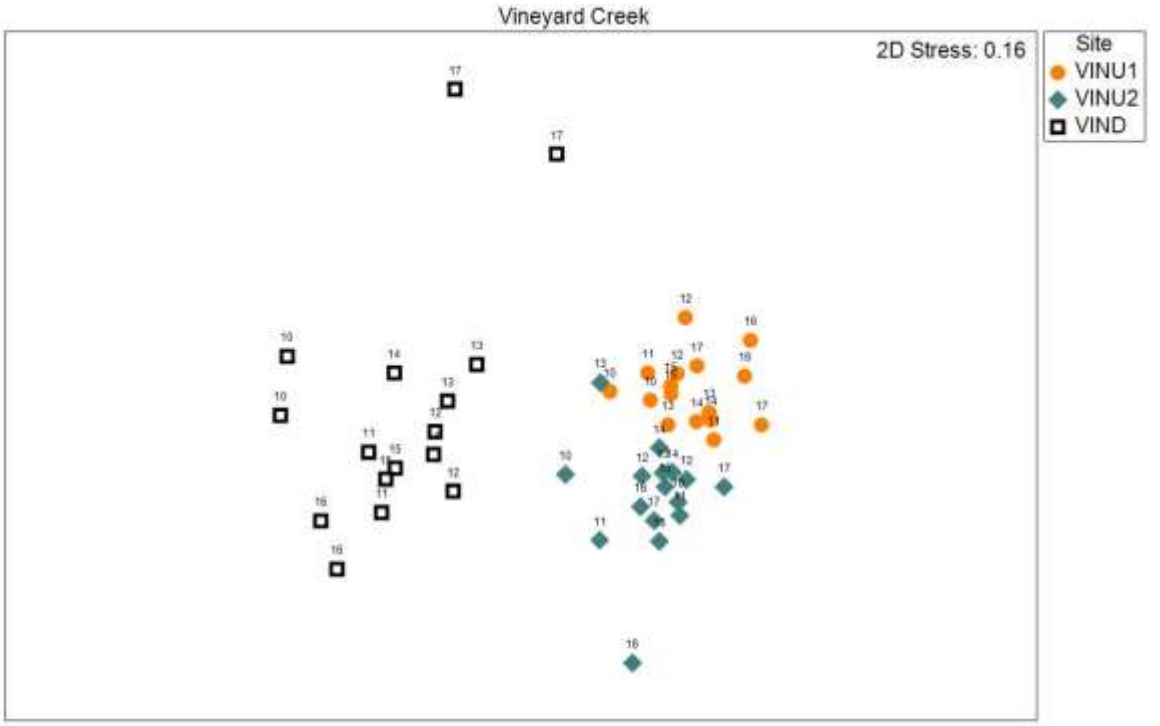


Figure 5-11: nMDS ordination plot of Vineyard Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

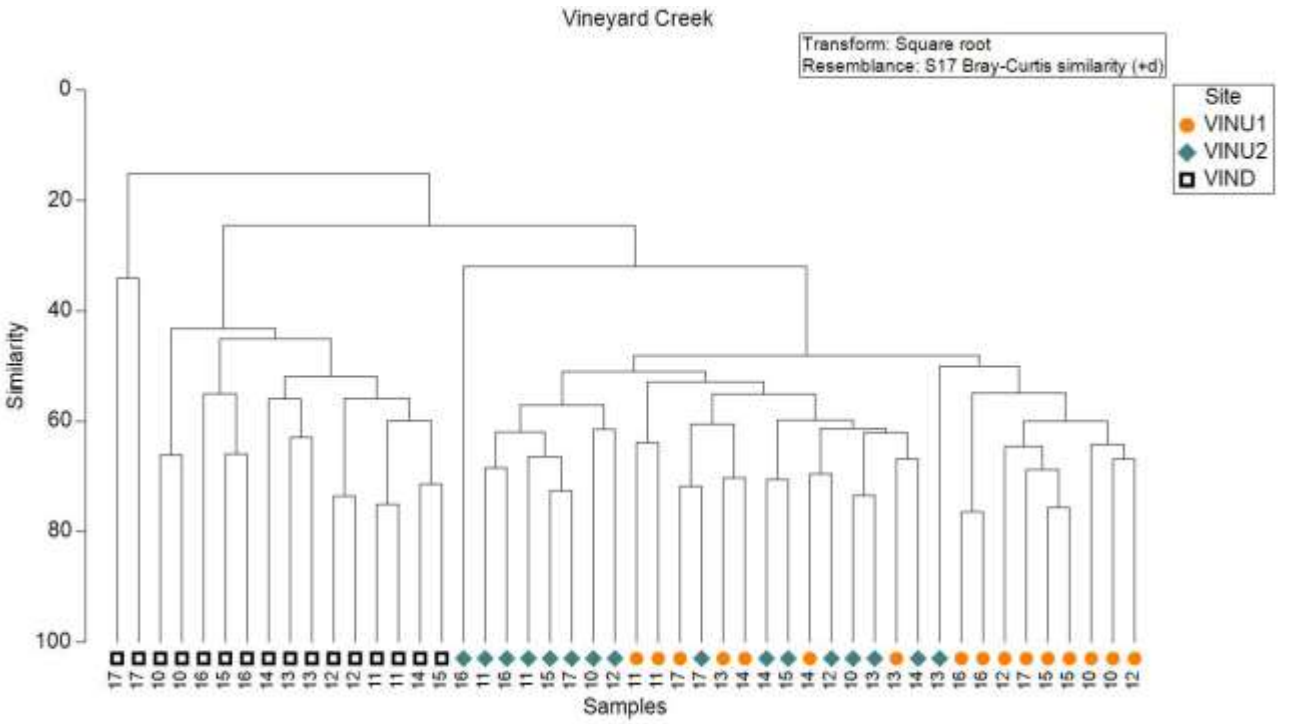


Figure 5-12: Tree diagram from classification analysis of MDS ordination plot of Vineyard Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

The clear separation of Vineyard Creek downstream site samples was also evident in the corresponding shade plot (Figure 5-13), where downstream samples displayed less diversity compared to the upstream site. The shade plot displayed the non-biting midge larvae, Chironimidae *Chironomus* that was commonly collected at the downstream site; while the damselfly larvae Argiolestidae *Austroargiolestes* and the water penny beetle larvae Psephenidae *Sclerocyphon* were commonly collected across the upstream sites. The corresponding SIGNAL-SG grades showed that the dominant taxa that occurred downstream have lower SIGNAL-SG grades such as *Chironomius* than those of the upstream site such as *Austroargiolestes* and *Sclerocyphon* (Figure 5-13), which is reflected in the separation of site SIGNAL scores displayed in (Figure 5-10).

The PERMDISP analysis indicated a dissimilar pattern of dispersion (between same site samples), although this is most probably predominantly influenced by the pair of downstream samples from period 17 after the major east coast low storm event under which WWO spilled across five days. The shade plot (Figure 5-13) shows that these downstream samples had few taxa, reflecting the major disturbance from this inflow and infiltration driven WWO event.

The ANOSIM test run on the factor 'Site' returned a high range value ($R = 0.66$; $P = 0.001$) confirming assemblage structure was distinct at each site. Pairwise tests indicated the two upstream (VINU1, VINU2) versus downstream (VIND) site comparisons had relatively high-level R-values (0.80, 0.75) close to the maximum R-value of 1. In contrast, comparison of upstream sites (VINU1 and VINU2) returned a mid-range R-value (0.51) at a level observed for natural differences between sites outlined in Section 5.2.3. These pairwise test results suggest clear differences in assemblage structure between the two upstream sites and the downstream site, that was situated below the atypical node of ERSs.

These results suggested downstream assemblage structure in Vineyard Creek was consistently different to that of the upstream site. WWO spills were the most likely influence on the downstream assemblage for these differences across collection periods 10 to 17, and a number of earlier collections as suggested by patterns in the ecological control chart (Figure 5-10). This outcome is supported by results of the dilutions study that identified ammonia was a contaminant of potential concern at the downstream Vineyard site (Section 4.4) and is also supported by outcomes of toxicity testing conducted at this location (Section 4.5.1).

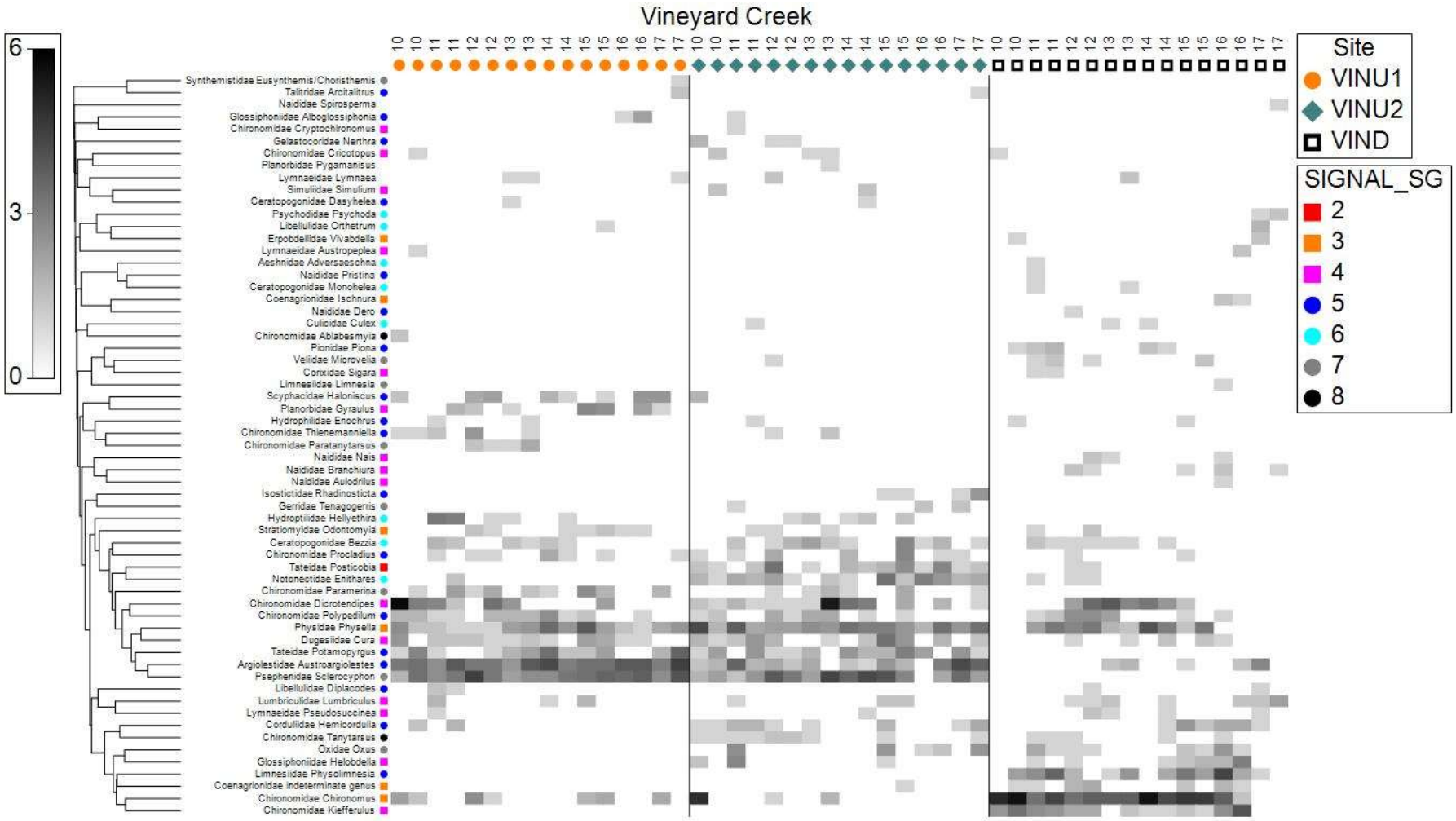


Figure 5-13: Shade plot of Vineyard Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

5.3.2 Buffalo Creek

Three sites were sampled in Buffalo Creek. The upstream sites were situated on each of the two arms of Buffalo Creek about 0.2 km for upstream site 1, and 0.5 km for upstream site 2, respectively above the downstream site.

The ecological control chart plot of SIGNAL-SG scores visualised stream health of Buffalo Creek through the WWOM study across February 2019–April 2021 by each 8-weekly collection (periods 3 to 17). This visual comparison illustrated downstream stream health varied marginally through time but was relatively consistently across the three sites within each collection period (Figure 5-14).

Statistical testing was conducted on data from collection periods 10 to 17. A comparison of the upstream-downstream SIGNAL-SG site scores indicated the downstream site had a mean score (BUFFD = 4.59 SD 0.22) that was similar to both the upstream sites (BUFFU1 = 4.58 SD 0.26, BUFFU2 = 4.36 SD 0.27), which supported the visual trends in the control chart plot (Figure 5-14). The accompanying ANOVA test of the factor 'site' was non-significant (P = 0.2064) and Brown and Forsythe's test for homogeneity of variance was non-significant (P =0.6779).

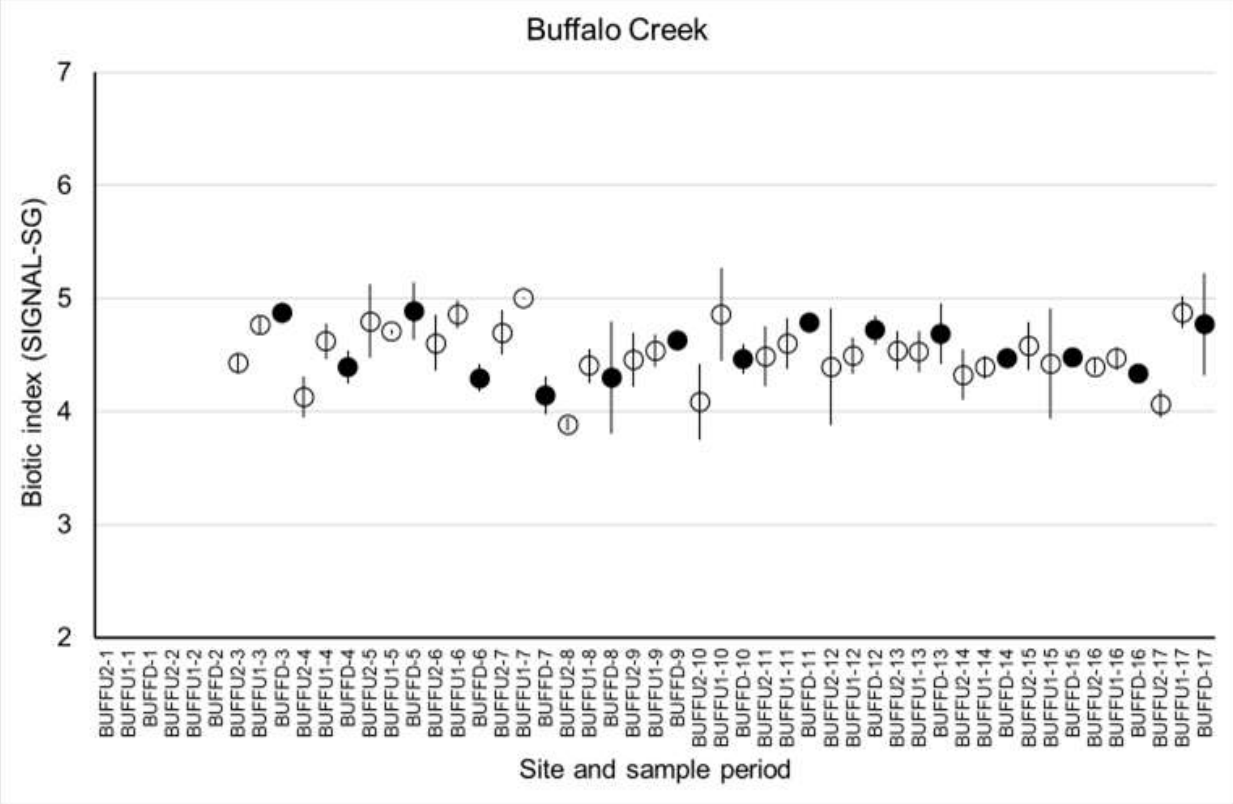


Figure 5-14: Buffalo Creek macroinvertebrate pool-edge water samples across 17 collection periods from three study sites

A two-dimensional nMDS ordination plot of Buffalo Creek samples from across collection periods 10-17, illustrated a somewhat interspersed cluster cloud of samples from across these three sites (Figure 5-15). This interspersed ordination pattern was confirmed in the corresponding tree diagram (dendrogram) from classification analysis as continued division of samples displayed a mixed pattern across sites (Figure 5-16). This initial separation was quite high at a similarity of 49% (Figure 5-16). The similar assemblage structure across Buffalo Creek site samples was evident in the corresponding shade plot (Figure 5-17).

The PERMDISP analysis indicated a similar pattern of dispersion (spacing between same site samples) for the three sites ($F = 0.557$, $P = 0.621$). This outcome suggests that the variability in taxonomic make-up of samples collected over time was at similar levels across sites through the periods tested (collections 10 to 17). This result then also implies subsequent results of ANOSIM tests are focused on community structure differences between sites.

The ANOSIM test run on the factor 'Site' returned a low value ($R = 0.18$; $P = 0.001$) confirming assemblage structure was relatively indistinct at each site. Large R values (close to unity) are indicative of complete separation of the groups, small R values (close to zero) imply little or no segregation. The R value itself is not unduly affected by the number of replicates in the two groups being compared; this is in stark contrast to its statistical significance, which is dominated by the group sizes (for larger number numbers of replicates, R values near zero could still be deemed significant, and conversely, few replicated could lead to R values to unity being classified as significant) (Clark et al., 2014). Pairwise tests indicated the two upstream (BUFFU1, BUFFU2) versus downstream (BUFFD) site comparisons also had relatively low-level R-values (0.01, 0.26) close to an R-value of 0 that indicates site samples are identical. A similar low level R value (0.26) was returned for the comparison of upstream sites (BUFFU1 and BUFFU2). These pairwise test results suggest minimal differences in assemblage structure between the two upstream sites and the downstream site, that was situated below a disparate ERS.

These results suggested that downstream assemblage structure in Buffalo Creek was not altered by WWO spills between collection periods 10 to 17. Rather stormwater influences were a consistent disturbance across these three urban stream sites.

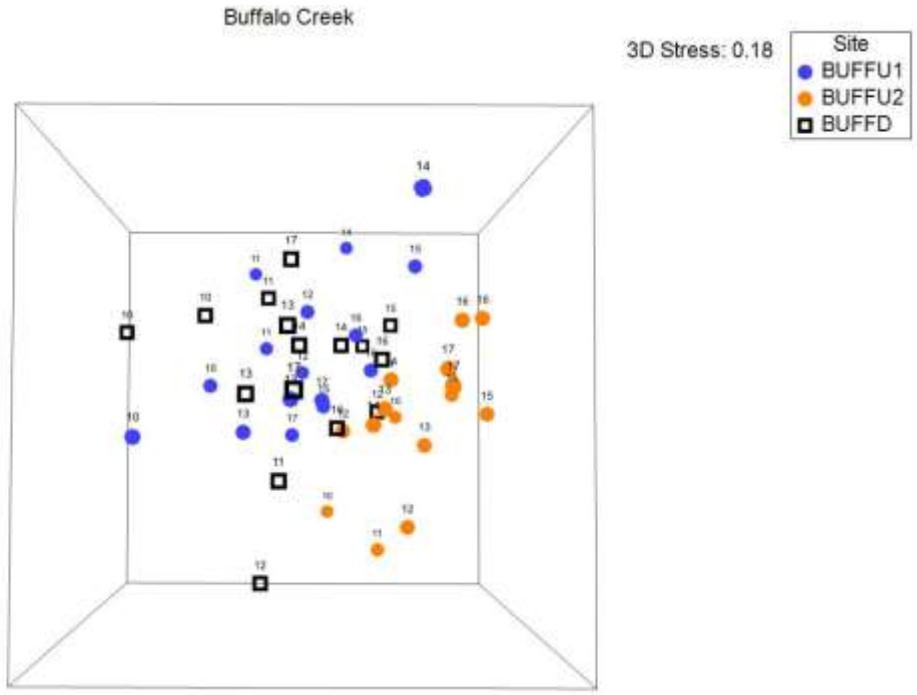


Figure 5-15: nMDS ordination plot of Buffalo Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

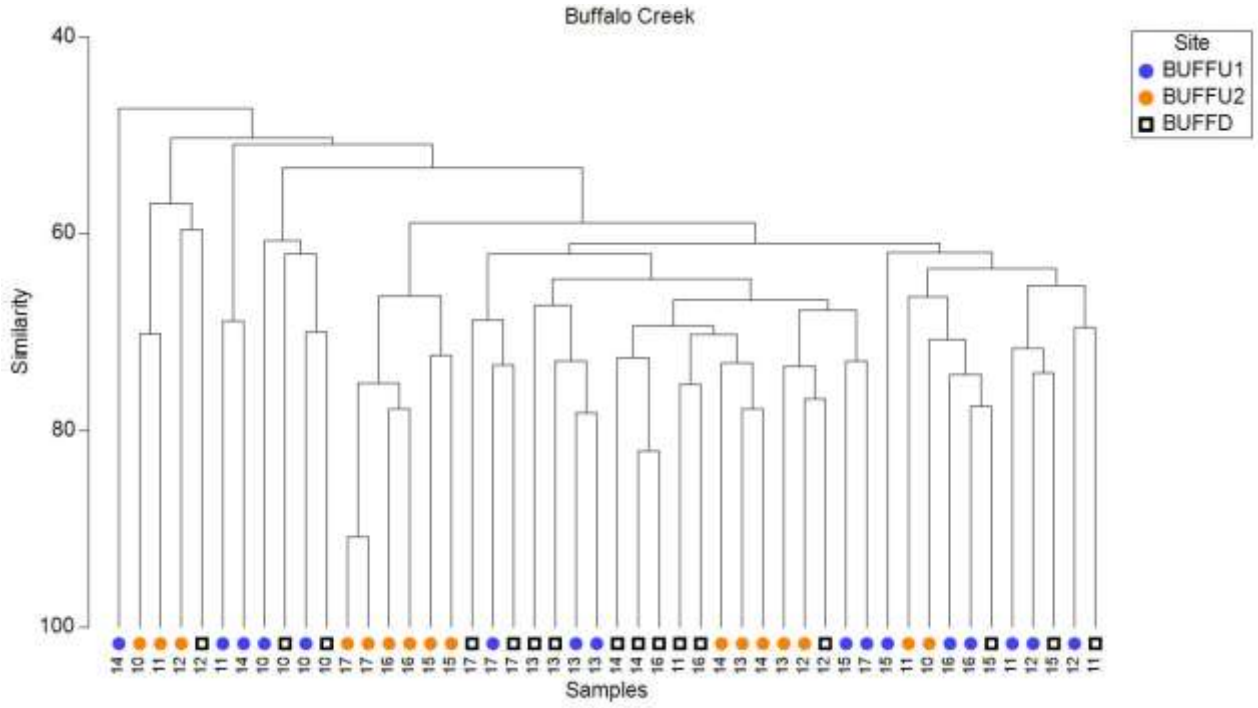


Figure 5-16: Tree diagram from classification analysis of MDS ordination plot of Buffalo Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

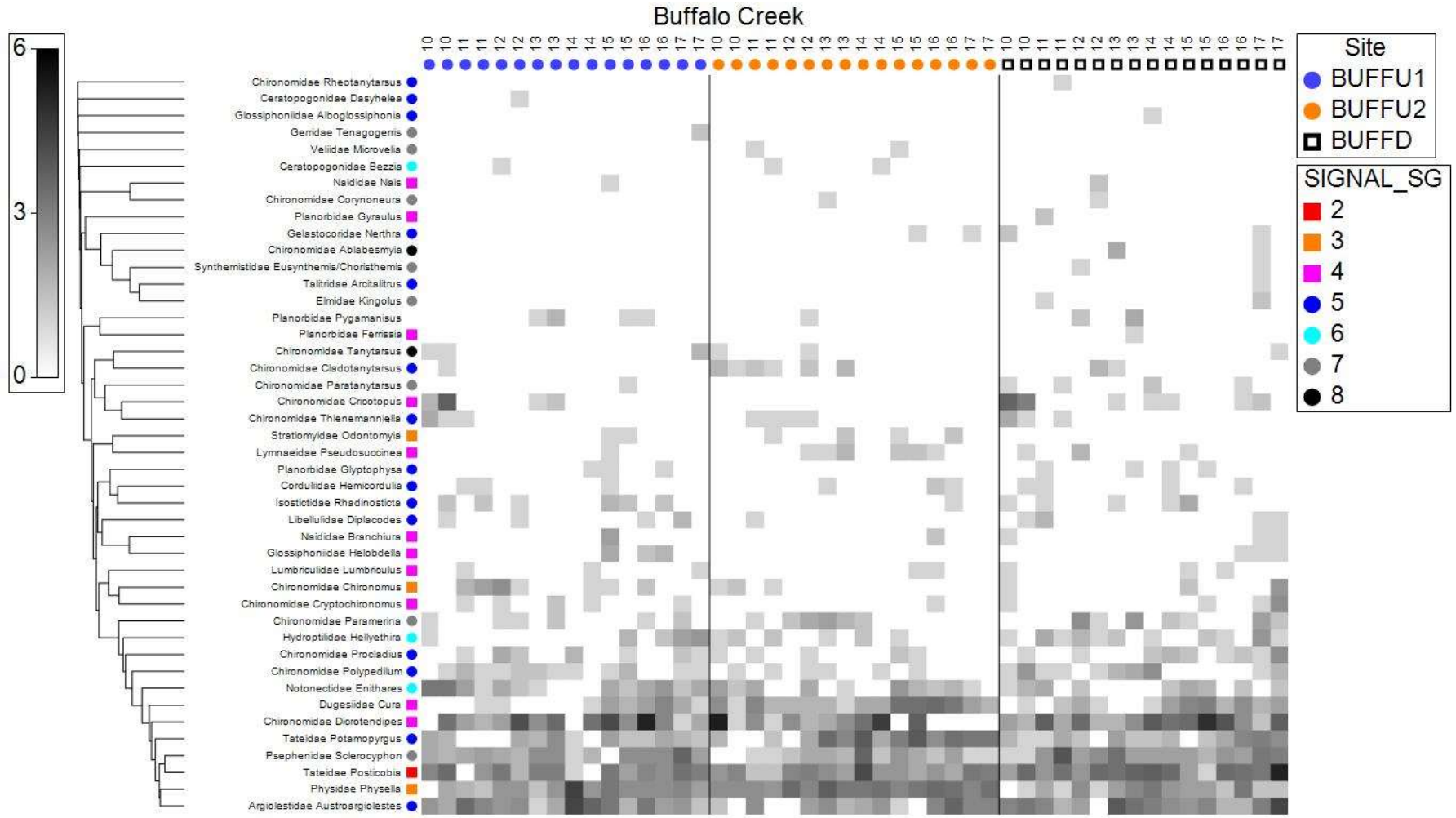


Figure 5-17: Shade plot of Buffalo Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

5.3.3 Avondale Creek

The ecological control chart plot of SIGNAL-SG scores visualised stream health of Avondale Creek through the WWOM study across October 2018–April 2021 by each 8-weekly collection (periods 1 to 17). This visual comparison illustrated that mean downstream stream health (represented by black infilled circle) was consistently lower than that of the upstream site in the post-drought La Niña period represented by collection periods 10 to 17 (March 2020 to April 2021) (Figure 5-18). The upstream and downstream sites were situated about 1.5 km apart (Figure 4-18).

Statistical testing was conducted on data from collection periods 10 to 17. A comparison of the upstream-downstream SIGNAL-SG site scores from a SNK multiple mean comparison test indicated that the downstream site had a statistically lower mean score (AVOND = 4.50 SD 0.30) than that of the upstream site (AVONU = 5.02 SD 0.24), which supported the visual trends in the control chart plot (Figure 5-18). The accompanying ANOVA test of the factor 'site' was significant ($P < 0.0001$) and Brown and Forsythe's test for homogeneity of variance was non-significant ($P = 0.7544$).

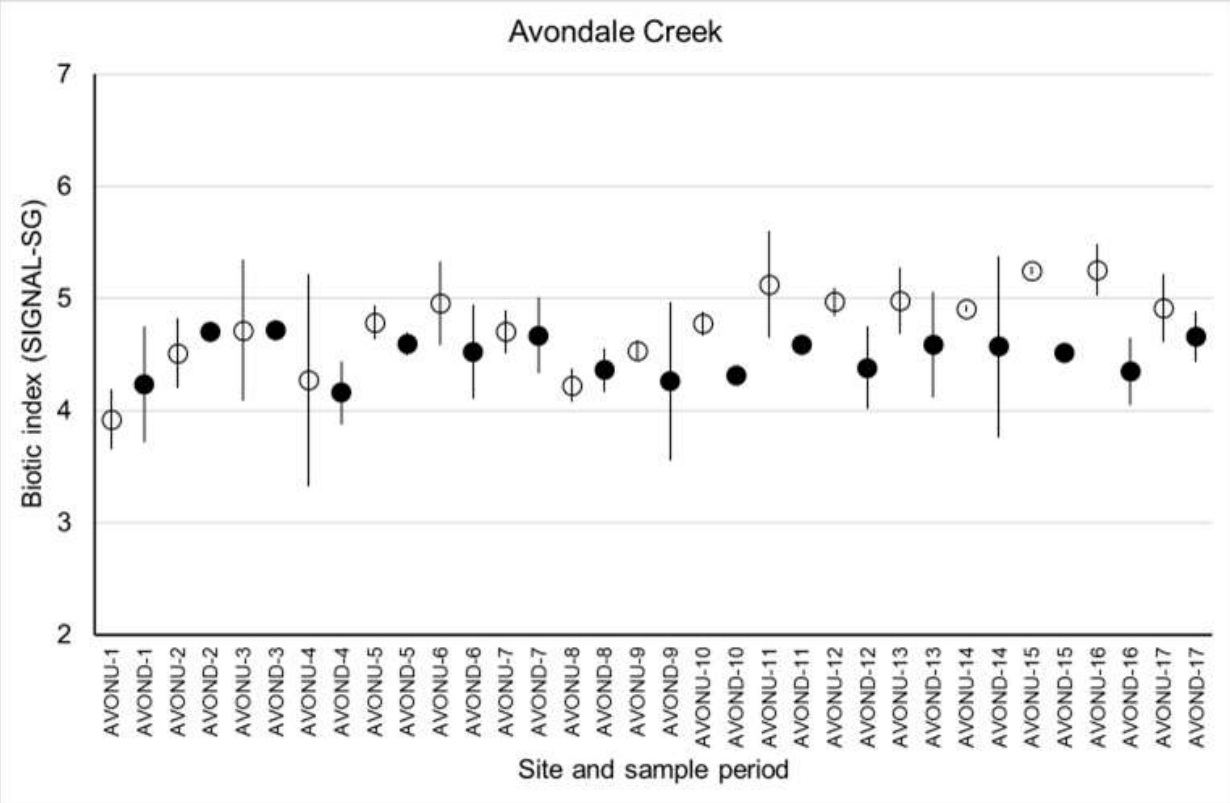


Figure 5-18: Avondale Creek macroinvertebrate pool-edge water samples across 17 collection periods from three study sites

In a two-dimensional nMDS ordination plot of Avondale Creek samples from across collection periods 10-17, a distinct group of downstream site samples was well-separated from a cluster of upstream sites samples (Figure 5-19). The ordination pattern was confirmed in the corresponding tree diagram (dendrogram) from classification analysis as the first division separated downstream site samples from upstream samples (Figure 5-20).

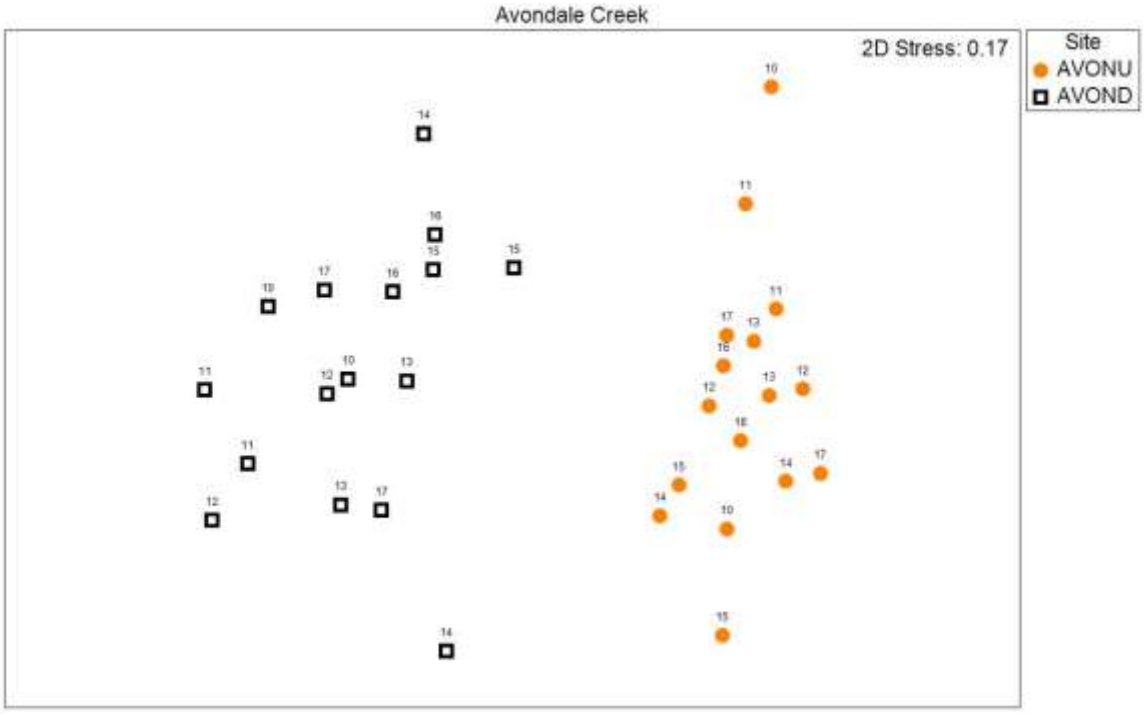


Figure 5-19: nMDS ordination plot of Avondale Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

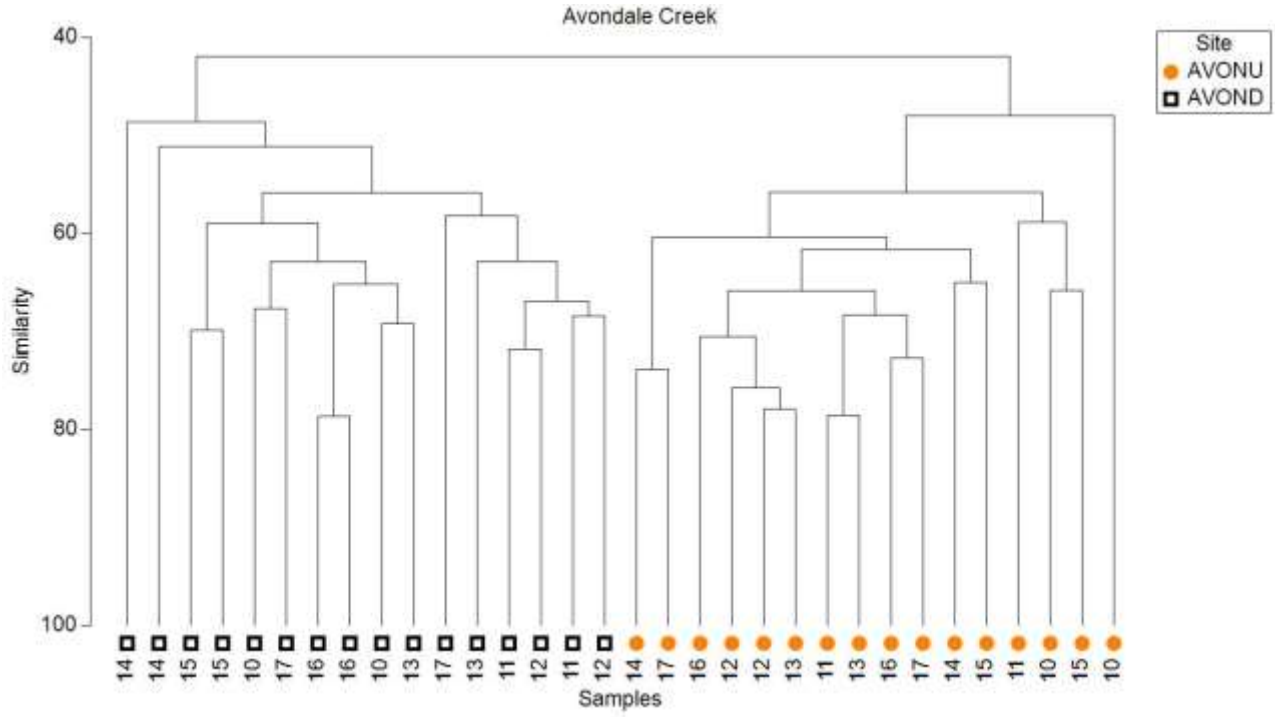


Figure 5-20: Tree diagram from classification analysis of MDS ordination plot of Avondale Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

The clear separation of upstream and downstream site samples illustrated in both the nMDS ordination plot and tree diagram are supported by the ANOSIM test run on the factor 'Site', which returned a high range value (R = 0.85; P = 0.001) confirming community structure was distinct at each site. This distinctness was also evident in the shade plot with only two taxa consistently collected at both sites (Figure 5-21). These taxa were the damselfly larvae Argiolestidae *Austroargiolestes* and the introduced snail Physidae *Physella*. Only two other taxa were commonly collected from the downstream Avondale Creek site, which were the Talitridae snails *Potamopyrgus* and *Posticobia*. Whereas from the upstream site the non-biting midge larvae Chironimidae *Procladius* and *Polypedilum* along with the flatworm Dugesiidae *Cura,* the water penny beetle larvae Psephenidae *Sclerocyphon* and the biting midge larvae Ceratopogonidae *Bezzia* were commonly collected across collection periods 10 to 17 (Figure 5-21). These consistently collected taxa contribute most to the separation of site SIGNAL scores displayed in (Figure 5-18).

The PERMDISP analysis indicated a similar pattern of dispersion (spacing between same site samples) for the two sites (F = 1.07, P = 0.331). This outcome suggests that the variability in taxonomic make-up of samples collected over time was at similar levels for both sites through the period tested.

These results suggested that downstream assemblage structure in Avondale Creek was consistently altered across collection periods 10 to 17 either by the observed network fault (Section 4.7, Figure 4-22) and or from WWO spills.

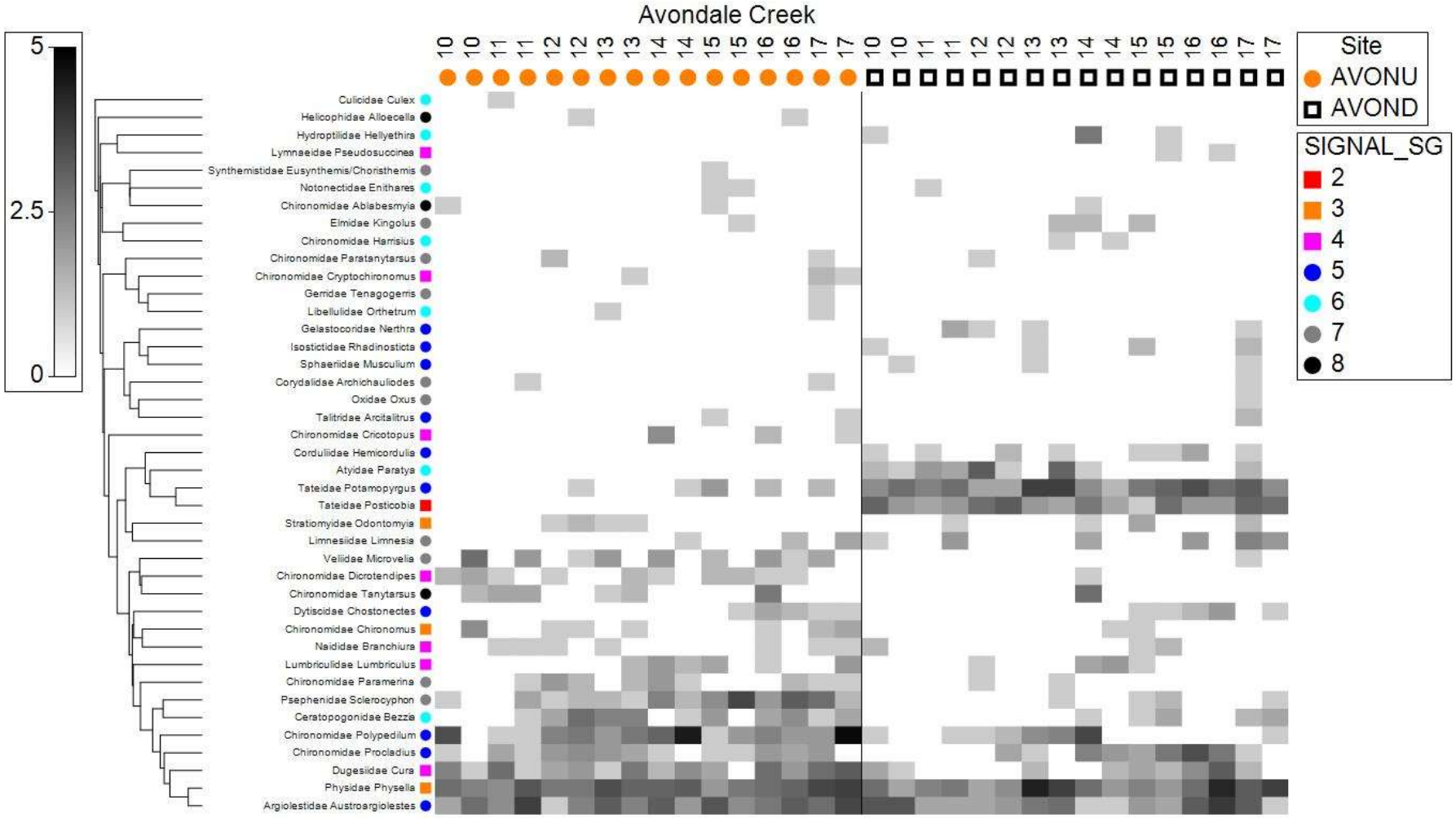


Figure 5-21: Shade plot of Avondale Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

5.3.4 Girraween Creek

Upstream and downstream sites are about 2 km apart and were sampled on Girraween Creek (Figure 4-18) with ERSs located between these sites as described in Section 4.7. The ecological control chart plot of SIGNAL-SG scores visualised stream health of Girraween Creek through the WWOM study across October 2018–April 2021 by each 8-weekly collection (periods 1 to 17). This visual comparison illustrated that downstream stream health was lower than that of the upstream site on 11 of 17 collection occasions (Figure 5-22). Disturbance at both sites appears evident from the major east-coast low storm events of February 2020 and March 2021 in results from collection periods 10 and 17.

Statistical testing was conducted on data from collection periods 10 to 17. A comparison of the upstream-downstream SIGNAL-SG site scores from a SNK multiple mean comparison test indicated that the downstream site had a statistically lower mean score (GIRD = 4.09 SD 0.26) than that of both the upstream sites (GIRU = 4.60 SD 0.41), which supported the visual pattern in the control chart plot (Figure 5-22). The accompanying ANOVA test of the factor 'site' was significant (P = 0.0002) and Brown and Forsythe's test for homogeneity of variance was non-significant (P =0.1973).

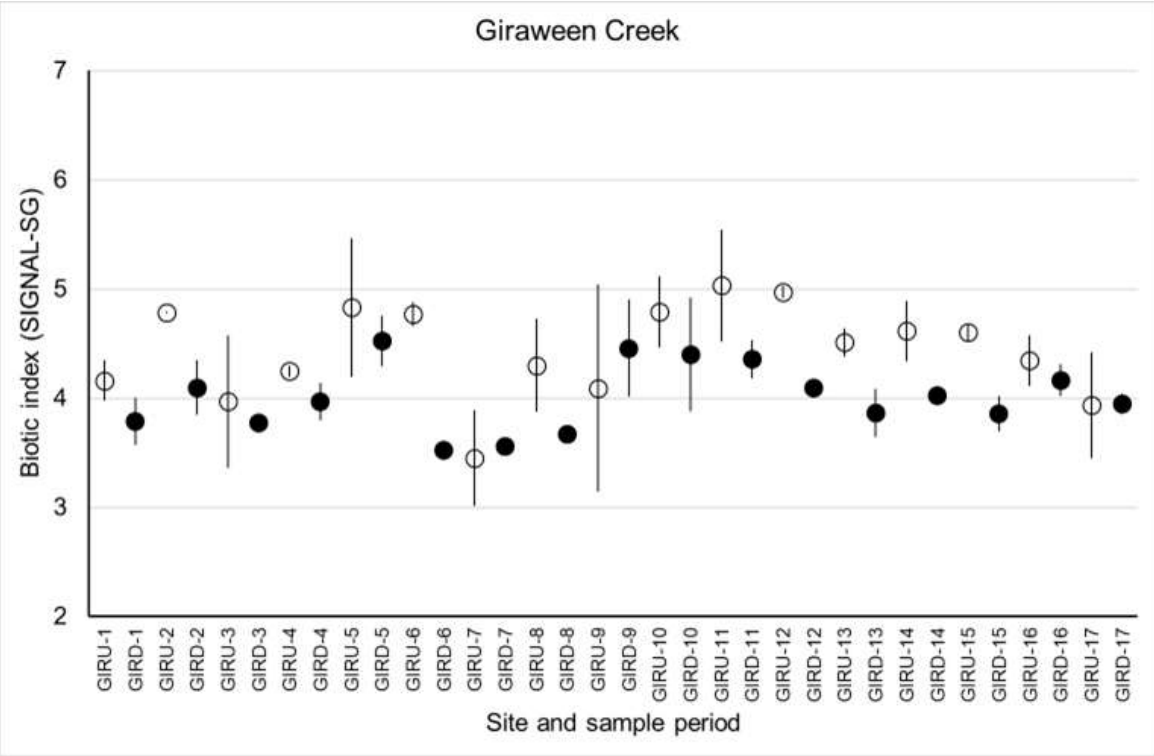


Figure 5-22: Girraween Creek macroinvertebrate pool-edge water samples across 17 collection periods from three study sites

In a two-dimensional nMDS ordination plot of Girraween Creek samples from across collection periods 10-17 a distinct group of downstream site samples was well separated from a cluster of upstream site samples (Figure 5-23), with the exception of period 17 upstream site samples. The ordination pattern was confirmed in the corresponding tree diagram (dendrogram) from classification analysis as the first division separated all downstream site samples and the two upstream site samples from period 17 from all other upstream samples (Figure 5-24). This initial separation also occurred at a somewhat low similarity of 33% (Figure 5-24).

PERMDISP analysis indicated a dissimilar pattern of dispersion (spacing between same site samples) for the two sites (F = 15.36, P = 0.001). The samples at the downstream site were relatively tightly spaced in the ordination plot (Figure 5-23), while a broader spread of upstream sites samples was evident within this ordination plot. This indicates that the assemblage structure of the upstream sites was less consistent though time than that of the downstream site. This variability in assemblage structure of the upstream site was evident in the shade plot as was the stability of the downstream site assemblage (Figure 5-25).

The ANOSIM test run on the factor ‘Site’ returned a moderate R value (R = 0.67; P = 0.001) indicating assemblage structure differed at each site.

These results suggested that downstream assemblage structure in Girraween Creek was distinct from that of the upstream site. Assessment with the biotic index SIGNAL-SG confirmed that the downstream assemblage was impaired compared to the upstream assemblage (Figure 5-22). Although in the upstream reach of this very small urban stream greater fluctuation in the macroinvertebrate assemblage was evident, which appeared to relate to larger storm events. Despite that disturbance, marginally higher SIGNAL-SG grade taxa on average occurred at the upstream site (Figure 5-25). This suggests that WWOs from the four ERSs identified as spilling into this stream reach (Section 4.7) may have influenced the downstream assemblage.

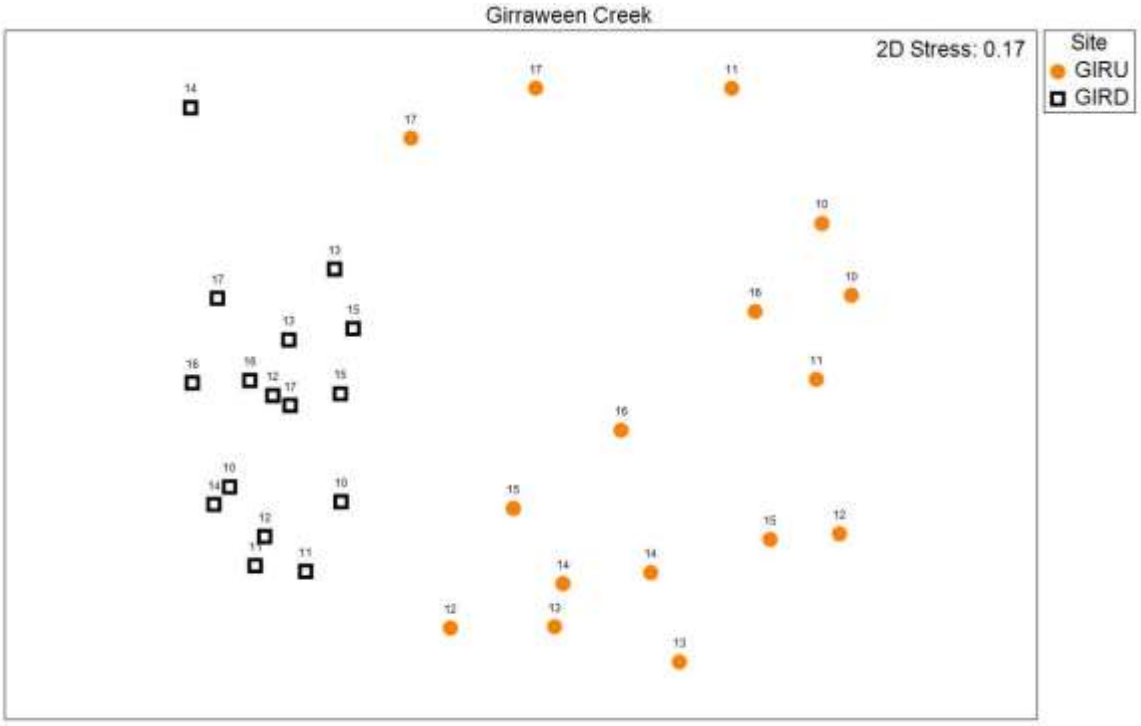


Figure 5-23: nMDS ordination plot of Girraween Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

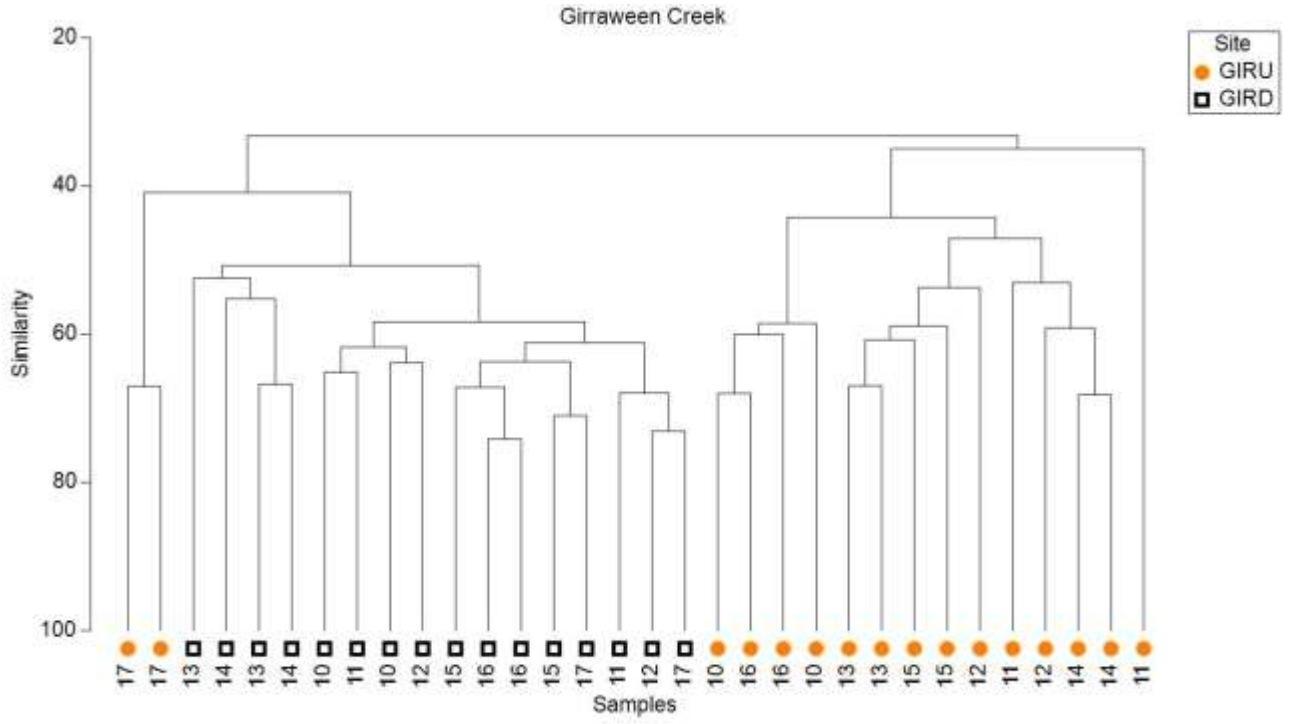


Figure 5-24: Tree diagram from classification analysis of MDS ordination plot of Girraween Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

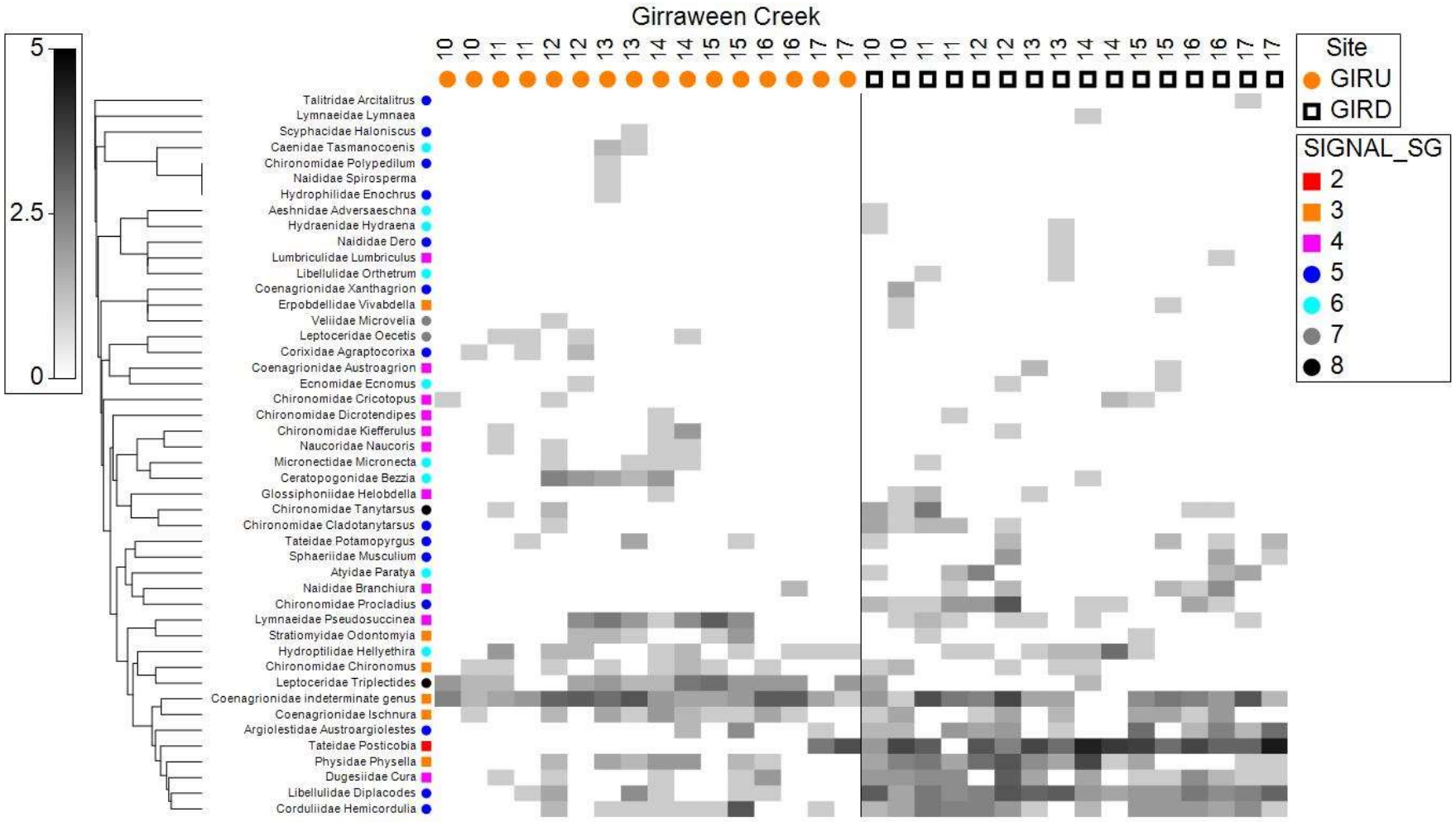


Figure 5-25: Shade plot of Girraween Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

5.3.5 Kittys Creek

Two sites were sampled on Kittys Creek about 0.4 km apart (Figure 4-18). The ecological control chart plot of SIGNAL-SG scores visualised stream health of Kittys Creek through the WWOM study across October 2018–April 2021 by each 8-weekly collection (periods 1 to 17). This visual comparison illustrated downstream stream health was consistently lower than that of both upstream sites on 12 of the 17 collection occasions, although fluctuations in the level of stream health were apart across the broader 17 collection periods.

To be consistent with assessment of other creeks, statistical testing was conducted upon data from collection periods 10 to 17. A comparison of the upstream-downstream SIGNAL-SG site scores from a SNK multiple mean comparison test indicated that the downstream site had a statistically lower mean score (KITD = 4.50 SD 0.44) than that of the upstream site (KITU = 5.11 SD 0.20), which supported the visual trends in the control chart plot (Figure 5-26). The accompanying ANOVA test of the factor 'site' was significant ($P < 0.0001$) and Brown and Forsythe's test for homogeneity of variance was also significant ($P = 0.0159$). In this case where the dataset remained heteroscedastic (Brown and Forsythe's test: $P < 0.05$), the level of significance was set at $P < 0.01$ (Chariton et al., 2011). The three levels of downstream scores across period 10 to 17 influenced this outcome.

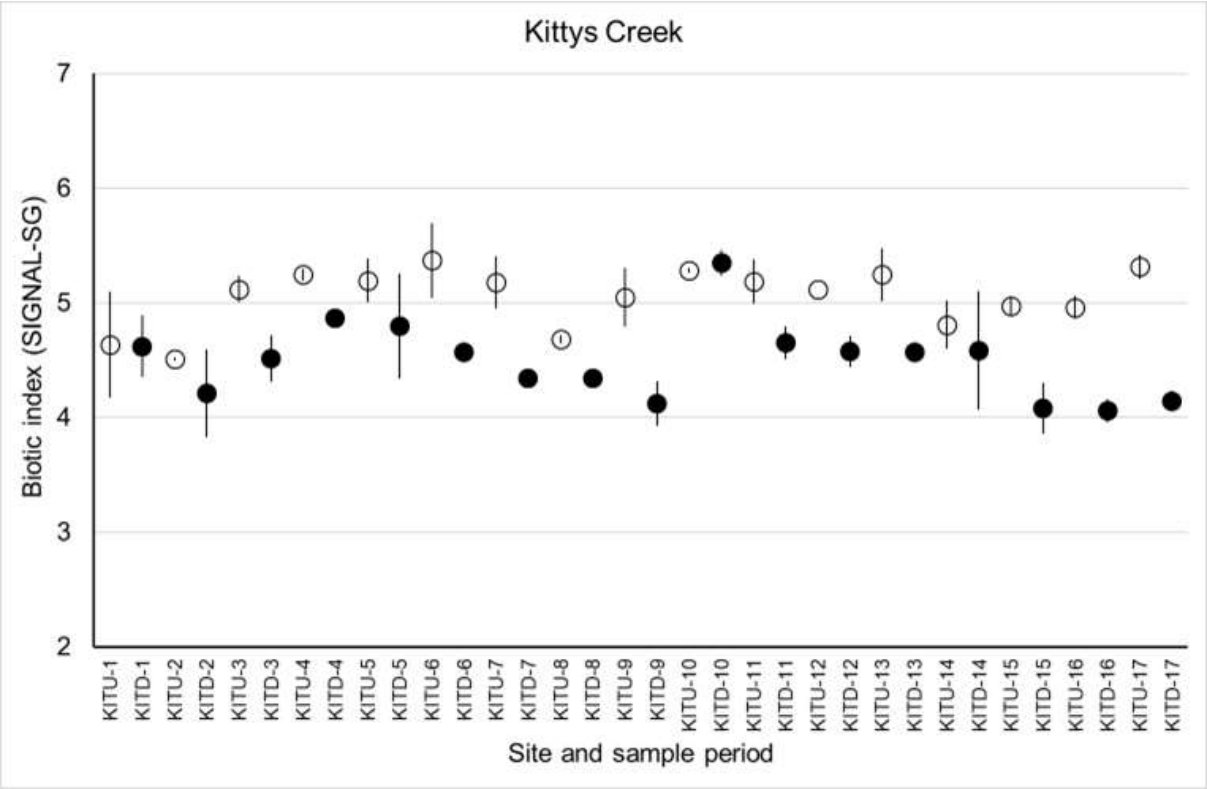


Figure 5-26: Kittys Creek macroinvertebrate pool-edge water samples across 17 collection periods from three study sites

In a two-dimensional nMDS ordination plot of Kittys Creek samples from across collection periods 10-17, a distinct group of downstream site samples was well-separated from a group of upstream

sites samples (Figure 5-27). The ordination pattern was confirmed in the corresponding tree diagram (dendrogram) from classification analysis as the first division separated all downstream site samples along with a replicate upstream site sample from collection period 13 from all other upstream samples (Figure 5-28).

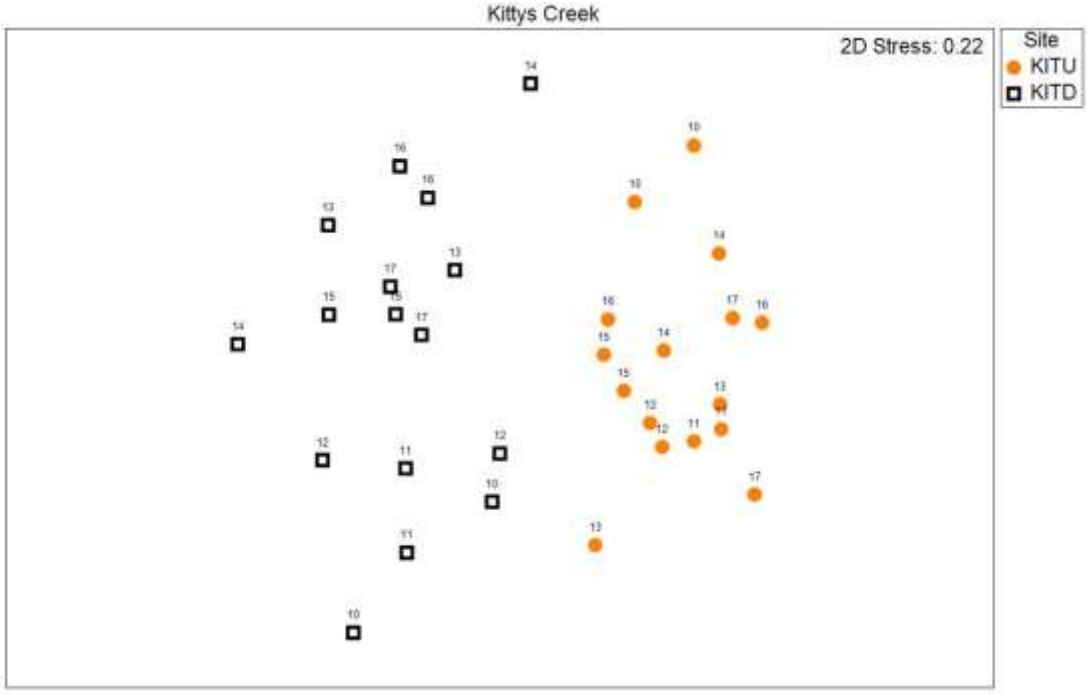


Figure 5-27: nMDS ordination plot of Kittys Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

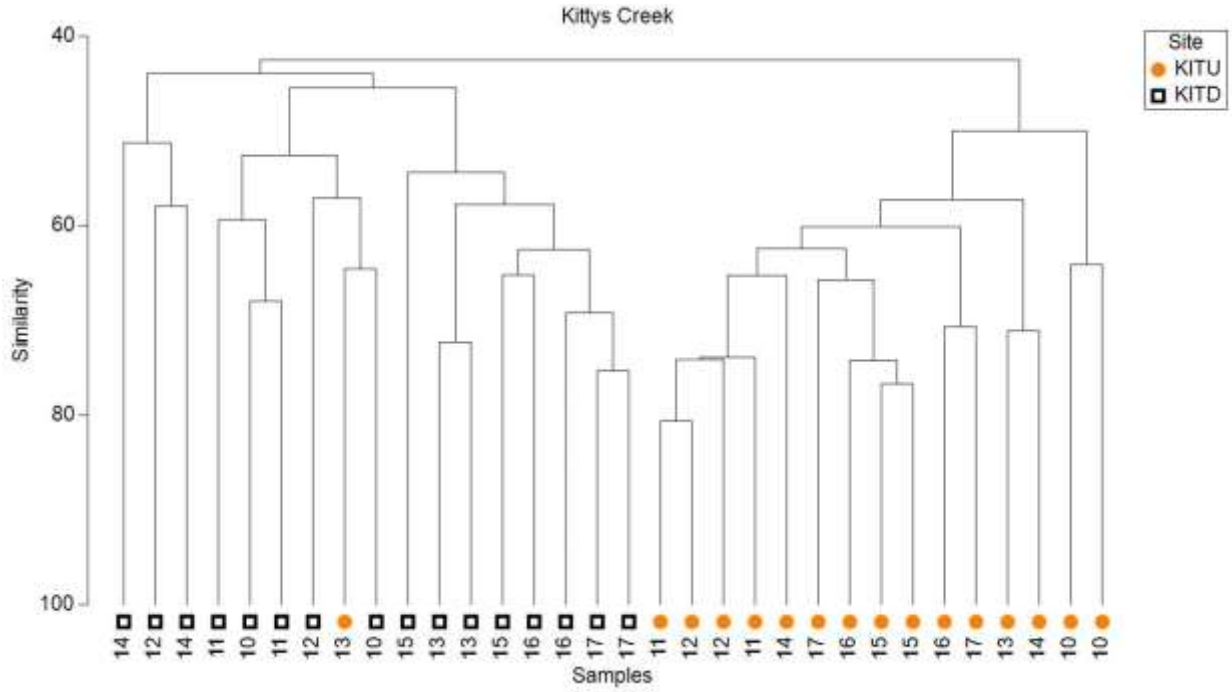


Figure 5-28: Tree diagram from classification analysis of MDS ordination plot of Kittys Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

The PERMDISP analysis indicated multivariate dispersion (differing spacing between same site samples) for the two sites (F = 7.51, P = 0.015). This result then also implies subsequent results of ANOSIM tests are focused on assemblage structure differences between sites and between collection periods. Within the ordination plot, samples from collections 10, 11 and 13 were slightly more distant that other downstream samples (Figure 5-27). The ANOSIM test run on the factor 'Site' returned a mid-range value (R = 0.64; P = 0.001) indicating community structure was distinct at each site. As these sites were about 0.4 km apart the returned R value is less likely to indicate a natural substrate difference in assemblage structure between sites.

Fewer taxa were consistently collected from the downstream site of Kittys Creek compared to the upstream site as illustrated in the shade plot (Figure 5-29). Also evident in the shade plot is the consistent collection of taxa, for example the dragonfly larvae Corduliidae *Hemicordulia* across collection periods 10 to 13, that probably influenced the slight separation evident in the ordination plot (Figure 5-27) and contributed to the multivariate dispersion test outcome.

The above analysis results clearly illustrate a difference between upstream and downstream assemblage structure in Kittys Creek with statistical testing confirming this difference across assessed collection periods 10 to 17. Analysis with the biotic index SIGNAL-SG suggested that the macroinvertebrate assemblage of the downstream site was more often impaired across the 17 collection periods and rarely attained the same level of stream health as measured for the upstream site (Figure 5-26). As this pair of sites on Kittys Creek were situated relatively close together (Figure 4-18) and around the study ERS, a natural difference in mesohabitat structure was unlikely at this close spatial scale. This then suggests that WWOs from this apparently oversized ERS spill WWOs in excess of the dilutions possible in the receiving waters of Kittys Creek (Section 4.7) and have potentially influenced this outcome.

Reference

Chariton A. A., Maher W. A., Roach A. C. 2011. Recolonisation of translocated metal-contaminated sediments by estuarine microbenthic assemblages. Ecotoxicology, 20, 706-771.

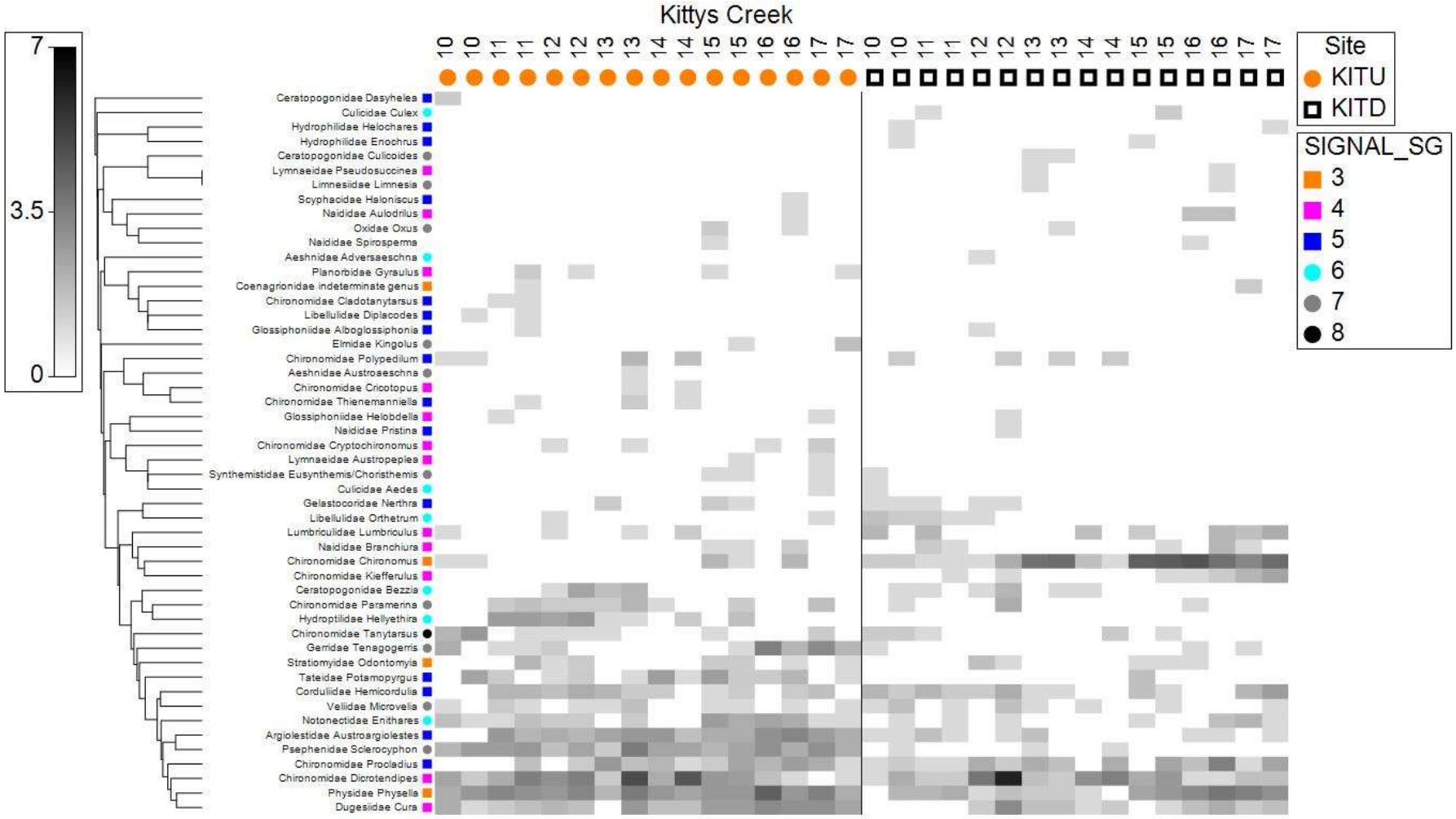


Figure 5-29: Shade plot of Kittys Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

5.3.6 Frenchs Creek

The pair of sites sampled on Frenchs Creek were about 0.5 km apart (Figure 4-18). The ecological control chart plot of SIGNAL-SG scores visualised stream health of Frenchs Creek through the WWOM study across October 2018 – April 2021 by each 8-weekly collection (periods 1 to 17). This visual comparison illustrated downstream stream health was consistently lower than that of the upstream site in 12 of the 17 collection periods (Figure 5-30).

To be consistent with assessment of other creeks, statistical testing was conducted upon data from collection periods 10 to 17. A comparison of the upstream-downstream SIGNAL-SG site scores from a SNK multiple mean comparison test indicated that the downstream site had a statistically lower mean score (FREND = 5.25 SD 0.24) than that of the upstream site (FRENU2 = 5.79 SD 0.22), which supported the visual trends in the control chart plot (Figure 5-30). The accompanying ANOVA test of the factor 'site' was significant ($P < 0.0001$) and Brown and Forsythe's test for homogeneity of variance was non-significant ($P = 0.8471$).

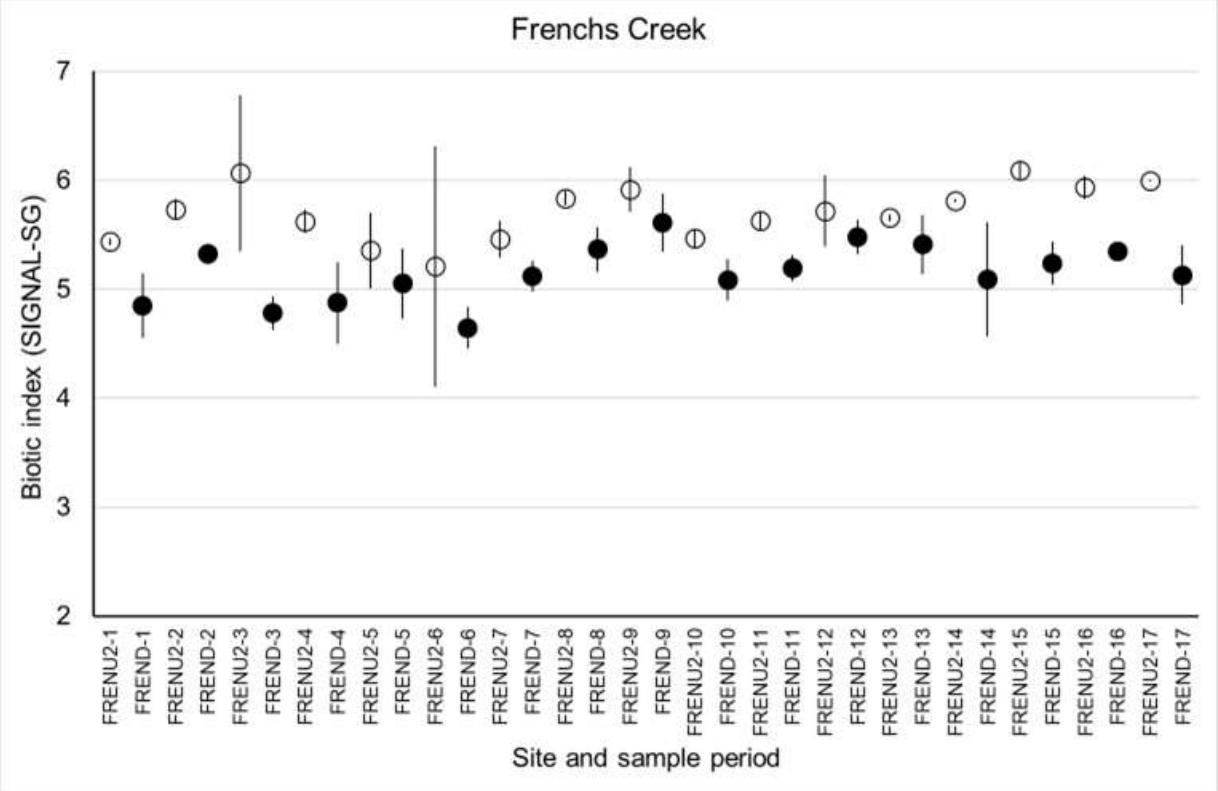


Figure 5-30: Frenchs Creek macroinvertebrate pool-edge water samples across 17 collection periods from three study sites

In a two-dimensional nMDS ordination plot of Frenchs Creek samples from across collection periods 10-17, a distinct group of downstream site samples was well separated from a tighter cluster of upstream sites samples (Figure 5-31). The ordination pattern was confirmed in the corresponding tree diagram (dendrogram) from classification analysis as the first division separated downstream site samples from upstream site samples (Figure 5-32).

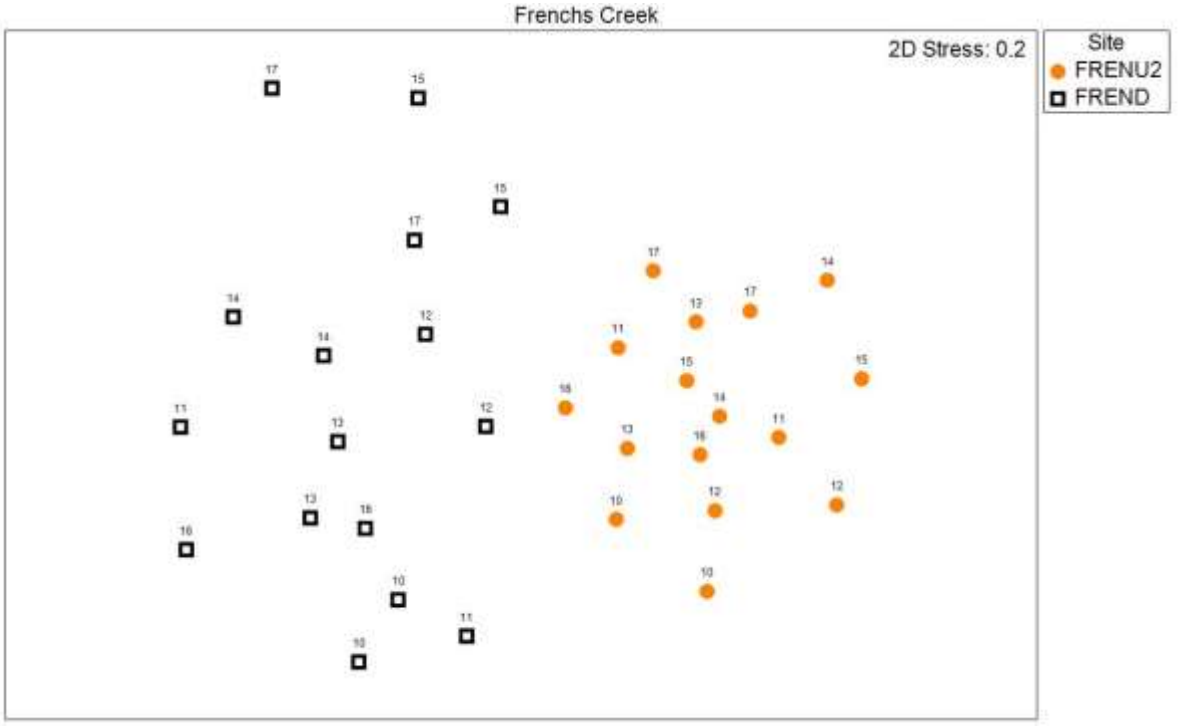


Figure 5-31: nMDS ordination plot of Frenchs Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

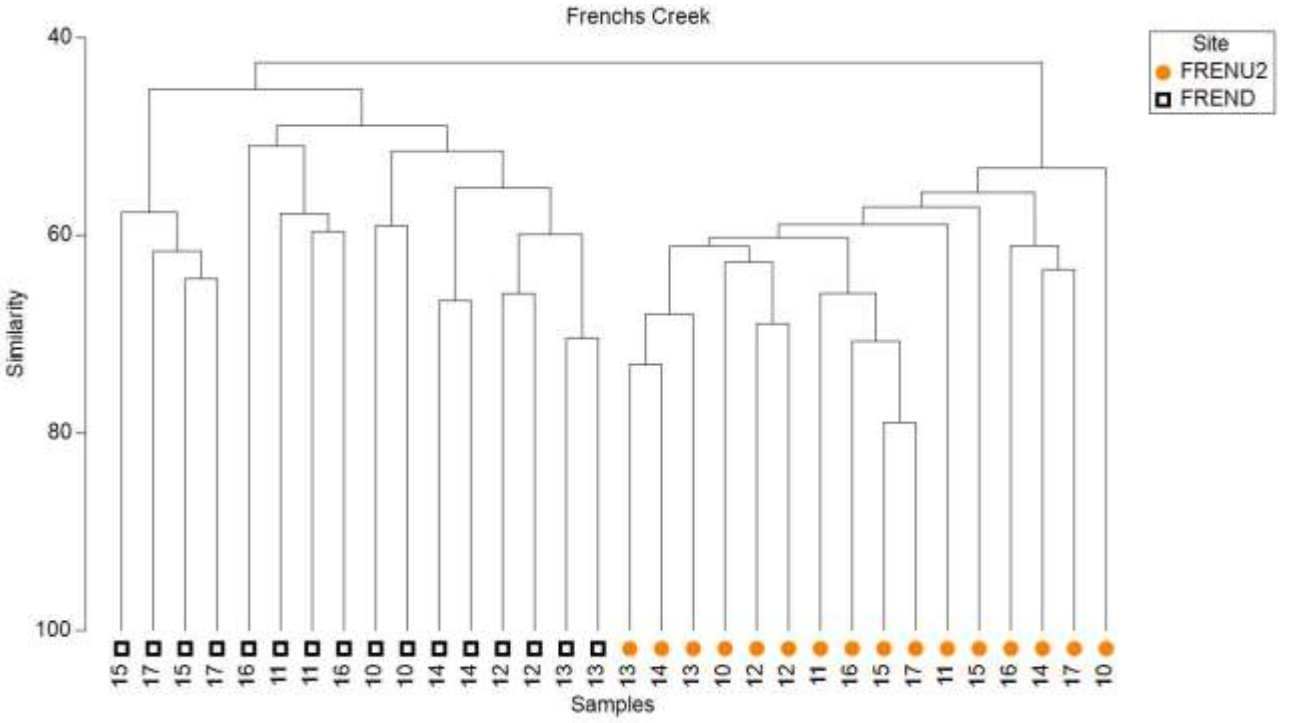


Figure 5-32: Tree diagram from classification analysis of MDS ordination plot of Frenchs Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

The PERMDISP analysis indicated a dissimilar pattern of dispersion (spacing between same site samples) for the two sites (F =15.55, P = 0.001). This outcome suggests a variability in the taxonomic make-up of samples collected over time. A greater separation (spacing) between downstream samples was evident in the ordination plot (Figure 5-31), and primarily influence this test outcome. Although for both sites there were taxa collected across a number of adjacent collection periods and then not collected in other periods as illustrated in the shade plot (Figure 5-33). The presence of multivariate dispersion in the dataset indicated that the ANOSIM test assessed both site differences and collection period differences in assemblage structure. The ANOSIM test run on the factor 'Site' returned a high range value (R = 0.70; P = 0.001) confirming assemblage structures had distinct characteristics at each site. As these sites were less than 0.5 km apart the returned R value is less likely to indicate a natural substrate difference in assemblage structure between sites.

Results from the above analysis suggested that downstream assemblage structure in Frenchs Creek was different to that of the upstream site across collection periods 10 to 17. Analysis with the biotic index SIGNAL-SG of Frenchs Creek suggested that the macroinvertebrate assemblage of the downstream site was more often impaired across the 17 collection periods and rarely obtained the same level of stream health as measured for the upstream site (Figure 5-30), which suggests impairment of the downstream assemblage. Similar to Kittys Creek, the pair of sites on Frenchs Creek were relatively close together (0.5 km) around the study ERS that spilt a potentially disproportionate volume of influent to this small urban stream (Section 4.7). This suggests that ammonia dilution from WWOs may have been inadequate in the receiving waters of Frenchs Creek.

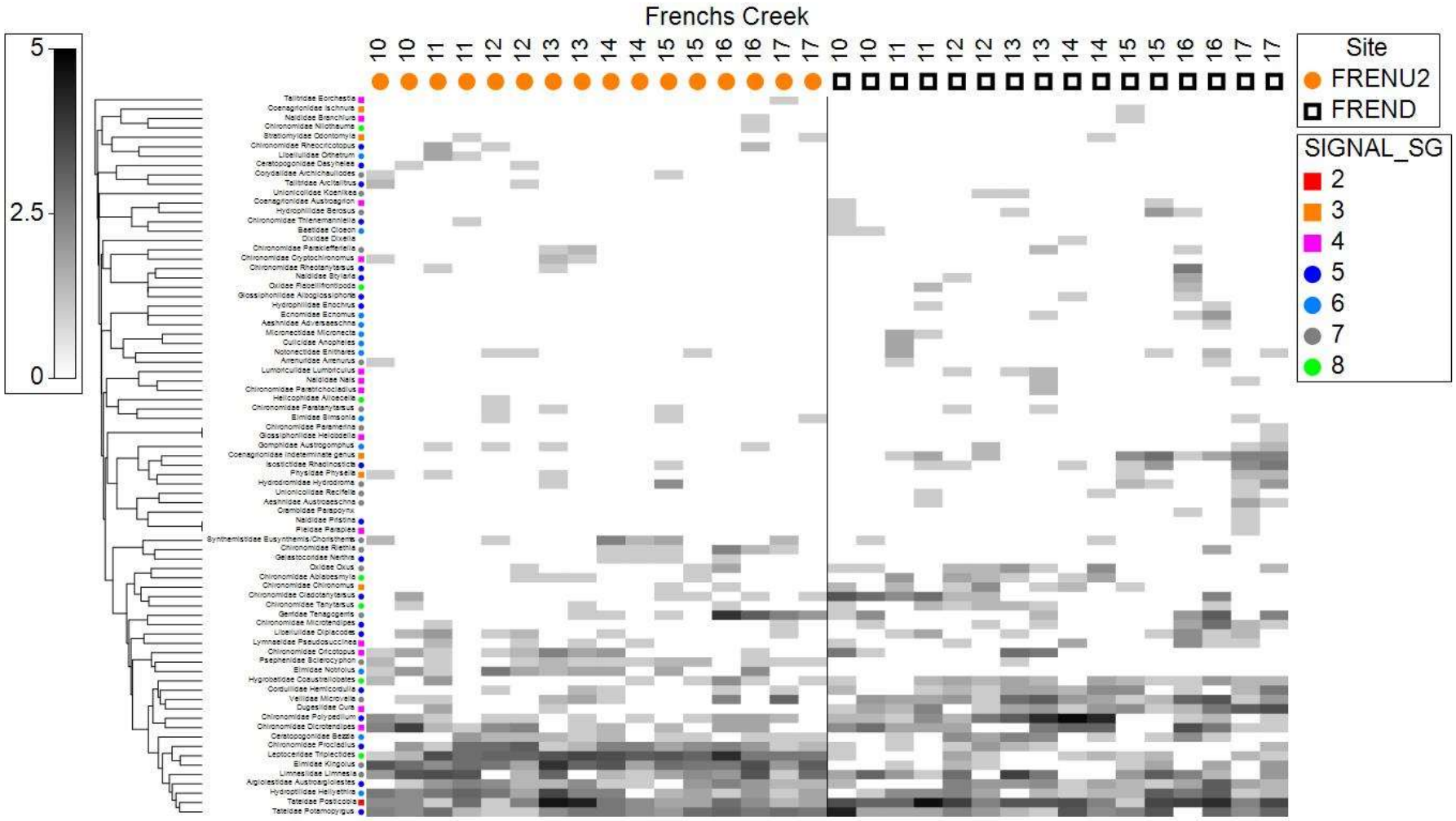


Figure 5-33: Shade plot of Frenchs Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

5.3.7 Darling Mills Creek system

Three sites were sampled in Darling Mills Creek system. The upstream sites were situated on each of two arms of Excelsior Creek. Upstream site 1 was 1.7 km above the downstream site while upstream site 2 was 1.8 km above the downstream site (Figure 5-34). The downstream site was located on Darling Mills Creek. This placement resulted in flows from Darling Mills Creek and tributaries of Bellbird Creek and Blue Gum Creek with a stream length of 10 km and that 0.4 km above the downstream site (Figure 5-34). This is reflected in catchment areas of 1.9 km^2 above each of the two upstream sites on the two arms of Excelsior Creek while the Darling Mills Creek and tributaries of Bellbird Creek and Blue Gum Creek have a catchment area of 15.4 km^2. This suggests that a greater stormwater exposure occurs at the downstream site.

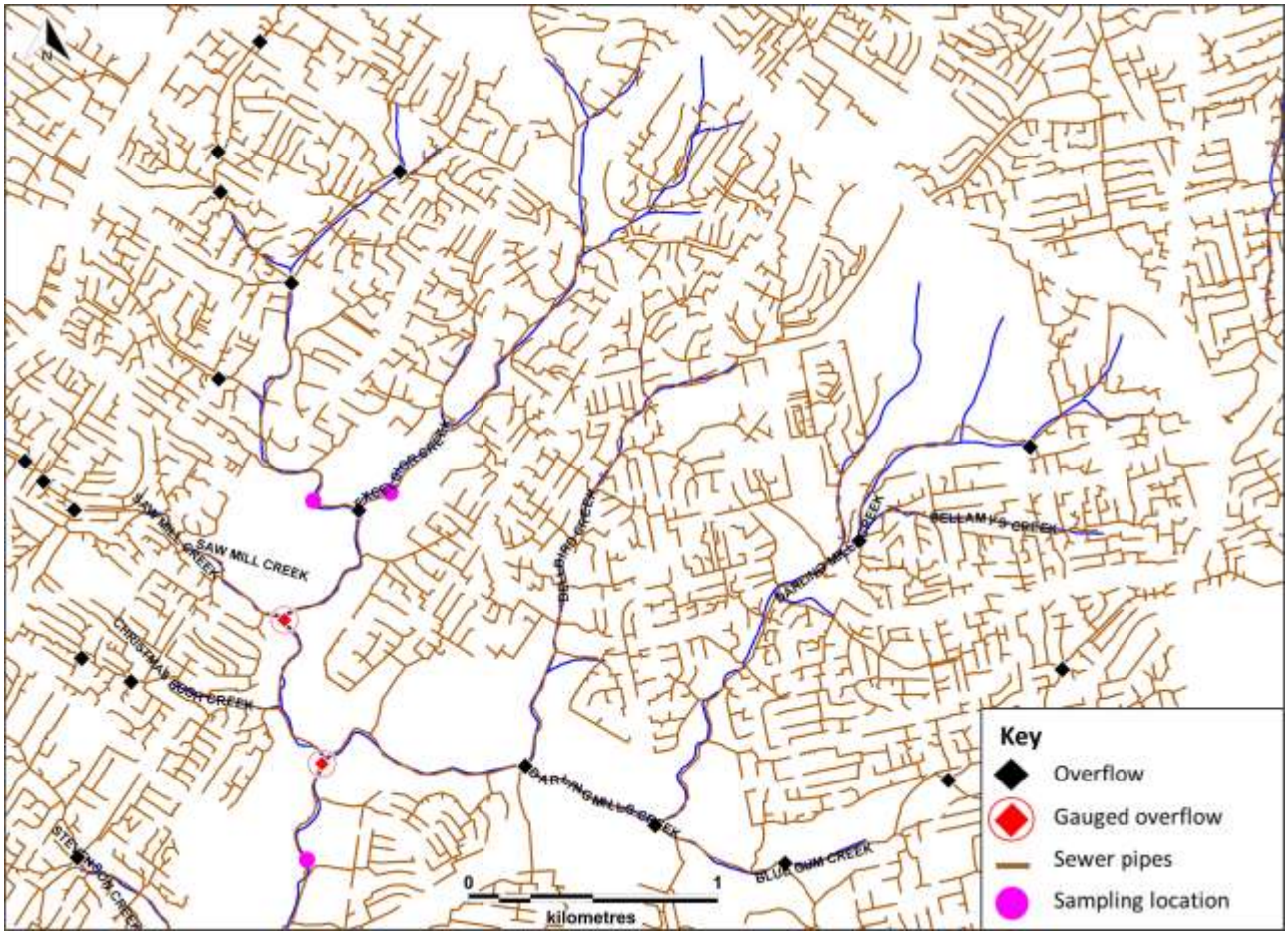


Figure 5-34: Location map of Darling Mills Creek system study sites

The ecological control chart plot of SIGNAL-SG scores visualised stream health of Darling Mills Creek system through the WWOM study across October 2018–April 2021 by each 8-weekly collection (periods 1 to 17). This visual comparison illustrated in the post-drought La Niña period upon return of frequent rainfall under collection periods 10 to 17 (March 2020 to April 2021) that three of these eight collection to have lower mean SIGNAL-SG scores for the downstream site (Figure 5-35).

Statistical testing was conducted upon data from collection periods 10 to 17. A comparison of the upstream-downstream SIGNAL-SG site scores from a SNK multiple mean comparison test indicated that the downstream site had a statistically lower mean score (DARLD = 5.24 SD 0.24) than that of both the upstream sites (EXCELU1 = 5.82 SD 0.16, EXCELU2 = 5.44 SD 0.24), which supported the visual trends in the control chart plot (Figure 5-35). The accompanying ANOVA test of the factor 'site' was significant ($P < 0.0001$) and Brown and Forsythe's test for homogeneity of variance was non-significant ($P = 0.4794$).

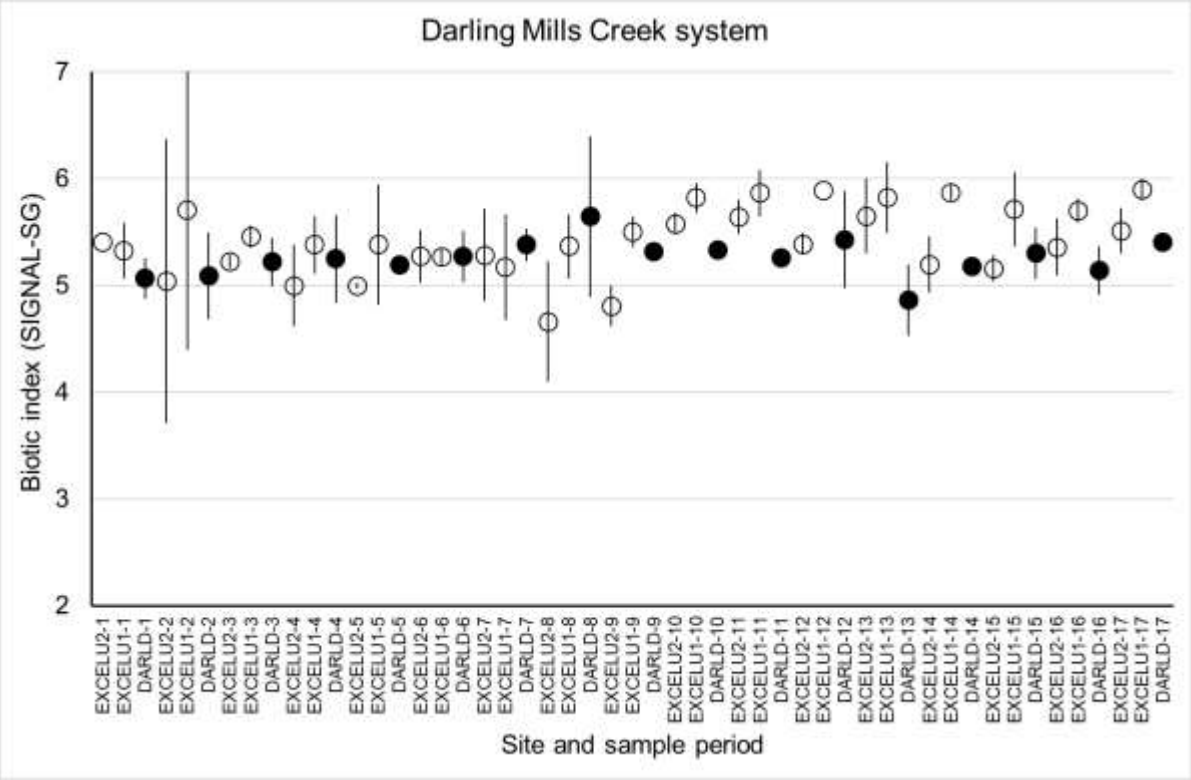


Figure 5-35: Darling Mills Creek system macroinvertebrate pool-edge water samples across 17 collection periods from three study sites

In a two-dimensional nMDS ordination plot of Darling Mills Creek samples from across collection periods 10-17, three groups of adjacent site samples were evident (Figure 5-36). In the tree diagram there was an initial separation of two downstream site samples. The next division had a similarity of 50% that separated the three groups of site samples, although in the group with samples from upstream site EXCELU1 were two downstream site samples and seven samples from the other upstream site EXCELU2 (Figure 5-37).

The PERMDISP analysis indicated a dissimilar pattern of dispersion (spacing between same site samples) for the two sites (F = 4.57, P = 0.03). The presence of multivariate dispersion in the dataset indicated that the ANOSIM test assessed both site differences and collection period differences in assemblage structure.

The ANOSIM test run on the factor 'Site' returned a high range value (R = 0.56; P = 0.001) confirming community structure was distinct at each site. Pairwise tests indicated that the two upstream (EXCELU1, EXCELU2) versus downstream (DARLD) site comparisons had mid-level R-values (0.56, 0.49 respectively) at a level observed for natural differences between sites outlined in Section 5.2.3. These R-values were lower than that of the pairwise comparison of upstream sites (EXCELU1 and EXCEL2) that returned a higher mid-range R-value of 0.65. These pairwise test results reflect the adjacent clusters of site samples observed in the ordination plot.

Evident in the corresponding shade plot (Figure 5-38) were moderately graded SIGNAL-SG taxa, a water mite Limnesiidae *Limnesia* (grade 7) and a beetle larva Elmidae *Kingolus* (grade 7) at the downstream and both upstream sites along with the biting midge larva Ceratopogonidae *Bezzia* that has a SIGNAL-SG grade of 6. Stoneflies were absent in these urban streams, as they are potentially too warm at the low altitudes of our urban study streams which were typically below 100 m. Pollution-sensitive mayflies were also absent, potentially reflecting these metal-sensitive taxa being already absent from urban streams (Iwasaki et al., 2018). Four caddisfly taxa were also collected infrequently at the downstream site including the pollution-sensitive Helicophidae *Alloecella* (grade 8) and semi-pollution sensitive taxa of Ecnomidae *Ecnomus* (grade 6) and Hydroptilidae *Hellyethira* (grade 6) (Figure 5-38). The presence of the frequently occurring higher SIGNAL-SG graded taxa contributed to influencing overall higher SIGNAL-SG scores of the Darling Mills system sites (Figure 5-35) compared to the highly urbanised Buffalo Creek sites that had SIGNAL-SG score a unit lower (Figure 5-14).

The higher SIGNAL-SG scores documented in the ecological control chart for the Darling Mills Creek system is potentially supported by above comments outlined in Section 4.1 of Birch and Taylor (1999) who indicated that one of the least affected areas of the Sydney Harbour catchment from human impact is the mainly forested sub-catchment of Darling Mills Creek. This forested aspect is expressed in the metric of 'road-density' (road length normalised by catchment area) with a value of 8.2 for the catchment including the Darling Mills downstream site DARLD, and for the more urbanised Buffalo catchment (BUFFD) had a road-density metric of 11.8. This metric in sandstone streams ranged from 6.1 to 14.5.

The dominance of these higher-graded SIGNAL-SG taxa that are relatively more pollution sensitive perhaps suggests that ammonia is well-diluted (and is not a contaminant of concern as demonstrated in Section 4.4) upon entering the receiving waters of the Darling Mills Creek system. Hence these results suggested downstream community structure in Darling Mills Creek, particularly after the return of more frequent rainfall after the drought, was influenced by stormwater loading due to the effective distance (10 km stream length) between upstream and downstream sites within this creek system. In retrospect, the downstream site would have been better placed above the junction of flows from the upper Darling Mills Creek and its tributaries of Bellbird Creek and Blue Gum Creek, although a limitation was that a key ERS was located at this stream junction.

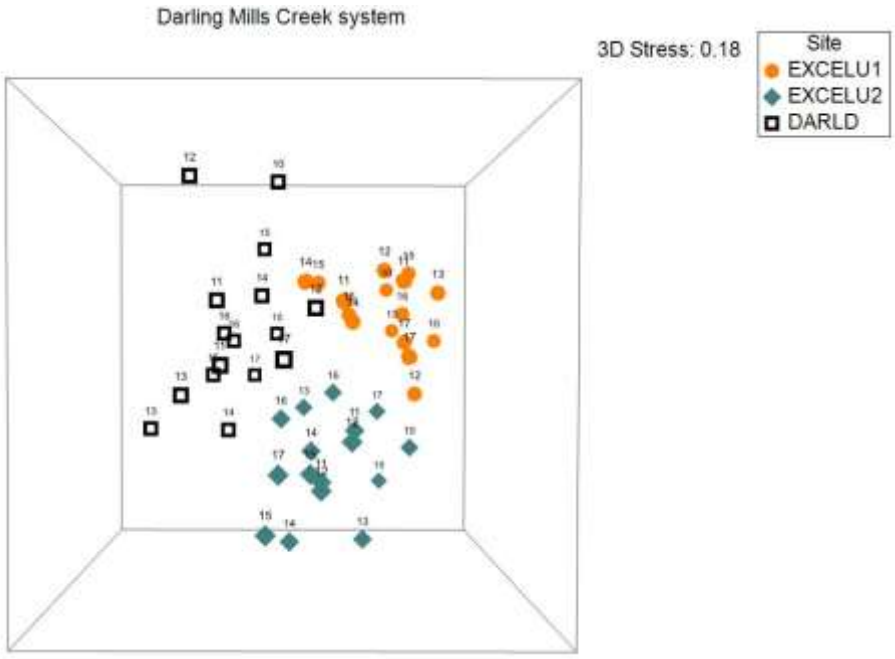


Figure 5-36: nMDS ordination plot of Darling Mills Creek system macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

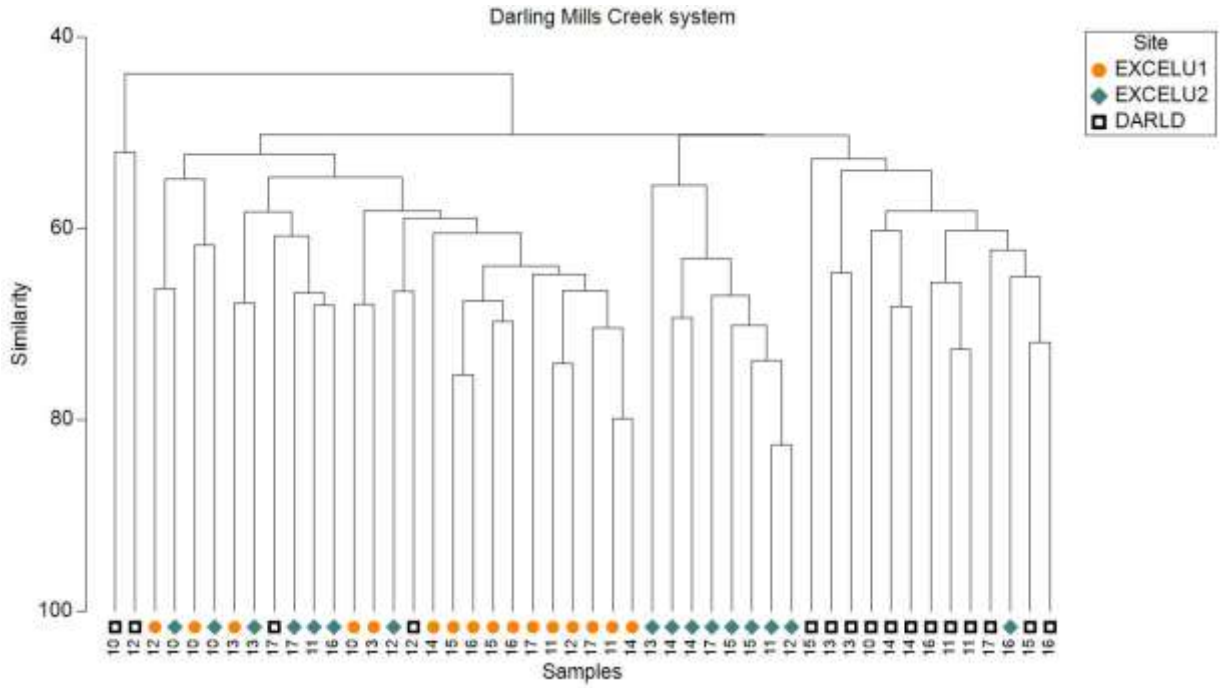


Figure 5-37: Tree diagram from classification analysis of MDS ordination plot of Darling Mills Creek system macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

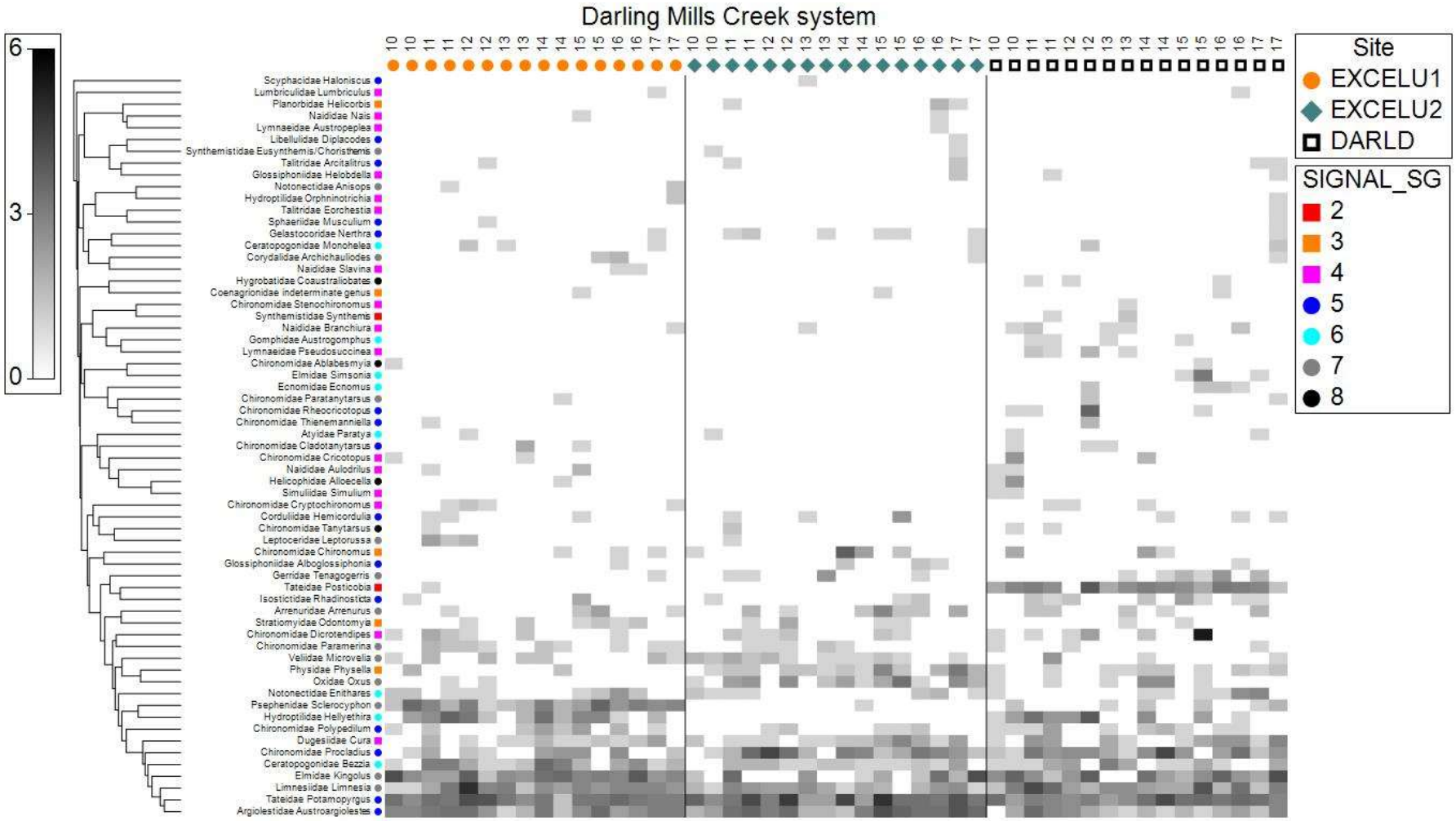


Figure 5-38: Shade plot of Darling Mills Creek system macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

5.3.8 Blacktown Creek

The Blacktown Creek study had two upstream sites situated on different creek arms that joined before the downstream site (Figure 5-39). The upstream sites were about 1.5 km above the downstream site. On the left-hand arm of the Blacktown Creek is upstream site 1 in a small shallow channel with a narrow strip of fringing vegetation (Figure 5-40A). In contrast, the longer upper arm of upper Blacktown Creek that arises near the scale bar in Figure 5-39 is formed as a concrete dish channel (Figure 5-40B) that enters a wetland section (Figure 5-40C) just above the second upstream site BLACU2 which was placed in the second of two wetlands (Figure 5-40E) with an emergent macrophyte reach of the stream (Figure 5-40D) between the two wetlands. The downstream site was sampled in the channel at the end of a lower wetland (Figure 5-40F).

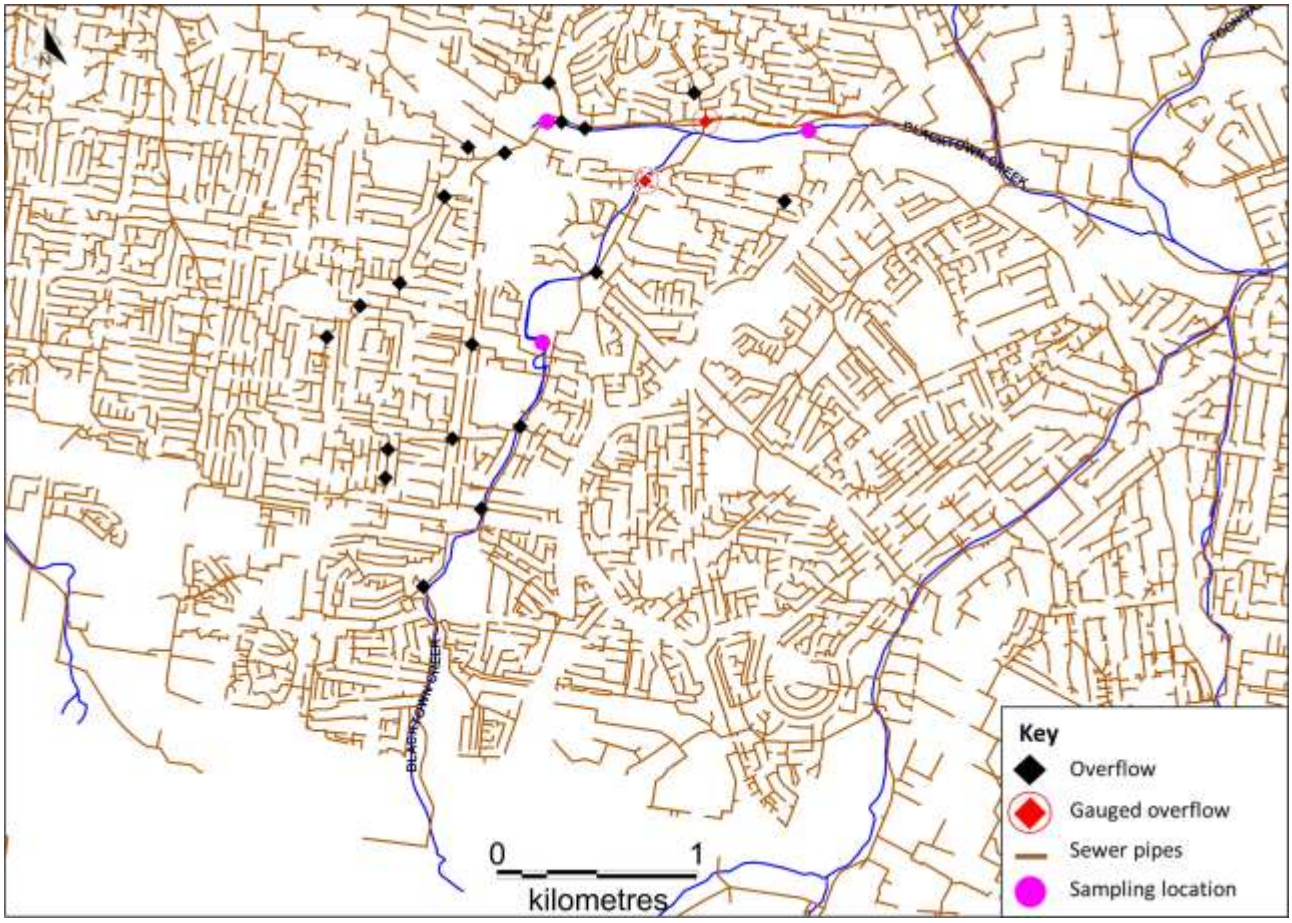


Figure 5-39: Location map of Blacktown Creek study sites

Upstream sampling site 1 is located on the top left stream; upstream site 2 is shown in the middle of the figure. The downstream sampling sites is located after the junction of the two streams on the right-hand side of the figure

Figure 5-40: Blacktown Creek after 7 mm of rain.

(A) upstream site 1 sampling location; (B) canal and start of wetland section upstream of sampling site 2; (C) the larger of the wetlands located upstream of sampling site 2; (D) showing the emergent macrophytes between the wetland and upstream sampling site 2; (E) upstream sampling location 2 in wetland; and (F) sampling location downstream

The ecological control chart plot of SIGNAL-SG scores visualised stream health of Blacktown Creek through the WWOM study across December 2018–April 2021 by each 8-weekly collection (periods 2 to 17). This visual comparison illustrated upstream site 1 stream health was consistently lower than that of upstream site 2 and the downstream site across collection periods 10 to 17 except for period 16 (March 2020 to April 2021) (Figure 5-41).

Statistical testing was conducted upon data from collection periods 10 to 17. A comparison of the upstream-downstream SIGNAL-SG site scores from a SNK multiple mean comparison test indicated the downstream site had a statistically lower mean score (BLACD = 4.99 SD 0.24) than that of both the upstream sites (BLACU1 = 3.71 SD 0.42, BLACU2 = 4.74 SD 0.46), which supported the visual trends in the control chart plot (Figure 5-41). The accompanying ANOVA test of the factor 'site' was significant ($P < 0.0001$) and Brown and Forsythe's test for homogeneity of variance was non-significant ($P = 0.2838$).

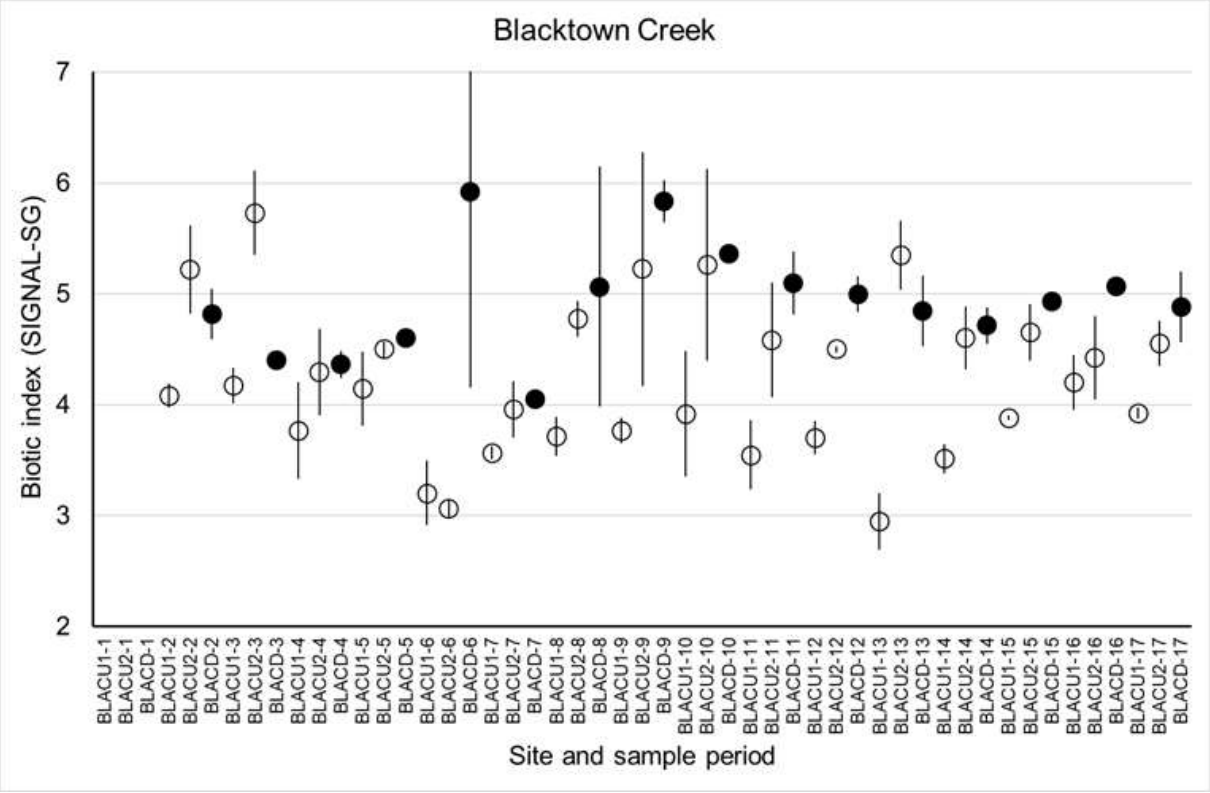


Figure 5-41: Blacktown Creek macroinvertebrate pool-edge water samples across 17 collection periods from three study sites

In a two-dimensional nMDS ordination plot of Blacktown Creek samples from across collection periods 10-17, a distinct group of upstream site 1 (BLACU1) samples was well-separated from a cluster of upstream site 2 (BLACU2) samples and downstream site (BLACD) samples (Figure 5-42). The ordination pattern was confirmed in the corresponding tree diagram (dendrogram) from classification analysis as the first division separated upstream site 1 samples from upstream site 2 and downstream site samples (Figure 5-43). This initial separation also occurred at a quite low similarity of 22% (Figure 5-43).

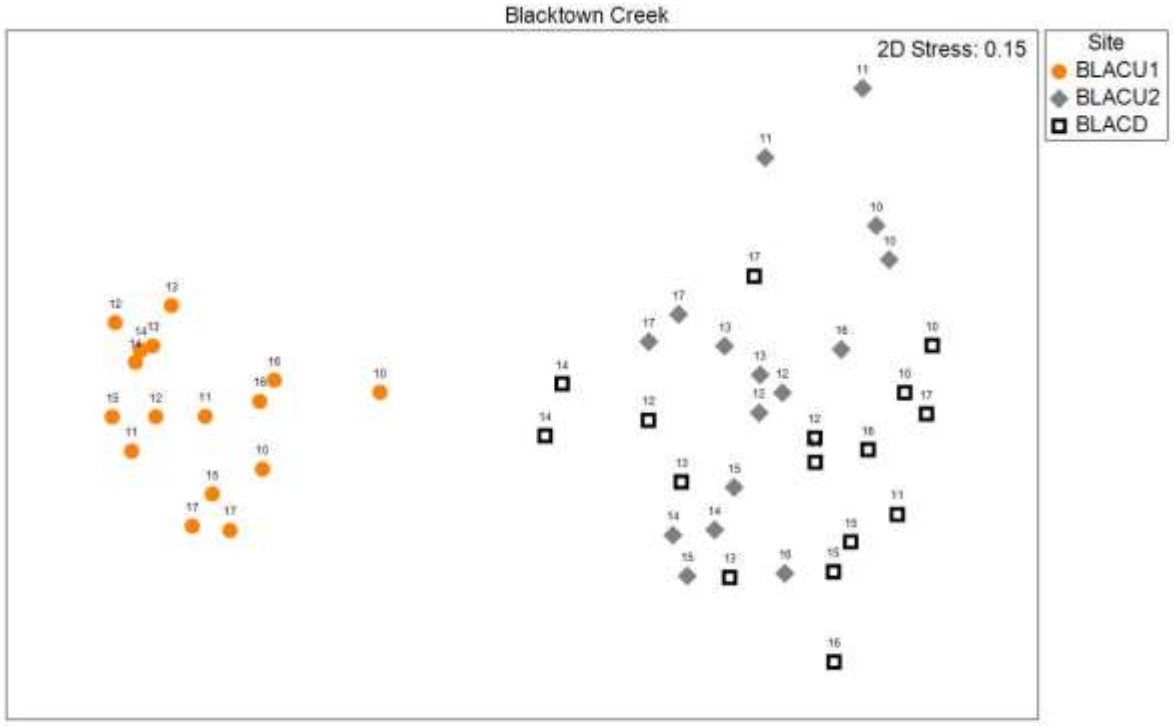


Figure 5-42: nMDS ordination plot of Blacktown Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

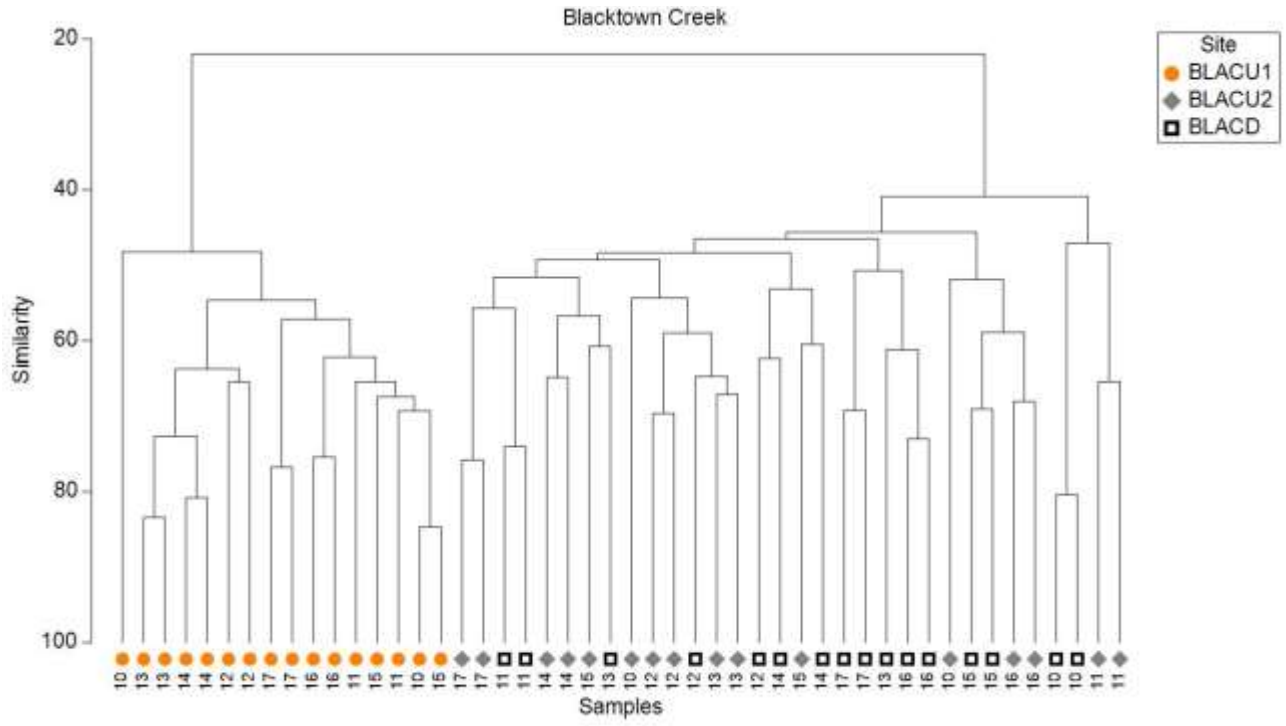


Figure 5-43: Tree diagram from classification analysis of MDS ordination plot of Blacktown Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

The clear separation of Blacktown Creek upstream site 1 (BLACU1) samples was also evident in the corresponding shade plot, which had few consistently present taxa and few in common with those documented in the downstream site (BLACD) samples and the upstream site 2 (BLACU2) samples as (Figure 5-44).

The PERMDISP analysis indicated a dissimilar pattern of dispersion (between same site samples) ($F = 8.36$, $P = 0.002$), although this is most probably predominantly influenced by samples from upstream site 2 (BLACU2) and the downstream site (BLACD) that were more variable in assemblage structure compared to a relatively stable assemblage structure at upstream site 1 (BLACU1) as evident in the nMDS ordination plot (Figure 5-42).

The ANOSIM test run on the factor 'Site' returned a high range value ($R = 0.75$; $P = 0.001$) confirming community structure was distinct at each site. Pairwise tests indicated the two upstream (BLACU2, BLACD) versus downstream (BLACU1) site comparisons had relatively high-level R-values (0.97, 0.97) close to the maximum R-value of 1. In contrast, comparison of upstream sites (BLACU2 and BLACD) returned a low-range R-value (0.15) at a level observed for natural differences between sites outlined in Section 5.2.3. These pairwise test results suggest clear differences in assemblage structure between the upstream site 1 and upstream site 2 and the downstream site.

Six ungauged ERSs were situated above upstream site 1 (BLACU1) and revised hydraulic modelling that became available in 2022 during the closing stages of the WWOM project indicated a modelled 10-year spill volume of 35 ML to this arm of Blacktown Creek. This contrasted with a modelled volume for the arm above upstream site 2 (BLACU2) of 54 ML and 221 ML from ERSs in the stream reach between where these upstream sites were situated and above the downstream site (BLACD). This suggests that WWO spill exposure may have been responsible for this difference in assemblage structure given the very narrow and shallow nature of the channel on which upstream site 2 (BLACU2) was situated (Figure 5-40A) and in which dilution may have been limited. During a site inspection where the photographs in Figure 5-40 were taken, 7 mm of rain had fallen in the early morning and the shallow flow observed in the stream was from stormwater inflows. The site inspection illustrated the downstream site was situated in a larger waterbody (Figure 5-40F) and the other upstream site 2 was below a relatively extensive constructed wetland choked in sections with emergent macrophytes (Figure 5-40B and Figure 5-40D). Perhaps the potentially larger overflow volume of the downstream site was diluted to a sufficient level, while buffering in the wetland may have occurred to ameliorate detrimental effects to upstream site 2. This suggestion may have been possible based on past observations, where a process failure at the North Richmond WWTP diverted untreated influent to Redbank Creek (in a stream reach with extensive macrophyte beds), and no adverse ecological effect was measured from application of the SIGNAL-SG index to collected macroinvertebrate samples.

Outcomes of the above analyses indicated that the assemblage structure in upstream site 1 of Blacktown Creek was consistently different to that of both upstream site 2 and the downstream site. The ecological control chart of stream health illustrated that upstream site 1 supported an impaired assemblage of taxa compared to the other two sites. Field observations suggest that it is plausible for the poor dilution of WWO spills at upstream site 1 of Blacktown Creek, as the likely influence on the documented impaired assemblage documented.

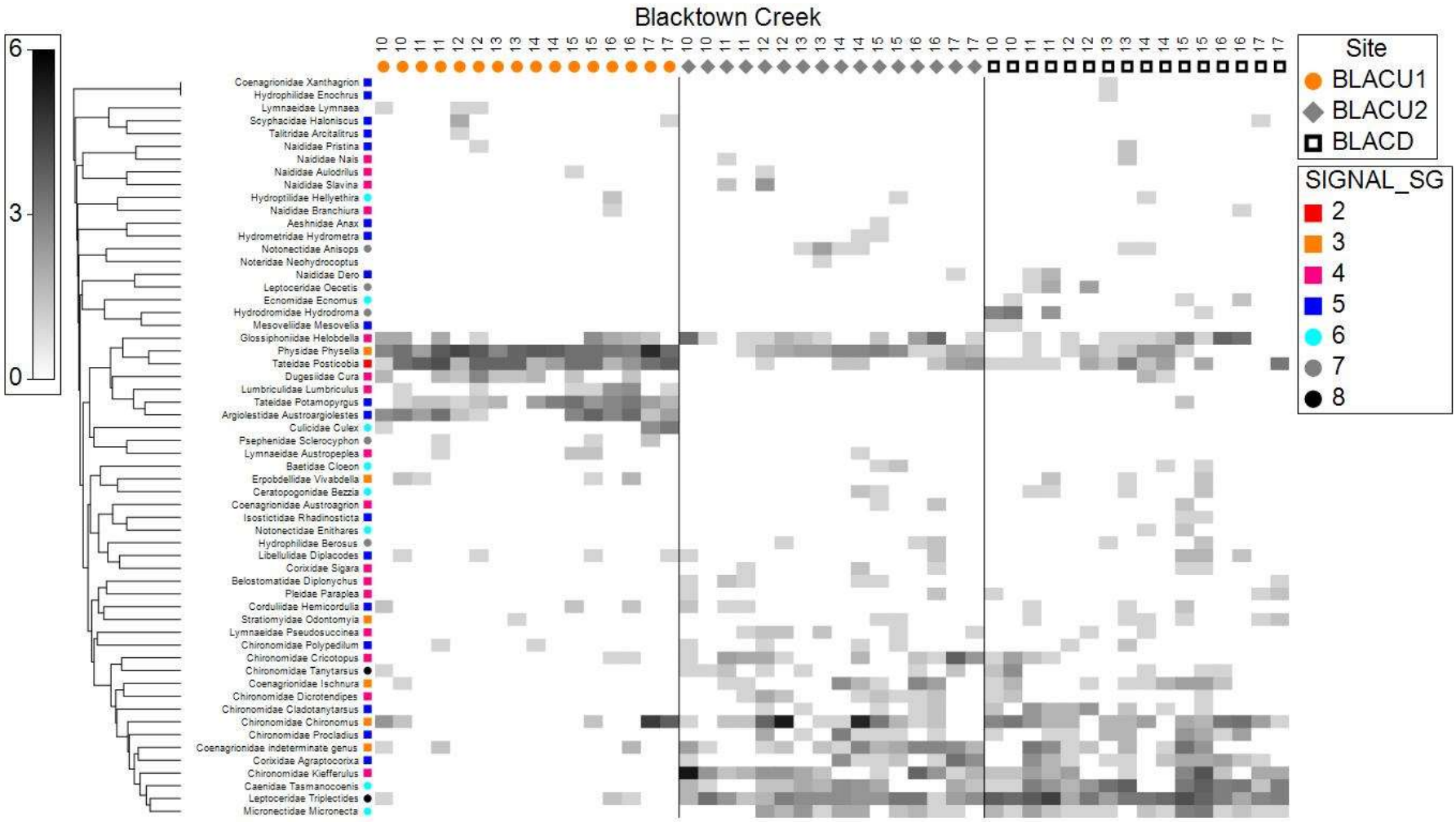


Figure 5-44: Shade plot of Blacktown Creek macroinvertebrate pool-edge water samples across collection periods 10-17 from three study sites

Morphometric macroinvertebrate summary

Results of the dilutions study identified ammonia as a contaminant of potential concern at the downstream Vineyard site (Section 4.4). This was then also supported by outcomes of toxicity testing conducted at this location (Section 4.5.1). Examination of morphometric macroinvertebrate assemblages in Vineyard Creek (Section 5.3.1) illustrated that a different assemblage of macroinvertebrates occurred at the downstream site compared to both upstream sites while results from the ecological control chart indicated an impaired assemblage downstream (Figure 5-10), suggesting that an ongoing disturbance is exerted by WWO spills into Vineyard Creek.

Outcomes of macroinvertebrate assemblage and stream health evaluations also demonstrated another example of ongoing disturbance in this case from inadvertent influent release into Avondale Creek (Section 5.3.3) from an observed network fault detected from site inspections (Section 4.7). It has to be noted that this was the worst sewage smell experienced at a receiving water site in over 30 years of field work (pers. comm. Colin Besley).

The placement of upstream and downstream sites with relatively close spatial proximity around ERSs in the smaller urban streams of Kittys Creek (Section 5.3.5) and Frenchs Creek (Section 5.3.6) led to the detection of assemblage differences that appeared to be attributable to WWO spills. These two outcomes potentially support the theory that when overflow volume spills exceed the capacity of the receiving water to sufficiently dilute, these WWO spills are impairing the stream health, as has been confirmed for Vineyard Creek. This lack of sufficient dilution in the receiving waters is potentially also the case for the impairment observed at upstream site 1 on Blacktown Creek (Section 5.3.8). And potentially also the case for Girraween Creek, where four category 1 ERSs spill into a relatively narrow and shallow urban stream along a 2 km reach between the upstream and downstream sites (Section 5.3.4). This case study from Girraween Creek illustrates an example of too much overflow volume from a number ERS can have the same effect as multiple ERSs discharging to the same stream reach as in the case of Vineyard Creek or a single or two ERSs that are potentially oversized for the stream as was the case for Kittys and Frenchs creeks.

Given that the impairment in Avondale Creek appeared to relate to a network fault, then the apparent ongoing disturbances can be attributed to WWO spills and occurred at 22% (five) of the 23 urban streams assessed (as outlined above). Across the other 17 urban streams assessed under the WWOM project, neither apparent ongoing nor intermittent-sporadic disturbances could be attributed to WWO spills with the traditional morphometric macroinvertebrate indicator. Using a morphometric macroinvertebrate indicator going forward would be cost-prohibitive in assessing the broader 3000 ERSs, even if knowledge gained enabled a subset of locations to be assessed. It would be much more cost-effective to assess ammonia and evaluate those results against the ANZG (2018) default guideline value (DGV) to establish locations where there is potential for adverse ecological effects.

5.4 Modelling morphometric macroinvertebrate assemblages

The multivariate regression techniques of canonical correlation of principal coordinates (CAP) and distance-based linear models (DISTLM) (McArdle and Anderson, 2001) were used to explore companion metadata collected with ecological samples. A focus of this data-exploration was to understand the variation explained by measures of overflow contribution in comparison to stormwater exposure and natural variables. Anderson et al. (2008) states 'The DISTLM routine asks: how much variability in the multivariate data cloud is explained by variable X? In contrast, the CAP routine asks: how well can I predict positions along the axis of X using the multivariate data cloud? So, the essential difference between these two approaches is in the role of these two sets of variables in the analysis. DISTLM treats the multivariate data as a response data cloud, whereas in CAP they are considered rather like predictors instead.'

Metadata variables employed in modelling are detailed below along with transformations (where applied) after inspecting draftsman plots to correct skewness.

- Modelled overflow contribution (natural log +1 transformed)
  - was represented by the modelled overflow volume (ML) from a ten-year time series adjusted to 15 months (to represent the time span of sample collection periods from March 2020 [10] to April 2021 [17]) then normalised by catchment area (km^2) above each sample site
- Gauged overflow contribution (natural log +1 transformed)
  - consisted of the overflow volume (ML) from gauging measurements in the 90 days before sampling normalised by catchment area (km^2) above each downstream sample site.

    The model runs with a reduced subset sites was limited to those downstream sites below ERSs with installed sewer gauges in the outlet structures of the ERSs.
- Stormwater exposure
  - comprised road density (road length [km] divided by catchment area [km^2]) multiplied by site specific total rainfall (mm in the 90 days prior to the sample collection day to each of collection periods 10 to 17).

    Rainfall was highly correlated (r = 0.96) for 60 and 90 days before sampling after viewing a draftsman plot and an associated correlation matrix. The total rainfall for the 60 days prior to sample collection was not included in model runs
- three variables characterising stormwater influence
  I. total impervious area in the catchment above a site
  II. road density (as defined above)
  III. site specific rainfall (as defined above)
- six variables characterising natural variation
  I. altitude (m) at a site
  II. riparian cover (%) from subjective assessment on three occasions at each site
  III. stream width (m), estimate of the widest pool reach of each site
  IV. riparian strip width (m) natural log +1 transformed

V. sediment size class mud (< 0.063 mm) fourth root transformed

VI. sediment size class gravel (> 2.0 mm) fourth root transformed

CAP modelling

The model runs described below were conducted on each of the four site groupings (Table 5-2). One group of sites was based on subsurface geology, with sites in sandstone catchments analysed separately to those sites in shale catchments. Another model run of downstream sites was based on the availability of gauged (measured) overflow data from study ERSs. The final model run was enabled by hydraulic modelling (2022) to allow both the upstream and downstream sites with modelled overflow data that included ERS in the catchments above the upstream sites.

CAP was initially used to specifically model for changes in the freshwater macroinvertebrate taxonomic assemblage structure (multivariate response data cloud) correlated with change along a gradient (Anderson and Willis, 2003). Two gradients were assessed, an overflow contribution gradient, followed by a stormwater exposure gradient. These gradients were based upon:

- overflow contribution (natural log +1 transformed) modelled for upstream and downstream site model runs while gauged information was employed in the reduced subset of downstream site model runs
- stormwater exposure

CAP analysis was also undertaken to assess the relationship of continuous variables with the freshwater macroinvertebrate taxonomic assemblage structure. Three sets of continuous variables were assessed:

- three variables characterising stormwater influences
    I. total impervious area in the catchment above a site
    II. road density (as defined above)
    III. site specific rainfall (as defined above)
- six variables characterising natural variation
    I. altitude (m)
    II. riparian cover (%)
    III. stream width (m)
    IV. riparian strip width (m) natural log +1 transformed
    V. sediment size class mud (< 0.063 mm) fourth root transformed
    VI. sediment size class gravel (> 2.0 mm) fourth root transformed
- ten variables comprised of the above nine variables along with overflow contribution

The CAP model runs conducted on each of the four site groupings explored the two gradient variables and three sets of continuous variables (Table 5-2). CAP uses principal coordinates (PCO, Gower, 1966) from a dissimilarity matrix and a check on over parameterisation is needed (that is, to avoid including too many axes and finding spurious relationships). This was achieved by choosing the number of PCO axes (m) that minimised a leave-one-out residual sum of squares and ideally total variation is above 60%. To base the CAP analysis, the companion freshwater macroinvertebrate taxonomic assemblage dataset was square root transformed before a

dissimilarity matrix based on the Bray-Curtis association measure was raised prior to conducting the above model runs.

The Canonical Analysis of Principal coordinates (CAP) of freshwater macroinvertebrate taxonomic assemblage structure to gradients of overflow contribution and stormwater exposure yielded negligible to weak correlation (δ = 0.07 to δ = 0.27) under all eight model runs (Table 5-2). CAP analysis undertaken to assess the relationship of three continuous variables characterising the influence of stormwater with the freshwater macroinvertebrate taxonomic assemblage structure returned moderate correlation (δ = 0.52 to δ = 70, Table 5-2). While moderate to strong correlation was returned from the four model runs of the six continuous variables characterising natural variation (δ = 0.67 to δ = 80, Table 5-2). Marginally higher correlations were returned across the four model runs of the ten continuous variables to freshwater macroinvertebrate taxonomic assemblage structure (Table 5-2).

As described above, the CAP routine asks how well can positions be predicted along the axis of X using the multivariate data cloud? The negligible to weak correlation of both gradient model runs (overflow contribution and stormwater exposure) inhibits these CAP models from being used in a predictive capacity with inclusion of new site data. While improved moderate correlation results were returned for model runs of the three stormwater variables, this level of correlation is still below a level where reliable predictions could be made of new site data. In contrast, the model runs of the variables of natural variation returned strong correlation which could be used for prediction, but this would not be useful in ranking ERS sites as an input into the risk prioritisation methodology.

Table 5-2: Canonical correlation of principal coordinates (CAP) model runs exploring the relationship between assembled metadata (environmental variables) and freshwater macroinvertebrate taxonomic assemblage structure for four site groupings

Subset of sites	Overflow contribution^	Stormwater exposure	Three variables characterising stormwater	Six variables characterising natural variation	Combined run of 10 variables^^
Sandstone upstream and downstream (32) sites	δ = 0.22, m = 6 % var* = 55%	δ = 0.13, m = 9 % var* = 67%	δ = 0.54, m= 8 % var* = 63%	δ = 0.68, m = 10 % var* = 70%	δ = 0.71, m = 11 % var* = 74%
Shale upstream and downstream (18) sites	δ = 0.27, m = 6 % var* = 59%	δ = 0.19, m = 7 % var* = 64%	δ = 0.52, m = 7 % var* = 64%	δ = 0.67, m = 9 % var* = 73%	δ = 0.69, m = 9 % var* = 73%
Sandstone downstream (13) sites	δ = 0.09, m = 9 % var* = 70%	δ = 0.14, m = 8 % var* = 66%	δ = 0.55, m = 8 % var* = 66%	δ = 0.80, m = 9 % var* = 70%	δ = 0.80, m = 8 % var* = 66%
Shale downstream (8) sites	δ = 0.07, m = 6 % var* = 0.65%	δ = 0.16, m = 6 % var* = 65%	δ = 0.70, m = 6 % var* = 0.65%	δ = 0.80, m = 5 % var* = 59%	δ = 0.87, m = 5 % var* = 59%

^ overflow contribution based on modelled information for upstream and downstream model runs while gauged information was employed in downstream model runs.

^^ 10 variables of combined model run were: altitude, riparian cover, riparian strip width, stream width, mud, gravel, total impervious area, road density, total rainfall in 90 days prior to sampling, overflow exposure.

* % var = the percentage of the total variation explained by the 'm' principal coordinate axis. All model runs were significant with p = 0.001 except for overflow exposure in the sandstone downstream model run where p = 0.016.

DISTLM modelling

Anderson et al. (2008) outlined an approach to analyse the data in sets under the DISTLM routine of PERMANOVA+. Under this approach one can explicitly examine the proportion of variation in the taxonomic data that is explained by either the overflow exposure, or stormwater variables over and above the amount explained in this case by the natural variables alone. This allows the test of the hypothesis that there is no relationship between the taxonomic assemblage and either the overflow contribution or set of three stormwater variables given the set of six natural variables.

DISTLM model runs assessed three sets of response variables (Table 5-3):

- natural variables (the six variables characterising natural variation)
- stormwater variables (the three variables characterising stormwater influences)
- overflow exposure (gauged or modelled as described in above for respective site groupings)

Two model runs were conducted on each of:

- Sandstone upstream and downstream sites
- Shale upstream and downstream sites
- Sandstone downstream sites
- Shale downstream site

Before running the DISTLM routine, each of the environmental variables (metadata) were normalised (subtracting the mean and dividing by the standard deviation for each variable). A dissimilarity matrix based on Euclidean Distance was raised for the environmental variables. Transformations were applied as outlined above and one representation of rainfall (90 days) was included to account for multi-collinearity of the rainfall mentioned above.

The four dissimilarity matrices based on the Bray-Curtis association measure raised from the freshwater macroinvertebrate taxonomic assemblage datasets for the CAP model runs were used in DISTLM analysis. DISTLM model runs were based on the all-specified selection procedure using the adjusted R^2 selection criteria.

The adjusted R^2 of the eight DISTLM model runs was relatively low (between 0.17 and 0.38) (Table 5-3). Inspection of sequential tests in each model run provided an understanding of the variation explained by each set of variables tested. The highest proportion of variation was explained in each of the eight model runs by the set of six natural variables (Table 5-3). In contrast, the stormwater set of variables explained three to five times less variation than the set of natural variables, and the overflow contribution explained only a paltry 1 to 3% of the variation (Table 5-3).

The DISTLM model run outcomes were supportive of CAP outcomes. The natural variables best explained the data patterns within the multivariate response data cloud that represented freshwater macroinvertebrate taxonomic assemblage structure of samples across collection periods 10 to 17.

Table 5-3: DISTLM model runs exploring sets of response variables and freshwater macroinvertebrate taxonomic assemblage structure for four site groupings

Site grouping tested	Adjusted R^2	Sequential tests of model run for sets of variables	Proportion of variation %
Sandstone upstream and downstream (32) sites	0.17	Natural	17
		Overflow contribution	1
Shale upstream and downstream (18) sites	0.20	Natural	21
		Overflow contribution	3
Sandstone downstream (13) sites	0.21	Natural	22
		Overflow contribution	1
Shale downstream (8) sites	0.32	Natural	35
		Overflow contribution	1
Sandstone upstream and downstream (32) sites	0.22	Natural	17
		Stormwater	6
Shale upstream and downstream (18) sites	0.22	Natural	20
		Stormwater	4
Sandstone downstream (13) sites	0.27	Natural	22
		Stormwater	8
Shale downstream (8) sites	0.38	Natural	35
		Stormwater	7

Summary of morphometric macroinvertebrate modelling

CAP outcomes indicated that natural variables best explained data patterns within the freshwater macroinvertebrate samples from collection periods of March 2020 [10] to April 2021 [17]. However, the negligible to weak correlation of both gradient CAP model runs (examining overflow contribution and stormwater exposure) limits meaningful predictive capacity from these gradient models.

The DISTLM model run outcomes were supportive of CAP outcomes that indicated natural variables best explained data patterns and that overflow contribution explained only a paltry (1 to 3%) of the variation. This was also found from model runs of the stormwater variables.

The assembled metadata variables did not enable a predictive capacity to allow a viable modelling approach for WWOs. Hence, predictive modelling is not proposed as a future option for the risk prioritisation methodology based upon morphometric macroinvertebrate taxonomic assemblage data.

References

Anderson, M.J., Gorley, R.N., Clarke, K.R., 2008. PERMANOVA+ for PRIMER: Guide to Software and Statistical Methods. PRIMER-E, Plymouth, U.K.

Anderson, M.J., Willis, T.J., 2003. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. Ecology 84, 511–525.

Gower JC. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325-338.

McArdle, B.H., Anderson, M.J., 2001. Fitting multivariate models to community data: a comment on distance based redundancy analysis. Ecology 82, 290–297.

5.5 Comparison of SIGNAL scores from the traditional and DNA-derived taxonomy

Macroinvertebrate surveys based upon morphological identification and companion counting of specimens present in samples are commonly used for assessing the health of freshwater systems around the world (examples are listed in Section 5.1). Biological indices, derived from taxonomic lists, provide convenient ways to summarise assemblage data as illustrated by application of the biotic index SIGNAL-SG in Section 5.3.

Raising sample biotic index (SIGNAL-SG) scores based on taxonomic data can also be performed on outputted molecular datasets from either:

- Community-DNA (the identification of species from mixtures of whole animals in bulk-samples) or
- environmental DNA (eDNA) from environmental samples (such as sediment)

Pawlowski et al. (2021, p. 2931) summarised three ways forward in the implementation of a molecular approach. They noted *'Astonishingly, these rapid advances in eDNA-based technologies are rather timidly implemented in routine biomonitoring (Hering et al., 2018; Shackleton et al., 2021). Although the concept of Biomonitoring 2.0 is widely endorsed, its acceptance in practice is hampered for various reasons. There is no consensus whether eDNA-based biomonitoring should only apply to:*

- *conventional bioindicators (1. **Renovate**), or*
- *new bioindicators (2. **Rebuild**), or*
- *new taxonomy-free approaches (3. **Revolutionise**).'*

In this section of the report, the conservative 'renovate' approach is evaluated with community DNA and eDNA datasets that were gathered under the WWOM project from freshwater urban streams.

Sections 5.6 and 5.7 below contain summaries of eDNA evaluations of the taxonomy-free 'revolutionise' approach that was explored by Macquarie University.

Reference

Pawlowski, J., Bonin, A., Boyer, F., Cordier, T., Taberlet, P. 2021. Environmental DNA for biomonitoring. Mol. Ecol. 30, 2931 - 2936. https://doi.org/10.1111/mec.16023

5.5.1 Community-DNA pilot study

As mentioned in the Pawlowski et al. (2021) comment above, this pilot study is an example of a conservative renovate approach of Biomonitoring 2.0. Shackleton et al. (2021) noted in recent years, molecular techniques for identifying taxa have become increasingly popular and metabarcoding approaches that offer the ability to identify species from' community-DNA or eDNA have gained much attention. However, generating accurate species lists from metabarcoded data is challenging and can be impacted by sample type, choice of primers, community composition within samples, and the availability of reference sequences in DNA libraries (Shackleton et al., 2021).

Objectives of pilot study

Under this early pilot study of the WWOM project, the performance of community-DNA extracted from bulk-samples of macroinvertebrates collected with the Chessman et al. (2007) methodology (as employed in Section 5.3 above) was compared against morphometric data obtained from these same samples before molecular processing. The taxonomy obtained from the morphometric dataset was compared with taxonomy from the molecular dataset by calculating biotic index scores from two versions of the SIGNAL (the Stream Invertebrate Grade Number Average Level) biotic index. A short, 313 base-pair, fragment of the mitochondrial cytochrome c oxidase (COI) barcoding region was used to determine whether SIGNAL2 and SIGNAL-SG biotic indices derived from DNA data were comparable to those derived from traditional, morphometric data. However, there were key differences from other studies on this topic, some of which aimed to reduce the cost and time involved with sample preparation.

- Firstly, DNA was extracted from whole bulk-samples rather than dissecting tissue from individual animals as was done by Carew et al. (2018).
- Secondly, only a single set of primers was used in this study, compared to three sets in Carew et al. (2018) and Marshall and Stepien (2020). While this may reduce taxonomic coverage it also reduces the sample preparation time and increases the sequencing read depth available per sample.
- Thirdly, past studies have investigated family-level metrics, whereas the present study includes a genus-level metric (SIGNAL-SG). Because taxonomic assignment of Operational Taxonomic Units (OTUs) can be affected by the taxonomic composition of reference DNA databases, metrics were derived using three molecular datasets containing the same OTUs but with differing taxonomic identifications applied from different reference databases in order to investigate how incomplete barcode libraries effect metric outcomes. The effect of filtering OTUs, based on their percent contribution to samples, had on index and metric outcomes was investigated at varying thresholds (Shackleton et al., 2021).

Taxa were assigned to OTUs using three reference libraries of COI barcodes, resulting in three sequence datasets:

- The first was a library of curated barcodes obtained from GenBank (Benson et al., 2012) (https://www.ncbi.nlm.nih.gov/genbank/ accessed May 2018), which contains data from species across the world
- The second was the Aquatic Invertebrates of Australia reference library (AIA), housed on the Barcode of Life Database (BOLD) (Ratnasingham and Hebert, 2007)

(http://www.boldsystems.org/ accessed July 2018), which contains only data from Australian macroinvertebrate species and many sequences from species that were not within the GenBank library

- The third was a combination of the two libraries, herein referred to as the Best of Both Worlds (BoBW) library. Taxonomic assignment for BoBW was achieved by taking the highest percent identity match from either the GenBank or AIA datasets

Text and citations in this section were drawn from the peer-reviewed journal publication of: Shackleton, M., Dafforn, K.A., Murphy, N.P., Greenfield, P., Cassidy, C., Besley, C.H. 2021. How does molecular taxonomy for deriving river health indices correlate with traditional morphological taxonomy? Ecol Indic. 125, 107537 https://doi.org/10.1016/j.ecolind.2021.107537

Main findings

This study demonstrated that the river health SIGNAL biotic indices can be derived from bulk sample community DNA data with results that are comparable to those derived through traditional morphometric analyses. Strong and significant correlations between morphometrically and molecularly derived SIGNAL scores were observed for both family (SIGNAL2) and genus-level (SIGNAL-SG) analyses. However, the choice of DNA reference library and data pre-processing influenced the significance and strength of correlations. In both generic and family level analyses, correlations greatly improved when at least some filtering of low contribution OTUs was performed (for example, filtering out of OTUs that contributed < 0.01%). At the genus- level, datasets with taxonomic identifications made using a reference database of local taxa (for example, AIA) performed better than using the GenBank reference library, with the AIA dataset returning the highest Pearson Coefficient Correlation (PCC) of 0.78 (p-value < 0.001) when using a 0.05% read number threshold. At the family level, the correlation of the GenBank data with the morphometric SIGNAL2 scores was greatly improved and when applying a read number filter at the 0.1% threshold, the GenBank dataset had the highest correlation of the molecular datasets (PCC = 0.79, p-value < 0.001). However, at lower thresholds the BoBW dataset performed marginally better, while the performance of the AIA dataset decreased.

Overall, molecular classification was relatively accurate with accuracies over 80% at the generic level and over 95% at the family level obtainable. At these levels, managers of river health could apply the molecular techniques described here with some confidence that their results will be relatively consistent with those of traditional methods.

An evaluation of what taxa did and did not amplify well, is included in Shackleton et al. (2021). Notably the flatworms Dugesiidae do not amplify, which has been encountered before. A discussion of taxonomic identification errors as false positive or negatives was also presented. An example from this discussion is the two mite genera *Oxus* and *Frontipoda* that were morphometrically recorded among the samples but were found in separate samples. This morphometric data pattern raised the following investigation of *Frontipoda* that did not occur in the molecular data; however, suspiciously, *Oxus* was recorded in all of the samples that contained *Frontipoda*. A review of the *Oxus* voucher specimens that contributed to the AIA barcode library, revealed, incorrect identifications placed onto some of the *Oxus* specimens, specifically those that genetically matched the misidentified *Frontipoda*, and that the vouchers were in fact *Frontipoda*. These voucher specimens were the same that genetically matched to the misidentified *Frontipoda* in the molecular analyses. This highlights the issue of incorrectly assigned taxonomy in DNA

databases and the need for adequately curated reference libraries, as has been emphasised by other authors (Nilsson et al., 2006; Tixier et al., 2012; Shen et al., 2013; Shackleton and Rees, 2016; Carew et al., 2017).

In general, this study attributed discrepancies between morphometric and molecular taxonomic assignments as errors in molecular assignment. Because the samples used were destroyed during the genetic extraction process, morphometric identifications could not be double-checked. However, errors in morphometric identification are probable and not unexpected. An error rate of about 4% was documented for the laboratory that performed morphometric identification and this corresponds with the rate found by Chessman et al. (2007) for genus level analyses on taxa in the Sydney region.

One possibility for unexpected positive results is that trace or environmental DNA (eDNA) may have been present in the samples (Beermann et al., 2020), including for species that may have been present as dietary components of the collected specimens (Zaidi et al., 1999; Sheppard et al., 2005; Hosseini et al., 2008). A further possibility, as recorded for *Oxus* above, is that specimens have been misidentified in the reference databases used and thus the best matches for sequence data are to incorrectly-assigned species. Pawlowski et al. (2018) suggested the most cited explanation for discrepancies between molecular and morphometric datasets is that the incompleteness and lack of accuracy of the molecular reference databases impedes the correct taxonomic assignment of DNA sequences. Elbrecht et al. (2017b) described DNA databases like BOLD (where AIA is hosted) as containing misidentified taxa or conflicting taxonomic assignments for the same Barcode Index Number. Some of the cases of false negatives were due to taxa missing from the reference databases, but these were limited to *Physolimnesia* (mite), *Illebdella*, *Vivabdella* (leeches), *Synthemis* (dragonfly) and *Pygmanisus* (snail).

The findings of this pilot study support those of Carew et al. (2018) who found little difference between DNA-derived and morphometrically-derived family level indices: SIGNAL2, AusRivAS (Reynoldson et al., 1997) and a Chironomidae-based pollution index developed as part of their study. However, unlike Carew et al. (2018), the present study used a single primer pair, reducing the time and costs involved in processing samples and increasing the available sequencing depth per sample. This does, however, come at the cost of increasing the number of undetected taxa due to primer bias and primer-template mismatches. While Carew et al. (2018) were able to recover 85% of families known to be in their samples, in the present study, the average number of families recovered ranged from 70.3% to 84.3% in the BoBW dataset with a range from 38.1% to 100% of families known to occur in the samples were recovered. Shackleton et al. (2021) noted that many of the species used in the present study have since been added to GenBank, and it is thus possible that these percentages will increase if analyses are performed on updated libraries. However, our work has also shown that databases with local taxa may be more useful, and thus GenBank identifications are unlikely to improve in regions where local data are depauperate. A further discussion is provided in Section 5.5.2 of work commissioned under the WWOM that added many of the species used in the present study to GenBank.

Application in biomonitoring

As discussed in Shackleton et al. (2021), in practice SIGNAL scores are interpreted as bands (water quality status classes) indicating gradients of pollution. Chessman (1995), who introduced the first SIGNAL score classified bands as greater than 6 = clean water, 5–6 = possible mild

pollution, 4–5 = moderate pollution and < 4 = severe pollution. In the present study, when applying at least some degree of read-number filtering, most molecularly derived scores were in agreement with the morphometrically derived scores in terms of water quality classification. When molecular classifications deviated from morphometric classifications they predominately classified to the next lower or higher classification; the notable exception being in the genus level analyses using the GenBank dataset which classified a few mildly polluted samples as severely polluted.

While application of the SIGNAL biotic indices can be used to assign water quality status classes, Besley and Chessman (2008) demonstrated the application of graphical and statistical assessment of SIGNAL-SG scores based on morphometric data collected from paired sites situated upstream and downstream of point source discharge of treated sewage wastewater. That graphical assessment illustrated that SIGNAL-SG scores do not neatly fall into a band, and often occur across two water quality status classes (bands). Addition of an overall upstream mean of SIGNAL-SG scores with error bars of ± 1SD for a temporal period allows presentation in a process control chart for ecological monitoring as advocated by Burgman et al. (2012). An example of this control chart approach is provided by the 25-year long-term study (1995 to 2020, Sydney Water STSIMP vol. 2 Appendices Data Report 2019 – 2020) of the Nepean River near the West Camden sewage treatment plant in the Sydney region, Australia, which illustrated the SIGNAL-SG range of morphometric derived scores of about a unit fluctuation as typical variation (see Supplementary Figure S1 of Shackleton et al., 2021). In adopting metabarcoding data as the basis for assessment with biotic indices such as SIGNAL, our study suggests that the quality of the underlying barcode library will cause slight differences in SIGNAL scores; and a period where both morphometric and metabarcoding data are obtained would provide an understanding of the potential site-specific ranges. This conservative approach would consider Buchner et al. (2019) advocation of the importance of properly evaluating the potential to link metabarcoding data to established indices and relating them to existing data. This approach seems prudent as management decisions can be expensive, for example the Blue Mountains Sewage Transfer Scheme in the Sydney region, Australia, was established to upgrade the sewerage system at a cost of AUD $360 million, by progressively closing small, local plants and diverting sewage to a larger, more efficient plant (Besley and Chessman, 2008). These considerations outlined above from Shackleton et al. (2021) need to be kept in mind when implementing Biomonitoring 2.0.

Summary

Under this pilot study, community-DNA derived SIGNAL2 and SIGNAL-SG scores correlated strongly with morphometric derived scores and both were strongest when using a reference library containing a mix of local and global data. This study demonstrated that bulk-sample community-DNA derived data can be used as an alternative for calculating family- and genus-level river health metrics with similar results to current practices with traditional freshwater biological assessment using morphometric macroinvertebrates. Although the DNA reference database gaps identified in this study highlighted the need to expand molecular information with taxa of the Sydney region in DNA databases to better facilitate implementation of the renovate approach of Biomonitoring 2.0.

A summary of work commissioned under the WWOM to infill taxonomy gaps in reference DNA databases is provided in Section 5.5.2.

The successful outcomes of this WWOM pilot study, based upon community-DNA from a single primer-pair (amplicon), provided confidence to explore eDNA (obtained from sediment samples) using an expanded suite of six primer-pairs (amplicons). This suite of primer pairs was applied across a broader range of sites sampled for the main component of the WWOM project between June 2019 to April 2021.

Section 5.5.3 presents an initial comparison of SIGNAL-SG sample scores raised from morphometric applied taxonomy to scores raised from eDNA sediment samples processed with COI and 16S primer-pairs and subsequent post-sequencing bioinformatic processing against DNA databases to provide two datasets with taxonomic assignments.

5.5.2 Gaps in taxonomy within DNA databases – a known limitation

Under the WWOM pilot study outlined in Section 5.5.1 (Shackleton et al., 2021), gaps in the DNA barcode reference library for the Sydney region were highlighted. This was unsurprising as at the commencement of the WWOM project, it was a well-documented issue in the scientific literature that many taxa are missing from molecular databases, which is a limitation to implementing Biomonitoring 2.0. For example, Erdozain et al. (2019) concluded that DNA metabarcoding of storage ethanol (in which macroinvertebrates had been kept) provided a promising approach for characterising stream macroinvertebrate assemblages, but that its full development in biomonitoring projects require developing more complete DNA reference libraries and enhancing the sensitivity for detecting taxa with low sample biomass.

In 2014, the Aquatic Invertebrates of Australia (AIA) DNA database hosted within the Barcode Of Life (BOLD) database was commissioned. This work has been coordinated by Dr Michael Shackleton of La Trobe University. It was understood at that time, in order to use DNA barcoding as a routine species identification tool, it was important to develop a database of morphometrically identified specimens and their barcodes. The main purpose for developing a comprehensive DNA barcode library was to be able to identify DNA metabarcoding sequences to a high taxonomic resolution. Applying DNA metabarcoding approaches to biomonitoring has the potential to greatly reduce the costs and increase detection of taxa. By 2018, the AIA hosted over 2000 specimens and in excess of 300 species. The AIA DNA database was the logical place to have sequences of the taxa from the Sydney region hosted. Dr Shackleton was engaged in a sub-project of the WWOM with the aim to fill in some of these gaps within the AIA by further sequencing targeted genera.

Text and citations in this section were drawn from: Shackleton, M. E., McPhan, L., Murphy, N. P. 2022. DNA barcoding and investigation into PCR bias among Sydney Water macroinvertebrates. La Trobe University, Centre for Freshwater Ecosystems, Wodonga, Victoria. Final report 275 for Sydney Water.

Findings of the pilot study

Sydney Water supplied 411 individual macroinvertebrate specimens for this sub-project. A total of 193 specimens and 47 genera were added to the AIA database; having either their mitochondrial cytochrome c oxidase (COI) or mitochondrial 16S, or both gene fragments barcoded. This increased the number of specimens from the Sydney region represented in the AIA from 253 to 446. Notable additions to the database are the first inclusion of:

- decapods (9 specimens from 3 genera),
- Ephemeroptera (47 specimens from 10 genera)
- Plecoptera (11 specimens from 4 genera)

An additional 10 Odonata, 17 Hemiptera, 27 Coleoptera, and 32 Trichoptera were also added.

The barcoding conducted under this sub-project filled in some important gaps for the Sydney region, however, there are still missing elements. A gap analysis of genera collected in Sydney Water projects and the DNA barcode data available showed that those taxa that are most regularly collected have a greater representation in the barcode datasets than those that are less frequently collected (Figure 5-45). Appendix III (Table 5 of Shackleton et al., 2022) provides a list of genera

collected by Sydney Water alongside and arranged by their abundances and the number of specimens barcoded from the Sydney region as well as represented in the AIA for both COI and 16S fragments. Figure 5-45 graphically summarises these data by providing the proportion of genera that have been barcoded for each family.

Chironomidae are the most abundantly collected family and include some 58 genera. Of these, only 13.8 and 15.5% have been barcoded for COI or 16S, respectively, in either the Sydney Water collection or the AIA as a whole. Chironomidae genera with relatively high abundances (> 500) that are missing from the DNA libraries included *Rheotanytarsus*, *Cladotanytarsus*, *Rheocricotopus*, *Ablabesmyia*, and *Paramerina*.

Not represented in either library were the highly abundant families of Dugesiidae and Glossiphonidae. For these taxa, primer (amplicon) choice needs to be adapted as the universal primers used in this and previous barcoding efforts generally do not work on these taxa.

The genus *Potomopyrgus* in the family Tateidae was regularly collected (1,443 records) but not represented in the barcode libraries whereas its sister taxon, *Posticobia*, was represented but only by a single specimen. Again, this may be the result of universal primers being not suited to all gastropod taxa.

Future updates to the Sydney Water DNA barcode library could aim to focus on those genera that would provide the greatest benefit (that is, genera that are regularly collected). Alternatively, an approach might be to ensure at least one genus from every family is represented in the DNA libraries.

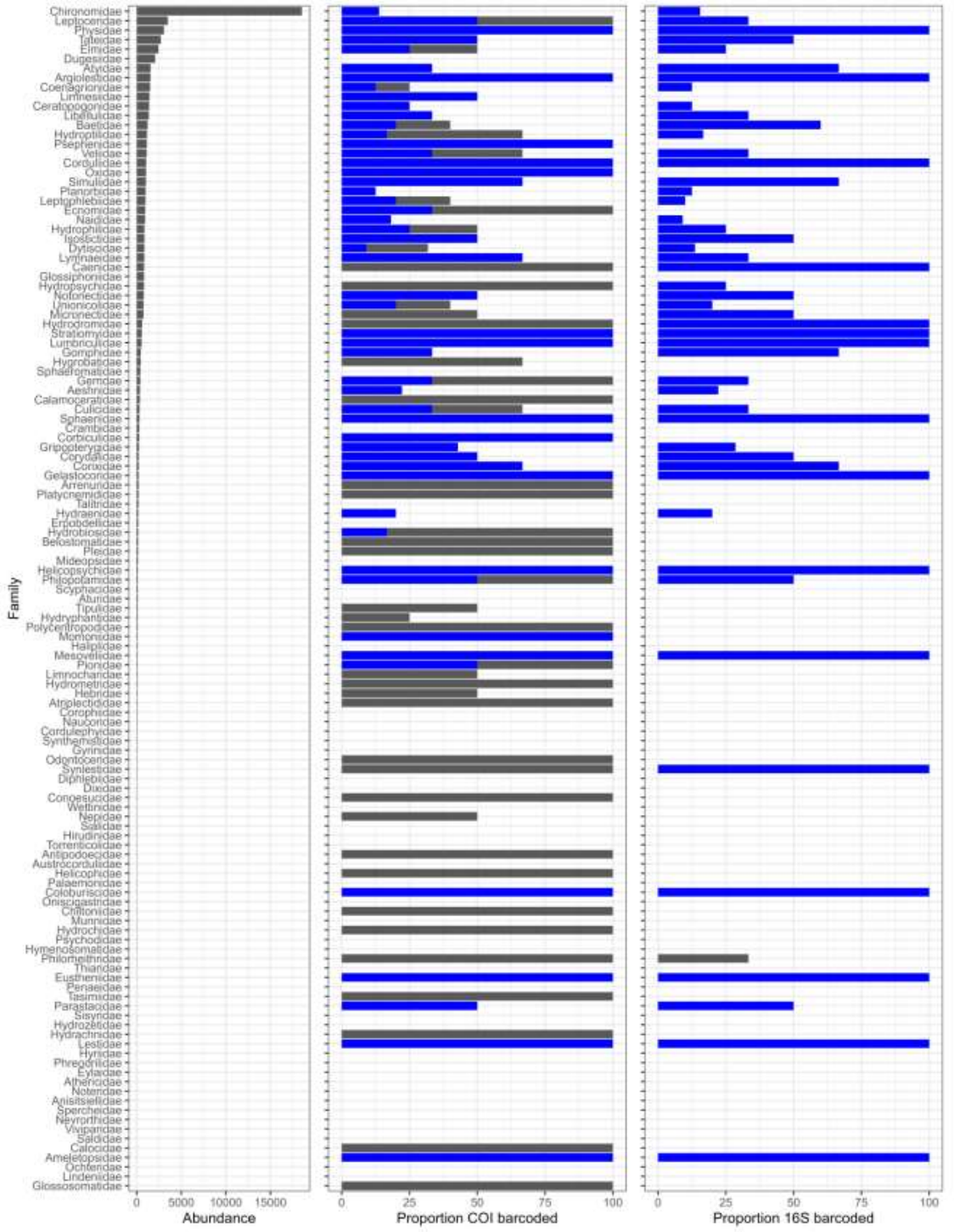


Figure 5-45: Abundances of families and their representativeness in COI and 16S DNA barcode libraries as a proportion of the number of genera barcoded.

Blue bars represent the proportion of genera from the Sydney region while grey bars represent the proportion of genera represented in the AIA database. This figure was reproduced from Figure 2 of Shackleton et al., 2022.

Summary

The addition of sequences for 193 specimens greatly increased the representativeness of taxa within the Sydney Water area of operations into the AIA DNA barcode library. Of particular importance are those that provide the first additions of taxa into the database, such as Decapoda, Ephemeroptera, and Plecoptera. The DNA reference database includes COI references from 158 genera of the Sydney region belonging to 82 families. As of this report, the DNA reference database now contains many individuals for which 16S data are available. These span 74 genera in 50 families. Over 380 genera of macroinvertebrates have been recorded over various Sydney Water projects, and around 200 are routinely collected in Sydney region based upon morphometric identifications under Sydney Water Aquatic Monitoring Program to assess WWTP discharges. The current study has increased the number of molecular specimens from the Sydney Water region represented in the AIA from 253 to 446, with 53 additional genera from the macroinvertebrate-sized taxa.

Comparing the molecular genera results to morphometric genera suggested that additional barcoding effort is required to further increase coverage within the AIA for the Sydney region. The gap analysis showed that while the most regularly collected genera were well represented in the DNA barcode libraries, barcodes were missing from many of the less regularly collected and from families that contained many genera. Future effort could focus on the Chironomidae, which are highly prevalent but for which only 13.8 and 15.5% of genera from the Sydney Water region have been barcoded for COI or 16S, respectively. Chironomids are an important group at the genus level as they provide a wide intrafamilial variation in tolerances in SIGNAL-SG grades (see Section 5.5.4 for further discussion). There are also many families that are not barcoded, and future effort could aim to barcode at least one genus from each family.

While the COI gene fragment has the greatest representation across species in DNA databases and is commonly used for species identification, the 16S gene is likely to become an important marker for eDNA or community DNA studies. COI suffers from strong PCR bias among taxa and detection of off-target taxa in eDNA samples, which can often dominate and swamp out target taxa. On the other hand, 16S is likely to exhibit less PCR bias and detect fewer off-target taxa (Elbrecht et al., 2016; Marquina et al., 2019).

Recommendation

A gap analysis should be conducted between DNA-barcoded taxa hosted in the AIA and the common taxa of the Sydney Water Aquatic Monitoring program. A further round of DNA barcoding should be undertaken to infill those defined gaps with a view to implementation of the ‘renovate’ Biomonitoring 2.0 into that long-term regulatory project. Some recommendations on further comparison studies with morphometric data are outlined at the end of the next Section 5.5.3, while considerations outlined under topic heading ‘Application in biomonitoring’ in Section 5.5.1 need to be applied.

5.5.3 Evaluation from raising SIGNAL-SG scores with environmental DNA

Objectives

The main objective of this section was to document the number of genera with SIGNAL-SG grades obtained from conventional morphometrically identified taxa and from those genera detected from post sequencing bioinformatic processing of COI and mitochondrial 16S against DNA taxonomy databases.

A secondary objective was to look for the same patterns of impairment at the downstream sites as contained within the control charts raised by:

- calculating sample SIGNAL-SG scores based on the genera with SIGNAL-SG grades in the COI and mitochondrial 16S DNA barcode datasets; and
- plotting ecosystem health control charts for the study streams of Vineyard Creek (Section 5.3.1) and Avondale Creek (Section 5.3.3)

Morphometric identifications to genus have been adopted as a precautionary approach under the WWOM project and mirrored the approach in the Sydney Water Aquatic Monitoring program. Chessman et al. (2007) stated that while a greater cost is incurred for genus level identifications, only a small difference in sensitivity between the family- (SF) and genus- (SG) level SIGNAL indices was observed for the Sydney region under bioassessment application. They further commented that finer-level morphometric taxonomy may be justified only in special circumstances, such as detection of subtle impacts. Revisiting these comments with Dr Bruce Chessman (pers. comm.) in setting up the WWOM, genus-level morphometric taxonomy was considered prudent to balance cost with sensitivity to detect potentially subtle adverse ecological effects from WWO spills over those from stormwater that have shaped the background taxonomic assemblages in urban streams. This consideration governs the decision to compare taxonomy assignments at the genus level in this report section with future application in the Sydney Water Aquatic Monitoring program as the NSW EPA would expect at least an equivalent replacement.

Molecular taxonomic assignment

GenBank was used for both mitochondrial cytochrome c oxidase (CO1) and mitochondrial 16S taxonomic assignment. In brief:

CO1

- blastDB=${code}/COI/GenBank_10-6-21_COI_id_C98_BLAST_64.udb
- blastTaxo=${code}/COI/GenBank_10-6-21_COI_C98_taxonomy.txt
- geneFilter=${code}/COI/GenBank_10-6-21_COI_C98_20.mer
- geneBLASTdb=${code}/COI/GenBank_10-6-21_COI_species_C98_BLAST_64.udb

mitochondrial 16S

- blastDB=${code}/m16S/GenBank_28-1-22_m16S_id.fa
- blastTaxo=${code}/m16S/GenBank_28-1-22_m16S_taxonomy.txt
- geneFilter=${code}/m16S/GenBank_28-1-22_m16S_20.mer
- geneBLASTdb=${code}/m16S/GenBank_28-1-22_m16S_species.fa

Known limitation of COI

DNA metabarcoding relies on PCR amplification of a fragment of a gene with primers, such as from COI. The advantage of COI is its taxonomic resolution and the availability of an extensive reference database. However, when universal COI primers are employed, co-amplification of abundant nontarget taxa (for example, fungi, algae, and bacteria) can occur (Deagle et al., 2014; Elbrecht et al., 2016; Leese et al., 2020). Elbrecht et al. (2016) suggested that the estimation of biomass might be less biased with 16S than with COI, although variations in read abundances of two orders of magnitudes were observed in their comparison. They suggested that the primer choice depended on the scientific question. If the goal is to obtain a taxonomic identification at the species level, then COI is more appropriate. While COI benefited from more established reference DNA databases use of general COI primer pairs is limited by existing gaps in DNA databases together with some taxa not amplifying. If the goal is to obtain a more comprehensive survey, the mitochondrial 16S marker (which requires building a local reference database) or optimised COI primers could be more appropriate (Elbrecht et al., 2016). These thoughts together with the outcomes of Section 5.5.1 have influenced inclusion of both COI and mitochondrial 16S in eDNA assessment among the six primer-pairs employed to assess freshwater sediment samples (Section 5.6) and estuarine sediment samples (Section 5.7).

Comparison of the number of genus level taxa with SIGNAL-SG grades

The traditional morphometric dataset was trimmed to include samples from the same six collection periods to match the sediment samples that had eDNA extracted from them and then sequenced from across 23 streams. Two replicate samples were collected per site for the traditional approach (636 samples), while four replicate samples (on all but a few occasions) were obtained from a site under the eDNA approach (1228 samples).

A total of 165 genus-level taxa were identified in the traditional hand-picked samples that had assigned SIGNAL-SG grade numbers, based upon the list of taxa in Chessman et al. (2007). This contrasted to 50 genera from the COI eDNA dataset and 30 genera from the mitochondrial 16S dataset that had assigned SIGNAL-SG grades. These returned genera under COI and mitochondrial 16S represented 30% and 18%, respectively, of the returned morphometric identified genera contained in companion-collected traditional live-picked field samples.

There were no visual differences evident within ecological control chart plots between upstream and downstream sites across collection periods for both Vineyard Creek (Figure 5-46) and Avondale Creek (Figure 5-47) based on DNA-assigned taxonomy. This is in direct contrast to ecological control chart plots based upon morphometric identification for Vineyard Creek (Figure 5-10) and Avondale Creek (Figure 5-18). These morphometric-based plots revealed dissimilar upstream and downstream site patterns across the collection period, and illustrated poorer ecosystem health at both downstream sites of these two urban streams where adverse ecological effects were attributable to sewer influent inflows (Section 5.3.1 and Section 5.3.3).

These comparisons raised the following two questions. Was the lack of separation of upstream and downstream site sample SIGNAL-SG scores derived from DNA barcode data due to:

1) the low percentage of genera detected with assigned biotic index grades? or
2) DNA barcode data returning only presence-absence data on which to base comparisons?

To explore the second question, count data in the morphometric dataset were converted to presence absence data and control charts were raised for Avondale and Vineyard creeks (Figure 5-48). In both plots, for most collection periods a distinction could be seen between upstream and downstream sites in each creek. The outcome suggested that the lack of separation of upstream and downstream site sample SIGNAL-SG scores derived from DNA barcode data, was mostly due to the low percentage of genera detected (that had assigned biotic index grades).

Summary

Results from these initial comparisons between morphometric and eDNA datasets of taxa clearly indicated low yields of 18% and 30%, respectively, of genera with assigned biotic index grades from taxonomy assigned from mitochondrial 16S and COI based on eDNA extracted from sediment samples.

The lack of detection of impaired downstream assemblages in ecological control charts based on eDNA data clearly indicated that these were not a viable option for implementation of Biomonitoring 2.0 under a renovate approach for Sydney Water. Ideally at least equivalent or more sensitive outcomes are required from genomic methods to satisfy the NSW EPA before implementation of Biomonitoring 2.0 under a renovate approach.

Another validation option would be to assign taxonomy using the Aquatic Invertebrates of Australia (AIA) DNA database hosted in Barcode Of Life (BOLD) database. This would look at taxa barcoded as outlined in Section 5.5.2.

Further pilot studies are recommended to evaluate more optimised primer-pairs (amplicons) that better target macroinvertebrate taxa if a renovate approach is to be pursued. The next Section 5.5.4 outlines a potential pilot study with an optimised primer-pair for evaluation with both community-DNA and eDNA.

Metagenomic sequencing discussed in Section 5.5.3, may enable a measure of quantitation to be obtained from DNA data. Sequencing results from metagenomics should also have taxonomy assigned from the AIA to avail of enhanced DNA-database taxonomy outlined in Section 5.5.2.

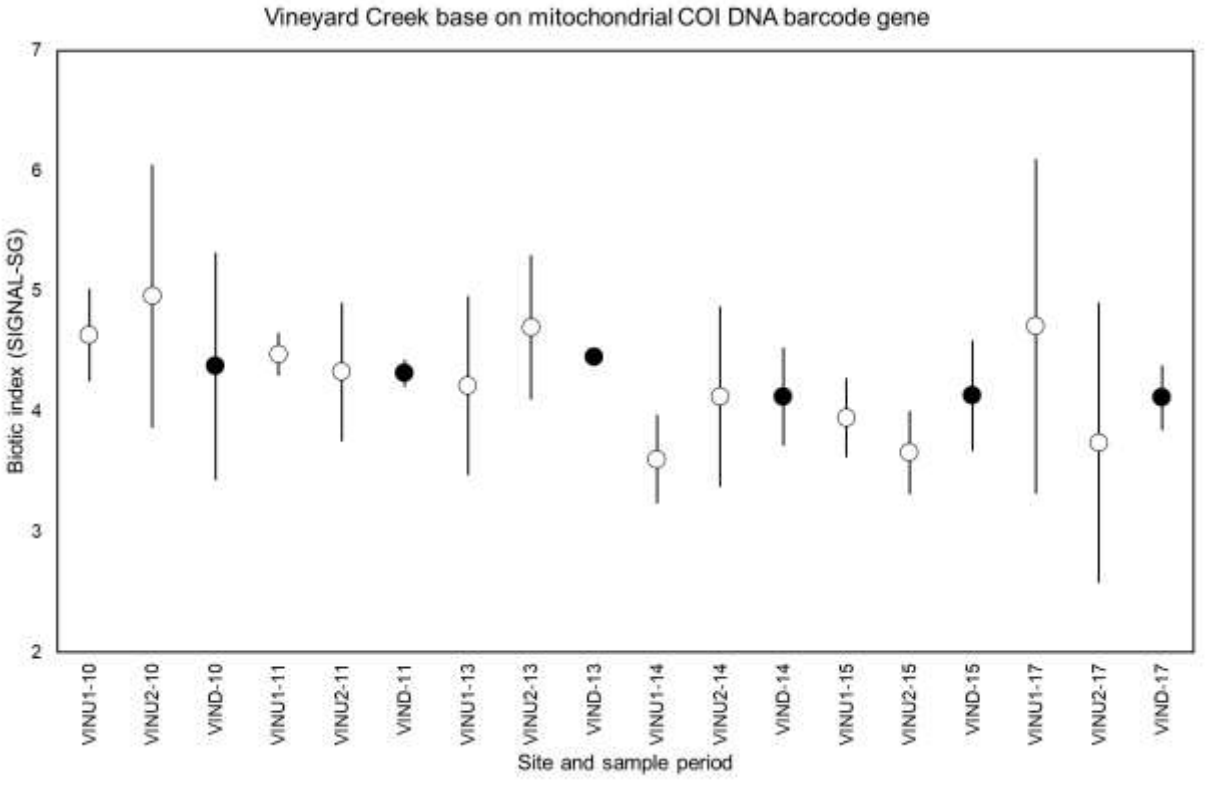
At present, a ‘rebuild’ approach to raise a new genomic-based version of SIGNAL does not seem practical until some of these issues can be resolved.

References

Deagle, B.E., Jarman, S.N., Coissac, E., Pompanon, F., Taberlet, P. 2014 DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. Biol. Lett. 10, 20140562. http://dx.doi.org/10.1098/rsbl.2014.0562

Elbrecht, V., Taberlet, P., Dejean, T., Valentini, A., Usseglio-Polatera, P., Beisel, J., Coissac, E., Boyer, F., Leese, F. 2016. Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects. *PeerJ* 4, e1966 https://doi.org/10.7717/peerj.1966

Leese, F., Sander, M., Buchner, D., Elbrecht, V., Haase, P., Zizka, V.M.A. 2021. Improved freshwater macroinvertebrate detection from environmental DNA through minimized nontarget amplification. Environ. DNA. 3, 261–276. https://doi.org/10.1002/edn3.177



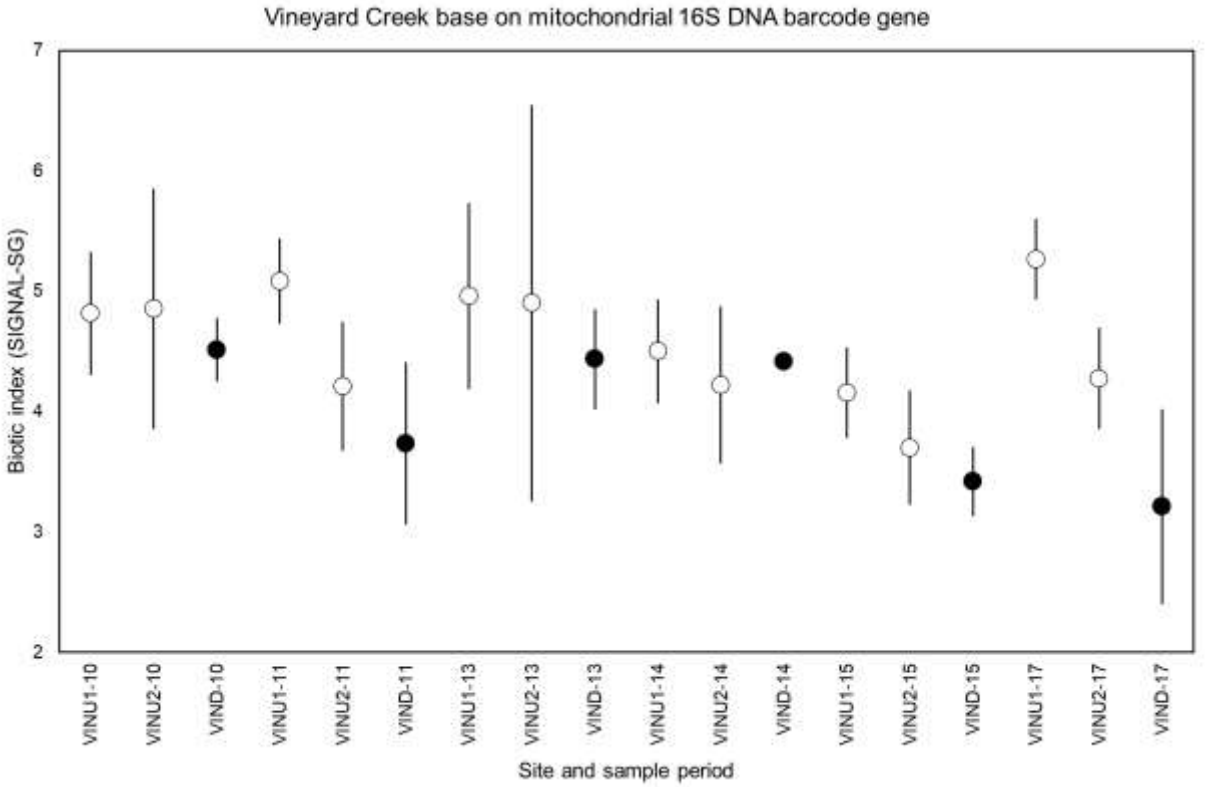


Figure 5-46: Ecological control chart plots of Vineyard Creek: upper based on COI DNA barcode gene; lower based on upper base on mitochondrial 16S DNA barcode gene

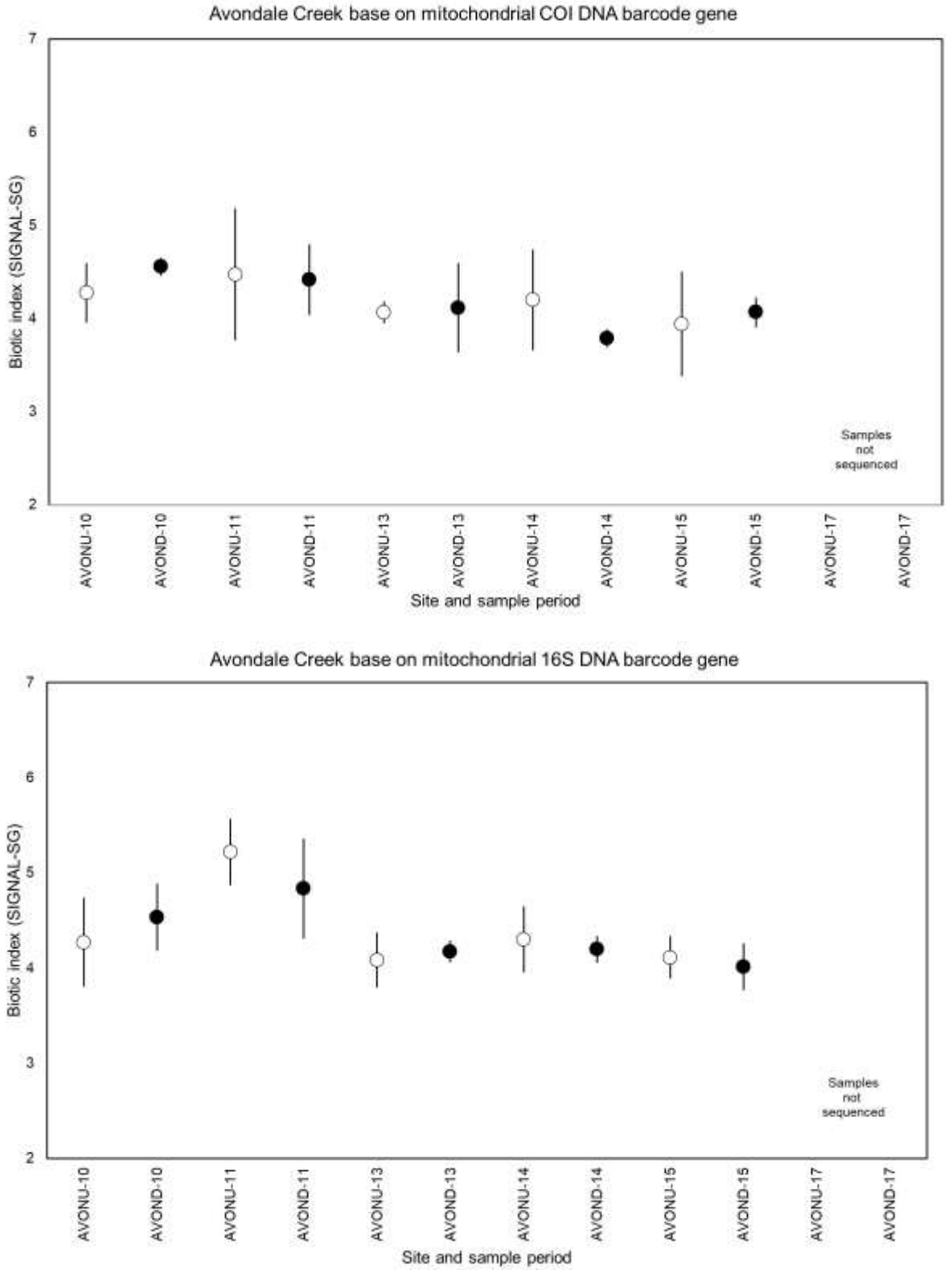
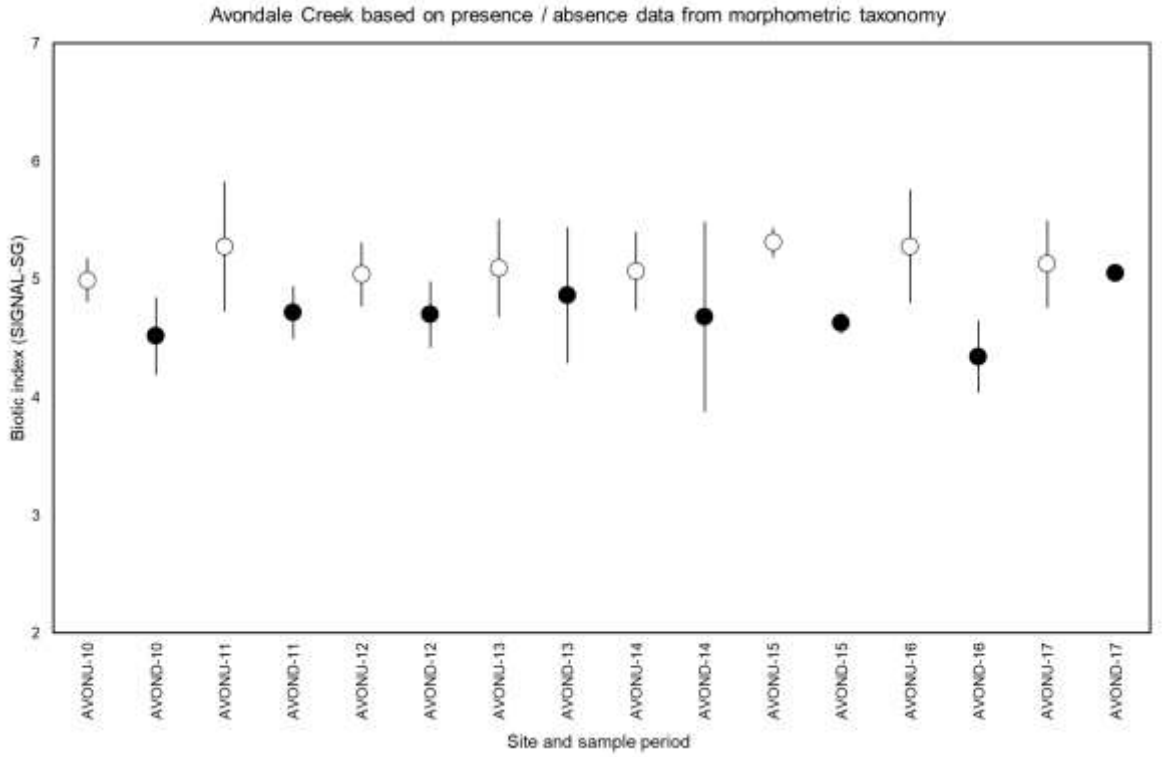


Figure 5-47: Ecological control chart plots of Vineyard Creek: upper based on COI DNA barcode gene; lower based on upper base on mitochondrial 16S DNA barcode gene



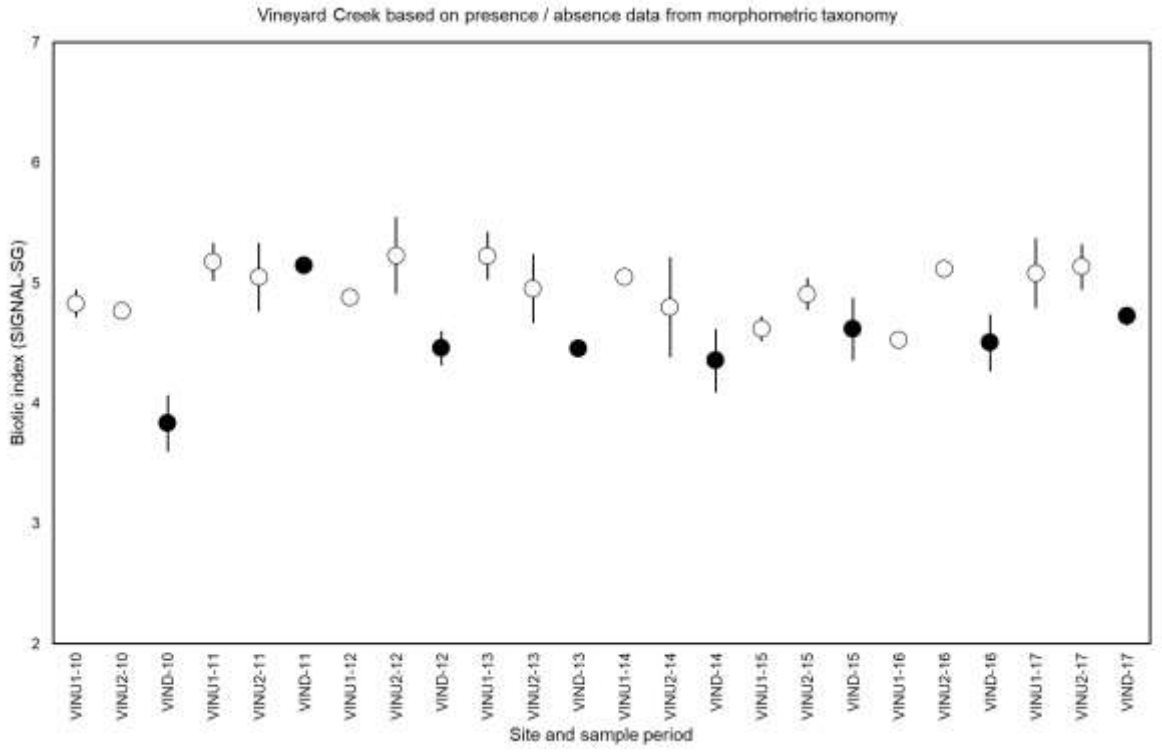


Figure 5-48: Ecological control chart plots based on presence / absence data from morphometric taxonomy: upper Avondale Creek; lower Vineyard Creek

5.5.4 Molecular refinement for enhanced recovery of invertebrate assemblages

Recoveries of mostly nontarget taxa, especially diatoms, are reported from several studies using universal COI primers under metabarcoding (Beentjes et al., 2019; Deiner et al., 2016; Gleason et al., 2020; Hajibabaei, et al., 2019; Macher et al., 2018), leaving only a small number of reads and OTUs of target taxa recovered from the extracted eDNA. Thus, under broad eukaryotic markers, the diversity of benthic invertebrate species can be underrepresented and read numbers dominated by meiofaunal groups that live in the habitat (Rotifera, Copepoda; see Li et al., 2018) but which are usually not part of the regulatory biomonitoring.

The WWOM COI dataset contained 24030 rows (sequence reads) of eukaryotes with 251 of those rows containing 112 metazoan genera confirmed at a 95% match threshold from 8711 rows of Metazoa. These 113 genera included: 10 molluscs, 34 annelids, 68 arthropods and one platyhelminth. Twelve chironomid genera were returned compared with 37 chironomid genera from the morphometric dataset. Within the Metazoa returns included Platyhelminthes (186 rows), other worms including Gastrotricha (199 rows), Nemertea (12 rows), Nematoda (141 rows), sponges (Porifera 348 rows), Tardigrada (64 rows), Rotifera (551 rows) and Chordata (68 rows). Among Chordata were the brown rat, cow, rabbit, birds including spotted dove, dusky moorhen, common myna, African sacred ibis, while fish included both short and long finned eels, introduced goldfish, eastern mosquitofish, Cox's gudgeon, striped gudgeon, firetail gudgeon, and flathead gudgeon. These eight fish species represent 12 of the species regularly surveyed by electrofishing in Botany Wetlands since 2009. The exceptions are the Empire gudgeon, catfish, Australian smelt and the introduced European carp.

Examples of non-target taxa within the COI dataset from kingdoms other than Metazoans included 14557 rows of:

- Alveolata (27 rows) dinoflagellates
- Excavata (71 rows) of the subphylum Tetramitia
- Fungi (5584 rows)
- Hacrobia (130 rows) a type of algae (Cryptophyta)
- Plantae (1132 rows) comprised of two types of algae (Chlorophyta and Rhodophyta)
- Protozoa (2239 rows) includes amoebas and apusominads
- Rhizaria (66 rows) types of Cyphoderiidae
- Stramenopiles (5308 rows) mainly diatoms

The WWOM mitochondrial 16S dataset also contained a large number of sequence reads (11898 rows) but these represent about half the reads of COI. A similar limited return of 286 rows contained 108 genera at a 95% match threshold for target macroinvertebrate genera. These 108 genera included 9 molluscs, 32 annelids, 53 arthropods along with other worms of the phyla Gastrotricha, Nemertea and Platyhelminthes, a Tardigrada, and five non-target Chordata including frog, cow, sheep, black rat, possum and humans. Two chironomid genera were returned compared with 37 chironomid genera from the morphometric dataset.

A total of 58 chironomid genera have been morphometrically identified from collections within the Sydney region and assigned SIGNAL-SG grades ranging from 3 to 9 out of the one to ten grade range under this biotic index, indicating wide intrafamilial variation in tolerances for these taxa

(Chessman et al., 2007). This and other families with genera having similar wide intrafamilial variation contribute to detecting subtle differences in ecosystem health levels between sites. This suggests that improving DNA recovery of taxa at the genus-level within families that have these intrafamilial variation signatures, would improve Bioassessment 2.0 outcomes. These signatures may also exist in other taxonomic groups not traditionally assessed, such as the meiofauna.

While optimised primer pairs should assist in better recovery of target taxa, metabarcoding remains dependent upon PCR, and the resulting lack of quantitative data will be another area to refine. As Shackleton et al. (2022) highlights, '*PCR bias occurs when the level of homology between primer sequences and priming sites differs among taxa. PCR bias results in the over amplification of some taxa and the under amplification of other taxa and can result in widely different read numbers among taxa that are present in samples in equal proportions. This bias is known to be one of the main impediments to deriving quantitative data from eDNA methods (Krehenwinkel et al., 2017)*'.

Text and citations in this section were drawn from Leese, F., Sander, M., Buchner, D., Elbrecht, V., Haase, P., Zizka, V.M.A. 2021. Improved freshwater macroinvertebrate detection from environmental DNA through minimized nontarget amplification. Environ. DNA. 3, 261–276. https://doi.org/10.1002/edn3.177, and

Chessman, B.C., Williams, S.A., Besley, C.H. 2007. Bioassessment of streams with macroinvertebrates: effect of sampled habitat and taxonomic resolution. J. North Am. Benthol. Soc. 26, 546-565.

Pawlowski, J. Bonin, A., Boyer, F., Cordier, T., Taberlet, P. 2021. Environmental DNA for biomonitoring. Mol. Ecol. 30, 1-6. https://doi.org/10.1111/mec.16023

Shackleton, M. E., McPhan, L., Murphy, N. P. 2022. DNA barcoding and investigation into PCR bias among Sydney Water macroinvertebrates. La Trobe University, Centre for Freshwater Ecosystems, Wodonga, Victoria. Final report 275 for Sydney Water.

Optimised primer pair

An example of a more appropriate optimised metabarcoding primer-pair that avoids nontarget amplification bias and substantially improved invertebrate detection based upon eDNA extracted from water samples outlined by Leese et al. (2020).

Leese et al. (2020) developed a new reverse primer (EPTDr2n) with specificity toward benthic invertebrate taxa and validated its specificity in silico together with universal forward primer fwhF2. This was then evaluated against two other COI primer-pairs. They found that the percentage of target reads was much higher for the new primer combination compared to two universal benthic invertebrate primer pairs, BF2/BR2 and fwhF2/fwhR2n (99.6% versus 25.89% and 39.04%, respectively). Likewise, the number of detected benthic invertebrate species was substantially higher (305 versus 113 and 185) and exceeded the number of 153 species identified by expert taxonomists at nearby sites across two decades of sampling. While some taxa, such as flatworms were not detected, they showed that the optimised primer avoided the non-target amplification bias and thus significantly improved benthic invertebrate detection from eDNA.

Leese et al. (2020) concluded that the detection of benthic invertebrate taxa from eDNA isolated from stream water, is greatly improved with a new specific primer combination that avoids non-target taxa amplification. The revised primer offers a solution to the common problem of

watered-down benthic invertebrate biodiversity in COI eDNA metabarcoding based on water samples and may thus improve future biodiversity assessment and monitoring of freshwaters (Leese et al., 2020).

While the new primer combination of Leese et al. (2020) provided a substantial gain in arthropods, in particular insect DNA amplification, the increased specificity also had a down-side because using the new EPTDr2n also excluded some derived arthropod target taxa that had the same nucleotide as the nontarget taxa. Likewise, flatworms and molluscs were also not reliably detected with the new primer combination. However, they advised caution that a non-detection of eDNA from a sample does not mean that a taxon is not there, but rather that we are still in the optimisation phase of method development.

Leese et al. (2020) suggested that the choice of primers is always a difficult quest and depends on the aim. Clearly, no “one-fits-all” solution exists (Clarke et al., 2017; Elbrecht et al., 2019; Elbrecht and Leese, 2017a; Grey et al., 2018; Hajibabaei et al., 2019; Taberlet et al., 2018). If the aim of a study is to capture the greatest number of benthic invertebrate taxa, in particular insects from eDNA for bioassessment with one primer pair, the newly designed fwhF2/EPTDr2n combination fits that purpose. If the aim is to maximise the phylogenetic diversity captured in a sample, either a more conservative marker such as 18S can be appropriate (Bagley et al., 2019; Deagle et al., 2014; Li et al., 2018), or for metazoan taxa it might be important to consider the use of multiple COI primers (Corse et al., 2019; Hajibabaei et al., 2019) as with the more specific primer combination, clearly a few macrozoobenthic taxa within Trichoptera, Mollusca, and Isopoda will be missed due to primer bias. These thoughts from Leese et al. (2020) support subsequent professional advice (Dr Anthony Chariton) to include a mix of somewhat targeted primers along with a universal COI primer pair in the WWOM. The following primer pairs (amplicons) for eDNA extracted from freshwater sediment samples targeted the below listed taxonomic groups:

- 16SV4 Bacteria
- Ant18S Eukaryotes
- COI Eukaryotes
- mitochondrial 16S Eukaryotes
- 18S Diatoms
- RBLC Diatoms

For eDNA extracted from estuarine sediment samples these primer pairs enabled detection of:

- 16SV4 Bacteria
- Ant18S Eukaryotes
- COI Eukaryotes
- 18S Diatoms
- Amphipods
- Polychaetes

Summary and recommendation

Both COI and mitochondrial 16S WWOM datasets of extracted eDNA from freshwater sediment samples unsurprisingly contain many non-target taxa as has been found in numerous previous studies. A low level of recovery of genera was observed under bioinformatics (Section 5.5.4) and a smaller subset of these genera had assigned SIGNAL-SG grades (Section 5.5.3). Both of these outcomes clearly reflect that optimisation in recovery of target macroinvertebrate taxa is required from metabarcode sequencing to enable application of a 'renovate' approach under Biomonitoring 2.0. The following two options are recommended for future pilot studies to evaluate recovery of genera under the primer pair raised by Leese et al. (2020) against Sydney region samples from metabarcoding sequencing.

Proposed future pilot study options under the Sydney Water Aquatic Monitoring program

1. Conduct sequencing of reserve aliquots of already extracted eDNA from previously collected WWOM sediment samples and compare to community-DNA extracted from previously collected WWOM bulk samples of hand-picked macroinvertebrates. This would allow evaluation of the morphometric taxonomy dataset with molecular taxonomy datasets from: sediment eDNA; and community-DNA
2. Collect new water and sediment samples at the same time as hand-picked macroinvertebrates are collected for the Sydney Water Aquatic Monitoring program. This would allow evaluation of the morphometric taxonomy dataset with e-DNA datasets from water and sediment, and with community-DNA

From a cost perspective, option 1 would be relatively cheaper as both traditional hand-picked macroinvertebrate samples along with companion sediment samples have already been collected, and DNA has been extracted from sediment samples. Comparison to the existing WWOM sequenced COI primer dataset is also available. Whereas that would of an additional cost under option 2. If improved recovery of genera with SIGNAL-SG grades is not obtained from sediment samples under option 1, then sediment samples would be omitted from option 2.

5.6 Environmental DNA assemblages in urban streams

The initial draft of text within this section of the synthesis report was kindly provided by the Macquarie University project team.

5.6.1 Overview of eDNA metabarcoding as biomonitoring tool

There is growing interest in the use of molecular tools, for example environmental DNA (eDNA) metabarcoding (here on referred to as metabarcoding) to capture a more comprehensive view of biodiversity. In sedimentary environments, metabarcoding can be used to potentially capture all of life; and depending on the approach, this can include bacteria, small eukaryotes, including diatoms, as well as macrobenthos. Thereby, enabling community assessments to be derived from thousands to tens of thousands of taxa. Metabarcoding is based on targeting and amplifying taxonomically informatic regions of genes from DNA extracted from an environmental sample. For example, a region of the 16S rDNA gene is used to target bacteria, while the COI gene, and others, focuses on metazoans. The products of numerous samples are then simultaneously sequenced using a high-throughput sequencer (Taberlet et al., 2018). The potential for using metabarcoding as a biomonitoring tool for benthic eukaryotes was first demonstrated by Chariton et al. (2010). Numerous studies have since demonstrated the utility of eDNA metabarcoding as a tool for assessing and monitoring aquatic systems. The technique is now being developed as part the European Union's (EU) Water Directive Framework (Hudson, 2008). In estuarine environments, metabarcoding has been shown to be 55% less costly and 72% less time consuming than traditional (morphology based) macrobenthic studies (Aylagas et al., 2018).

While metabarcoding is providing a paradigm shift in the way aquatic systems are being assessed, it is not without limitations (Cordier et al., 2020). Firstly, accurate taxonomic assignment is determined by the quality and relevance of sequences available in online and custom depositories. And as highlighted in Section 5.5.2, for a vast majority of regionally relevant aquatic taxa, sequence data are not available. Secondly, the genes and loci targeted, as well as the bioinformatics pipeline can profoundly affect the compositional matrix used for downstream interpretation. This is particularly true where 'universal primers' are used for levels of taxonomic assignment below that which is realistically viable (Taberlet et al., 2018). Thirdly, most published examples are study-specific and cannot be upscaled as they are based on operational taxonomic units (OTUs), which are clusters of similar sequences unique to that study. While this *de nova* approach is powerful for discriminating communities exposed to varying natural and anthropogenic stressors, the lack of formal taxonomy associated with OTUs often hinders the transferability of the findings to other studies (Chariton et al., 2015).

Globally, there are a number of initiatives and projects developing scalable taxonomically based metabarcoding indicators for both freshwater and coastal environments. However, to date, these have not been developed or applied for monitoring WWOs. For example, in Australia there is an initiative to explore the possibility of replacing the morphometric endpoints of AUSRIVAS with metabarcoded macroinvertebrate data (Nichols et al. 2019). AUSRIVAS, regardless of whether its traditionally or eDNA derived it is not applicable to coastal freshwater or estuarine environments nor is it designed to monitor episodic responses (Hose et al., 2004, Besley and Chessman, 2008). Furthermore, such approaches should be viewed as a replacement of traditional macrobenthic surveys and not a true metabarcoding survey, as they only focus on the taxa already used in the traditional programs, negating the tens of thousands of other taxa which make up the sampled

communities (for example bacteria, diatoms and metazoans). This is critically important as it has been repeatedly shown that microbial (eukaryotes and prokaryotes) and meiofauna are better indicators of environmental condition than macrofauna (Chariton et al., 2015; Sutcliffe et al., 2019).

One interesting metabarcoding index is micro-AMBI, which is derived from metabarcoded bacterial data (Borja, 2018). However, micro-AMBI is based on the sensitivities of bacteria taxa to metals, organics and salinity in estuarine environments, and consequently the metrics have not been specifically developed or tested on comparing the various states associated with wet-weather overflows. Taxonomic-based eDNA indices have some significant challenges. Firstly, they require a reference database, and as such a vast majority of sequences are excluded as they lack taxonomic assignment (Chariton et al., 2015; Graham et al., 2019). Secondly, and most pronounced in eukaryotes, is the poor relationship between taxa abundance (or biomass) and the number of sequence reads, especially when comparing the relative abundances between taxa (Vivien et al., 2015, Dowle et al., 2016). This can make it challenging to develop a single marker (for example one gene or loci) indices which are derived from either quantitative responses to a stressor(s) or the relative abundances between taxa (Cordier et al., 2018). A novel way to circumvent these issues is via the use of taxonomic-free machine-learning based approaches (Cordier et al., 2018). These approaches ignore the taxonomy, but still use the sequences, in essence, they are your taxonomy. For example, Cordier et al. (2017) demonstrated this approach comparing benthic foraminifera eDNA and macrobenthic faunal responses to aquaculture (Cordier et al., 2017). Cordier et al. (2019) went on to show that taxonomic-free machine-learning approaches using a range of single eukaryote and prokaryote markers outperformed those based on traditional taxonomy (Cordier et al., 2019). Consequently, machine learning based approaches show great promise for classifying conditions for wet-weather overflow events, with the caveat being that a priori conditional states need to be identified, that is, clear states are defined using representative samples to train the model.

Regardless, if using traditional statistical multivariate techniques or machine learning for comparing system exposed to varying states of wet-weather events, some common attributes are required. This includes: the need to use a range of loci and genes to target the different taxonomic communities (for example bacteria, diatoms and metazoans); and robust and reliable abiotic data which reflect both catchment and in-stream conditions. In some cases, abiotic variables will be static, that is, will not markedly change during the length of the study, for example percentage of impervious surfaces and catchment area. In other cases, abiotic variables will be dynamic, reflecting different temporal trends in wet-weather events, for example rainfall, flow, and ideally, the volume and duration of untreated waters. To reiterate, eDNA-based machine learning approaches such as neural networks are founded on assigning samples to predefined states, and then predicting their ability to assign new samples to those states. Other approaches are available that look at correlative relationships (random forest), although the approach is novel, it is within the context of eDNA aquatic biomonitoring.

Objectives of the freshwater biomonitoring study

In this study, we used eDNA biomonitoring to examine benthic communities from a range of freshwater sites across the Sydney catchment exposed to varying levels of wet-weather events. This includes, dry periods, small rainfall events, and pronounced rainfall events, where unprocessed waters are likely to enter the system. Specifically, we aimed to:

1. Examine spatial and temporal patterns in biotic communities (for example bacteria and diatoms) across sites and catchments
2. Identify the environmental variables, both static and dynamic, which were correlated with eDNA metabarcoded communities
3. Compare machine learning derived predictions from metabarcoded data (structure and function) scores against traditionally morphometric data (SIGNAL-SG)
4. Examine the potential for applying metabarcoded derived diatom data to assess the condition of sites
5. Explore the potential for machine learning to predict community composition from the abiotic and biotic attributes of the system (Table 2 in Section 5.6.2).

In addition to the ecological-based questions described above, we also examined the spatio-temporal characteristics of the data as well as its quality, ensuring that it was suitable for addressing these questions. Details regarding this approach are provided in Section 5.6.2. Given the immensity of the data and analysis, the methods and findings here are restricted to the bacterial and diatom datasets, as these were deemed the most suitable for addressing the aims of this study.

5.6.2 Overview of sampling, sequencing, bioinformatics and statistical approaches

Experimental design

Sediment samples were taken from 56 sites on six occasions (here on referred to as a collection period). The samples were collected during periods reflecting a range of rainfall and flow conditions between March 2020 and April 2021. Sites included both upstream (prior to major exposure inputs) and downstream positions. The assumption was that upstream sites would be markedly less exposed to any potential sewage inputs during wet-weather events. At each site, three sediment samples were taken for eDNA. DNA extracted from each sample was amplified for six loci/genes to capture a wide range of biota: 16S rDNA (bacteria); 18S rDNA (eukaryotes); 18S rDNA (diatoms); rbcl (diatoms); mt16S (invertebrates) and COI (metazoans). For each loci/gene, 3 separate polymerase chain reactions (PCRs) were performed. In addition, 1,152 synthetic sequences (positive controls) and 6,912 negative controls (blank samples using DNA-free water) were sequenced for quality assurance. Samples were sequenced in two parallel lanes (same sample in both lanes) on an Illumina NextSeq800 at the Ramaciotti Centre for Genomics (UNSW). Due to the substantial number of samples, each of the three PCR replicates was sequenced on separate run.

Bioinformatics

Bioinformatics was performed using the Greenfield Hybrid Amplicon Pipeline (GHAP), an in-house clustering and classification pipeline build around tools from USEARCH (Edgar 2010) and RDP (Cole et al., 2014), combined with locally written tools for demultiplexing and generating operational taxonomic unit (OTU) tables (Greenfield Hybrid Analysis Pipeline (GHAP)). The poor-quality tails of each read were trimmed prior to pair-merging. The merged reads were then filtered, the reads were then clustered at 97% similarity using the cluster_otus command in Usearch v11.0.667 (Edgar 2010). Representative OTU sequences from each OTU were assigned a taxonomic lineage using the RDP Naïve Bayesian Classifier with a minimum confidence threshold of 50% (Cole et al., 2014) and matched to its closest sequence in a RefSeq-based set of reference sequences. Functional inferences of 16S rDNA (bacteria) were performed using the package FAPROTAXA. In addition to using the sequenced diatom compositional data (Louca et al., 2016).

Statistical approaches

Visualisation of the metabarcoded communities was examined using Principal Component Analysis (PCA). PCA is a dimensionality reduction technique used to transform high dimension data into a lower dimensional space whilst still preserving as much information as possible. The number of variables is reduced by the analysis by identifying a new set of orthogonal axes, known as principal components (PCs), which represent the directions of maximum variance in the data. Each PC is orthogonal to each other in terms of variance. A Hellinger transformation was applied to the data before carrying out the PCA using the scikit-learn package v1.2.0 from Python (Virtanen et al., 2020). PCAs were performed for each gene/loci at various taxonomic levels from the OTU to phylum. Numerous PCAs were produced to examine at a range of spatial and temporal scales, capturing different catchment geologies (sandstone and shale) and differences between upstream and downstream sites. Statistical comparisons between treatments (for example temporal periods, locations and geology) were performed by PERMANOVA using the function adonis2 in the R package vegan.

The relationships between metabarcoded derived communities and environmental variables were examined using Canonical Correspondence Analysis (CCA) (Ter Braak and Verdonschot, 1995). In addition to the CCA approach, dominant gradients in species composition were identified using Generalised Linear Latent Variable Models (GLLVMs). GLLVMs extends generalised linear modelling (GLM) by incorporating latent variables (LVs) to model correlated responses among species. LVs can be constrained (like canonical correspondence analysis, CCA) or unconstrained (like detrended correspondence analysis or DCA). GLLVMs have a number of advantages over the standard CCA/DCA approaches. First, OTU counts in samples can be modeled using standard statistical distributions (for example Poisson, negative binomial). Second, standard methods can be used for evaluating model assumptions, selecting covariates (using for example, an information criterion (AIC)), and testing hypotheses (using log likelihood ratio tests). Finally, prediction uncertainty is explicitly quantified while controlling for systematic changes in variance with mean OTU abundance.

GLLVM analyses were undertaken on OTU counts aggregated to multiple of levels of taxonomic resolution for Bacteria (phylum, class, order, family, genus) and diatoms (family, genus) for the sandstone sites. Analyses were restricted to taxa found in at least 90 samples. Counts were assumed to follow negative binomial distributions. Two community-level LVs were inferred in each

GLLVM after controlling for three sources of variation: (i) sample variability within a collection period was controlled using the logarithm of the mud-gravel ratio; (ii) variation among collection periods within a site was controlled for rainfall using a 3rd order polynomial; and (iii) sequencing depth was controlled using an offset term for the log sequencing depth. After generating the latent variables using GLLVM, LVs were correlated with landscape variables to assess responsiveness to impervious area, road density, and riparian habitat. GLLVM models were fitted using the R package gllvm.

Variance partitioning of taxonomic composition

As previously stated, given the immense scale and novelty of the eDNA metabarcoding approach in this study, analysis was also performed to examine the variation in the data, capturing both technical (runs, which includes PCRs, and lanes), as well as spatial and temporal aspects of the biological data. Variance in taxonomic composition was hierarchically partitioned among the following nested factors: site (56 in total), collection period (6 per site), sample (3 per site-collection-period combination), DNA extraction (3 per sample), and lane (2 per extraction).

Prior to analysis, lane-level estimates of proportional abundance

$$Pi\text{Pi}$$

were converted to a Hellinger-distance matrix by first applying the Hellinger transformation to the proportional abundance values

$$Pi\text{¯}\sqrt{}\text{Pi}$$

and then calculating pairwise Euclidean distances between lane-level estimates of taxon composition. Variance in taxon composition was then partitioned among hierarchical levels using permutational multivariate analysis of variance (PERMANOVA) using the function adonis2 in the R package vegan. Analyses were separately conducted for bacteria (16s rDNA) and diatoms (18s rDNA) at multiple taxonomic levels including phylum, class, order, family, genus, species, and OTU.

Analysis of microbial source tracking markers

Given that no true measure of sewage inputs could be obtained (see Section 5.6.3), measurements of microbial source tracking (MST) markers (Section 3) were used as a potential surrogate for sewage. For this, two HFMGs were used *Bacteroides* HCF183 and crAssphage CPQ_056. Analysis for these two HFMGs was performed by Sydney Water using Quantitative PCR (qPCR). Statistical analysis of the two markers was performed to examine occurrences across the sites (both temporally and spatially), thereby determining their suitability as indicators of sewages, as well as examined to see whether these markers correlated with the collected static and dynamic environmental variables.

The qPCR data for each analyzed HFMG encompassed 918 samples, 53 sites, and 6 periods. Correlates of these data were identified using generalised linear mixed effects models with random effects for site and the site-collection-period interaction. Informative (that is, those that explained a substantial amount of variance) fixed and random effects were identified from among candidates using forward variable selection procedures based on the AIC criterion. Models were fitted using the R package lme4, including random effects for site and the site-collection-period interaction. Informative fixed effects were identified sequentially in three steps using a forward variable

selection procedure. First, we considered the importance of two substrate variables that varied among samples at a site within a collection period: log[(%sand)/(%gravel)], log[(%mud)/(%gravel)]). Second, we considered the importance of variables that may account for variation among collection periods within a site, including rainfall and sewage totals over various time windows ending 1 or more days prior to the day of sampling (1-5 days, 1-10 days, 1-30 days, 1-45 days, 5-10 days, 5-30 days, 5-45 days, 10-30 days, 10-45 days, 30-45 days). Third, we considered the following landscape variables that could account for variation in average conditions among sites: logit-transformed riparian habitat, log(road density), and log(impervious area). Candidate predictors were modelled using 1st-3rd order polynomials because visual inspection indicated that many relationships were non-linear. The categorical variable geology (levels: sandstone, shale) was always included, and interactions of each variable with geology was considered at each step.

Machine learning

Two machine-learning approaches (neural network and random forest) were performed in this study to: see if there was a relationship between the metabarcoded data and SIGNAL-SG (morphometric based biotic index scores); examined the potential for sites to be classified into different states based on diatom indices (Table 5-4); and to identify potential abiotic variables which were predictive of ecological condition (Table 5-5).

A neural network (NN) is a supervised machine learning algorithm that can be used for a solving complicated classification and regression tasks. The basic structure of NNs consists of layers comprised of neurons that are interconnected. The three different types of layers are known as the input layer, hidden layers, and the output layer. The input layer is the starting point that receives the initial data. The hidden layers are what perform the computations by taking input from the previous layer and applies a transformation, via weights and biases, which are then passed onto the next layer. Then the final output layer which provides the result, for example, a prediction of a biotic index. The connections between each neuron to each adjacent layer are weighted which represents the strength of the connection and thus determines how much influence the output of a neuron has to the next layer. During the training process, NNs use a method called backpropagation which adjusts weights and biases to minimise the error or loss function which yields a better performing NN. Each neuron also typically applies an activation function to its weighted sum of inputs before passing the signal on to the next layer. These activation functions introduce a non-linearity into the network, allowing it to learn complex patterns and relationships in the data. There are additional processes, such as dropout probability and batch normalisation that can be used to prevent overfitting and improve generalisation.

For the neural network approach, the TensorFlow package v2.11.0 from Python (Abadi et al., 2016) was used. The biotic data used included the 16S, 18S, and *rbcL* OTU tables, which were normalised, with SIGNAL-SG (Besley and Chessman, 2008) as the biotic index being predicted. Phylogenetic resolutions were also investigated from species, genus, family, and order. Environmental data were also investigated which included transformed data of grain size ratios, rainfall, road density, road impervious area, riparian strip width, sewer volume, and sewer duration (Table 5-5). The data were split into a training dataset, testing dataset, and validation dataset at a ratio of 6:2:2, respectively. The data were grouped in different ways prior to splitting the data into the different sets, that is, no grouping, grouping by site, and grouping by site and collection period.

Three-layer models were generated with each layer using the rectified linear activation function (ReLU) (Agarap, 2018). Multiple parameter settings were investigated by permuting through a list of settings for each hyperparameter to find the best model. The number of nodes for each layer were 16, 32, 64, or 128, the dropout probability was set to 0.1 or 0.2, learning rate was set at 0.01 or 0.02, and the batch sizes investigated were 8, 16, or 32. Each model was run for 100 epochs and compiled with the Adam optimiser (Kingma and Ba, 2014) using mean squared error as the loss function.

Random forest (RF) is a supervised machine learning algorithm that can be utilised for both classification (discrete values, for example to compare against SIGNAL-SG) and regression (continuous values) tasks. It is an ensemble learning method that uses multiple decision trees and merges their predictions to obtain accurate and stable results. The first step of the process involves splitting the data into a training set and testing set. The training set contains the input data (features), along with the corresponding output (labels), which the algorithm uses to capture patterns and relationships between the features and labels make accurate predictions when presented with unseen data. The testing set is the unseen data that were set aside to be used as a tool to evaluate the performance of the trained model. After splitting the data into their representative sets, the RF creates multiple random subsets of the training set through a process known as bootstrapping. This makes each training subset the same size as the original data, but some data points are left out, while others may be repeated. For each subset, a decision tree is created by splitting the data based on the values of different features, with the goal of minimising the measure of uncertainty or disorder in a set of data points within a node of the decision tree. After all the trees have been built and predictions made, the final prediction is determined by aggregating all the previously calculated predictions. One of the major benefits in using RF is that it can provide a measure of feature importance which is useful for understanding which features are most important in relation to the label used.

For the RF approach, the scikit-learn package v1.2.0 from Python (Virtanen et al., 2020) was used as a regression model predicting biotic indices. The input data that were investigated included the environmental data, normalised OTU tables of 16S rDNA and diatom 18S rDNA amplicons, and the putative functional profile of the 16S rDNA community using FAPROTAX v1.2.7 (Louca et al., 2016). The environmental data included grain size ratios, rainfall, road density, road impervious area, riparian strip, overflow contribution, and CrAssphage data from the MST data (these data were transformed, seeTable 5-5). The biotic data were separated by the geology type (sandstone or shale) based on the subsurface bedrock of each site. Each sample was labelled with the biotic index SIGNAL-SG (Besley and Chessman, 2008) and 23 diatom indices generated using the DiaThor R package v0.1.0 (Nicolosi et al., 2022). For each RF regression run, the data were split with a group K-folds approach into a training dataset and a testing dataset. The number of folds was set to 5 and split groups were based on site to reduce overfitting. The group K-folds setting is important as it ensures that samples that are related, in this case samples that were collected at the same site, are grouped together before splitting into the training or testing set. This guarantees that the model does not learn from related samples which may lead to overfitting or an overestimation of the performance of the model. The hyperparameters for the RF were as follows: number of trees = 1000, max features = 1.0, ‘sqrt’, and ‘log2’, minimum samples split ranged from 2-10, and the maximum samples split ranged from 1-10, number of iterations = 200, and cross-validation folds = 5.

Creating molecular-based diatom indices

There are a number of diatom indices used to assess river condition, much in the same way SIGNAL-SG is used. As in the case of SIGNAL-SG, these are derived from morphometric measurements. Here we explore the potential to use these indices using the molecular diatom data, potentially providing an additional approach for assessing the condition of sites. This information could then be incorporated into a random forest model. A total of 13 diatom indices were investigated which were obtained using the DiaThor R package v0.1.0 (Nicolosi et al., 2022). In addition, four different reference databases were used to identify the sequence reads, they were SILVA128 (Quast et al., 2013), SILVA 138 (Quast et al., 2013), EukRibo (Berney et al., 2022), and the Protist Ribosomal Reference v5.0.1 (Guillou et al., 2013). Pre-processing of the OTU tables was carried out using a custom in-house Python script to correctly match the genus and species names to the DiaThor list for accurate calculation of biotic indices.

From a comparison of the different reference sets, the EukRibo reference database had the highest percentage of recognised taxa across all the diatom indices, and hence was used to develop all indices.

Across both geologies (sandstone and shale), the transformed diatom indices all performed poorer in comparison to the untransformed versions. Not all the RF models run using the environmental variables to predict diatom indices had a positive R^2 value, meaning there was a positive correlation between the predictions and true values. For the sandstone data, there were six diatom indices with a positive R^2 value which were EPID20, EPID, IPD, PBIDW, SLA, SPEAR, and TDI. While for the shale data, there were four diatom indices with a positive R^2 value which were DISP, EPID, ILM, and IPS. From the in-depth comparisons of all indices, including modelling them using RF, and their responses to the measured static and dynamic variables, only one index, EPID, was chosen was found to be potentially suitable for inclusion in the RF model.

Environmental variables

Initially, the environmental variables used in all relevant analysis (both traditional and machine learning based approaches) were derived on what was believed would classify the sites into five broad states (Table 5-4). However, initial analysis based on these variables was deemed unsuitable, as it: failed to capture the potential number of inputs into a site; did not assist in determining whether a site was pre-exposed to untreated waters from other sources; and lacked fundamental information on catchment modification, for example the density and area of pervious surfaces and riparian cover. Consequently, additional static and dynamic variables were used, whilst others were replaced, aggregated, normalised or removed (Table 5-5).

Table 5-4: Variables and parameters used to initially classify the sites into five wet-weather overflow states

State	Stormwater exposure (rainfall mm)	Sewage exposure (duration or volume)	Water chemistry (passive samplers)
Baseline (B)	< 40 mm in 8 weeks prior to sampling*	No wet-weather overflows	Stormwater-chemical markers detected*
Urban stormwater (U)	> 40 mm in 8 weeks prior to sampling*	No wet-weather overflows	Stormwater-chemical markers detected*
Sewage overflow (low)		Wet-weather overflow duration < 1 h or < 100 kL	Sewage-chemical markers detected **
Sewage overflow (medium)		Wet-weather overflow duration 1 to 6 h or 100 to 1000 kL	Sewage-chemical markers detected **
Sewage overflow (high)		Wet-weather overflow duration > 6 h or > 1000 kL	Sewage-chemical markers detected **

**Rainfall is adjusted for the impervious area of each catchment. ** Stormwater chemical markers: N,N-dicyclohexylurea , aurodeoxycholic acid, 2-(dibutylamino) ethanol, atrazine-desethyl, chlorantraniliprole, dimethoate, dimethomorph isomers 1 and 2 flutolanil, metalaxyl, omethoate, tebuconazole, terbuthylazine, thiamethoxam, triadimefon, triadimenol, diflufenican. **Sewage chemical markers: acetaminophen, amitriptyline, atenolol, benzophenone-3, carbamazepine, carvedilol, cotinine, diclofenac, diltiazem, distyrylbiphenyl disulfonate, fluorescent brightener (FB71), gliclazide, griseofulvin, hydrochlorothiazide, ibuprofen, losartan, mefenamic acid, metformin, naproxen, octocrylene, sotalol, sucralose, sulfapyridine, theobromine, 4-methylbenzylidene camphor, 1H-benzotriazole, 4-methyl-1H-benzotriazole, 5-bethyl-1H-benzotriazole. Chemical measurements obtained from POCIS passive samples.*

Table 5-5: The refined list of environmental variables, including their abbreviations and descriptions, used for assessing the relationships with the metabarcoded derived communities

Variable	Abbreviation	Description
Log riparian	lr_rip	Log_{10} of riparian cover
Natural log impervious road area	ln_imp	Log of impervious road area)
Natural log road density	ln_rod	Log of road density
Natural log CrAssphage	ln_cpq_056	Log of the MST marker CPQ_056
Overflow contribution	overflow_contribution	Modelled overflow volume normalised (divided by) catchment area
Road density by rainfall 90 days	road_density_rain90	Road density multiplied by 90 days rainfall period
Log mud sand ratio	lr_mud_sand	Log to the ratio of mud (<0.063 mm) and sand (>=0.063 mm to <2.0 mm)
Log sand gravel ratio	lr_sand_gravel	Log to the ratio of sand and (>=0.063 mm to <2.0 mm) and gravel (>=2.0mm)
Riparian strip	rip_strip	Measured in metres

5.6.3 Key trends and findings

Variance partitioning in the metabarcoding-derived bacterial and diatom compositional data

To reiterate, given the novelty of the study, it was pertinent to gain a fundamental understanding of how eDNA-derived communities varied biologically across space and time, as well as from technical aspects, for example sequencing lanes and between runs, with each run having a technical replicate.

The results of PERMANOVA analyses on the bacterial data yielded remarkably similar results at all taxonomic levels, phylum to OTUs (Table 5-6). OTUs are the lowest level of taxonomic resolution used here and are derived from the clustering of sequences with a 97% similarity to account for natural variation and sequencing errors. with latter being derived from the clustering of similar sequence reads. The mean squares (MS) for site were about 7-fold higher than for collection period (range: 6.7 - 7.9 fold), with site alone accounting for 36-40% of the variance (see R^2 values in Table 5-6). These findings indicated that there was about 7-fold more variation in taxonomic composition among sites than among collection periods within a site. Hence, the stream where the sample was collected accounted for the most variation in bacterial communities. The MS for collection period was about 4-fold higher than for sample (range: 4.0 - 4.5 fold), indicating that there was substantially more variation in taxonomic composition among collection periods than among samples within a collection period. The MS for sample was about 3-fold higher than for extraction (range: 3.2 - 4.2 fold). While the variance attributable to extraction was highly significant ($P < 0.001$), extraction consistently accounted for < 20% of the total variance, suggesting that replicate DNA extractions from samples is likely unnecessary, although is generally considered a critical QA component of any study. Variance attributable to lane, where the same sample was run in parallel, was negligible. Collectively, these findings emphasise that variability was due primarily to the where the samples were collected, with the technical attributes associated with eDNA metabarcoding only providing a small amount of variation. Importantly, the collection period, which reflects how communities respond across various wet-weather events, whist still important, played a smaller role than the location. Hence, the analysis suggested that the site and not the response to wet-weather events was the primary driver for variations in bacterial communities.

Table 5-6: Results of PERMANOVA analyses for bacteria at multiple taxonomic levels.

Taxon		df	SS	MS	F	R^2	MS ratio
Phylum	site	57	200	3.52	2909	0.357	6.7
	period	276	145	0.53	435	0.258	4.0
	sample	991	129	0.13	108	0.230	4.2
	extraction	2650	82	0.03	26	0.147	25.9
	lane	3975	5	0.00		0.009	
	site	57	200	3.52	2909	0.357	6.7
Order	site	57	694	12.18	2125	0.396	7.9
	period	276	426	1.54	269	0.243	4.5
	sample	991	341	0.34	60	0.195	3.4
	extraction	2650	268	0.10	18	0.153	17.7
	lane	3975	23	0.01		0.013	
OTU	site	57	1238	21.72	967	0.364	7.8
	period	276	773	2.80	125	0.227	4.3
	sample	991	649	0.65	29	0.191	2.7
	extraction	2650	652	0.25	11	0.192	10.9
	lane	3975	89	0.02		0.026	

All pseudo-F statistics were deemed to be highly significant ($P < 0.001$) on the basis of 1000 permutation samples. MS ratios were calculated based on the quotient of MS values for the variance component and the variance component below it in the hierarchy. To aid visualisation, data on the level of Phylum, Order and OTU are provided, however, the patterns were similar across all taxonomic levels

PERMANOVA analyses of the diatom data yielded much the same story as the bacteria (

Table 5-7). The MS attributable to site was about 6-fold higher than for collection period (range: 6.3 - 6.5 fold), the MS attributable to collection period was about 4-fold higher than the MS attributable to sample (range: 4.4 – 4.7 fold). Interestingly, in contrast to results for the bacteria, for diatoms, the MS for sample was about 9-fold higher than for extraction (range: 7.7 - 10.4 fold), suggesting that samples were more homogenous in taxonomic composition for diatoms than bacteria. Again, the technical components associated with the sequencing (runs, PCRs and lanes) had a negligible effect on diatom composition. These findings highlight the influence of the site and to a slightly lesser extent the effect of collection period, on diatom composition, while emphasising the relatively homogenous structure of diatom communities at a site scale when compared to bacteria.

Table 5-7: Results of PERMANOVA analyses for diatoms at multiple taxonomic levels.

Taxon		df	SS	MS	F	R^2	MS ratio
Family	site	57	1399	24.5	8617	0.393	6.4
	period	276	1058	3.8	1345	0.297	4.3
	sample	991	873	0.9	309	0.245	10.4
	extraction	2650	224	0.1	29.6	0.063	30.1
	lane	3975	11	0.003		0.003	
Species	site	57	1816	31.9	6373	0.396	6.4
	period	276	1367	5.0	991	0.298	4.7
	sample	991	1049	1.1	212	0.229	8.5
	extraction	2650	331	0.125	25.0	0.072	25.0
	lane	3975	20	0.005		0.004	
OTU	site	57	1894	33.2	5264	0.387	6.3
	period	276	1463	5.3	840	0.299	4.7
	sample	991	1122	1.1	179	0.229	7.7
	extraction	2650	388	0.147	23.2	0.079	24.5
	lane	3975	25	0.006		0.005	

All pseudo-F statistics were deemed to be highly significant ($P < 0.001$) on the basis of 1000 permutation samples. MS ratios were calculated based on the quotient of MS values for the variance component and the variance component below it in the hierarchy. To aid visualisation, data on the level of Phylum, Order and OTU are provided, however, the patterns were similar across all taxonomic levels

Human faecal-associated marker genes (HFMGs) as potential surrogates of sewage exposure

One of the key challenges of this project was determining whether a site was exposed to untreated water, and if so, whether the exposure was episodic due to a wet-weather event, additional upstream sources, or sustained due to dry leakage. To address this, two microbial HFMGs were measured, *Bacteroides* HF183 and crAssphage CPQ_056.

HF183 was detected in only 27% of samples and 81% of sites, with non-zero detections ranging by over 5 orders of magnitude of these HFMG concentrations. Given the small number of detections for HF183, presence/absence of HF183 was modelled using a binomial logit generalised linear mixed effects model. The only informative predictor of HF183 presence was rainfall, and its effects were only marginally significant ($\chi 23 = 8.42$, $P = 0.04$). Hence, this marker gene was only a weak indicator of rainfall and any potential carrier signal associated with untreated waters after a pronounced wet-weather event. The random effect of site (sd = 3.09) was very small relative to the

site-collection-period interaction (sd = 14.06), suggesting that HF183 is highly variable within a site among collection periods, with little capacity to discriminate among sites. Collectively, these findings suggest that HF183 was unsuitable for classifying the states of the sites and their responses to overflow events at the scale of this study.

CPQ_056 was detected in 82% of samples, and all but 1 of 53 sites (DEVU2). Non-zero detections ranged by over 5 orders of magnitude. The presence-absence of CPQ_056 was modelled using a binomial logit generalised linear mixed effects model. Informative predictors included: impervious area (AIC = 9.63), log[(% sand)/(%gravel)] (AIC = 6.74), and a 3rd order polynomial for rainfall over the time interval 1-10 days (AIC = 2.39). The effect of impervious area was highly significant ($\chi 21\chi 12$ = 8.68, P = 0.003). Heavily urbanised sites nearly always had concentrations of CPQ_056, and the site without CPQ_056 was among the least urbanised. The probability of the presence of CPQ_056 rapidly increased towards a plateau with increasing urbanisation (Figure 5-49).

The random effect of site (sd = 2.64) was comparable in magnitude to that of the site-collection-period interaction (sd = 4.92), suggesting that CPQ_056 offers some capacity to discriminate among sites even after accounting for impervious area. In addition to modeling presence-absence of CPQ_056, log(CPQ056) was also modeled for the subset of samples with CPQ_056 present. Informative predictors of log(CP0_Q56) included the relative proportion of sand to gravel (AIC = 7.15), sewage volume over a 1-30 day window (AIC = 5.42), and riparian area (AIC = 0.96). The effect of sewage volume was moderately significant ($\chi 21\chi 22$= 9.42, P = 0.009); however, it is worth mentioning that the effects of rainfall and sewer variables were hard to discriminate because many models had other variables within 2 AIC units of best model presented here (generally accepted as meaning the models are of equivalent performance, Burnham and Anderson 2002). The fitted random effects indicated that site to site variation (sd = 0.97) is of similar magnitude to date-to-date variation within a site (site:collection_period sd = 0.92), suggesting that CPQ_056 differs among sites even after accounting for other variables.

The CPQ_056 marker provided a potentially more robust measure of anthropogenic activities than HF183. However, discriminating between the effects of rainfall and sewage volume was limited, suggesting that it was not an accurate measure of sewage exposure during marked wet-weather events. Importantly, CPQ_056 was present in most samples, including those from upstream sites, with these sites initially classified as Baseline or Urban Water (Table 5-4). This emphasises the limitations associated with the initial classification of sites (Table 5-4), and the need for the subsequent analysis to be transitioned from a factorial approach to more correlative-based approach, utilising a more comprehensive list of environmental variables (Table 5-5). Given its potential utility as an indicator of catchment condition, CPQ_056 was included as a dynamic variable in subsequent analyses (Table 5-5).

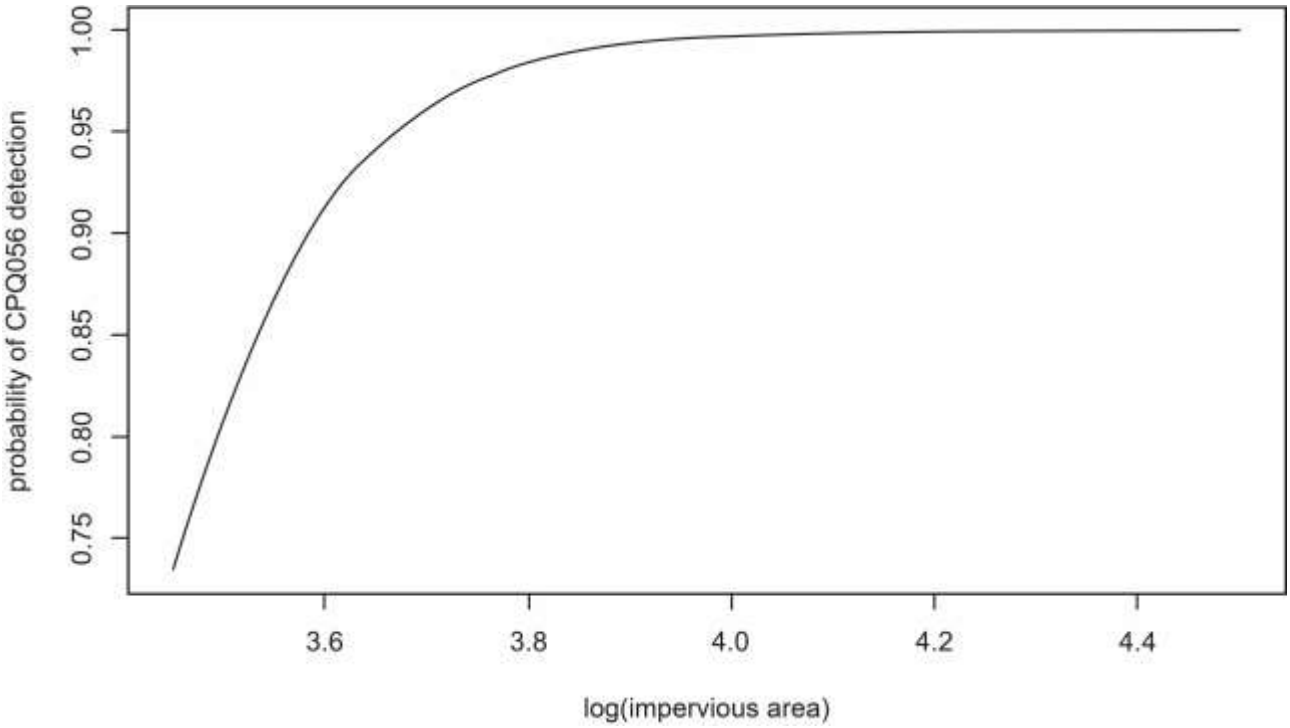


Figure 5-49: Predicted relationship between CPQ_056 detection probability and the amount of impervious area

Spatial and temporal patterns in eDNA metabarcoded bacterial and diatom communities

The bacterial communities from across the study are presented in the PCA plot (Figure 5-50). PERMANOVA analysis found that composition differed between catchments ($F=25.2$, $P<0.001$), sites ($F=18.8$, $P<0.001$), position in the stream, that is, upstream or downstream ($F=13.2$, $p<0.001$) and the collection period ($F=6.58$, $P<0.001$). Qualitatively, the figure shows a separation between those communities sampled from the Georges and Parramatta River catchments, with this reflecting their differences in subsurface geology. This is supported by Figure 5-51, where the ordinations are presented for both catchments with sandstone and shale geologies. Given the marked effect of subsurface geology, subsequent statistical analyses were performed separately for both sandstone and shale environments. Figure 5-52 provides separate ordination plots for each catchment. In some catchments, for example the Georges River, Middle Harbour and Hawkesbury River, clear groupings of some upstream and downstream sites were evident, indicating a downstream effect, possibly due to wet-weather inputs, however, other attributes to the streams may also be contributing to these differences, for example flow, riparian cover and road density.

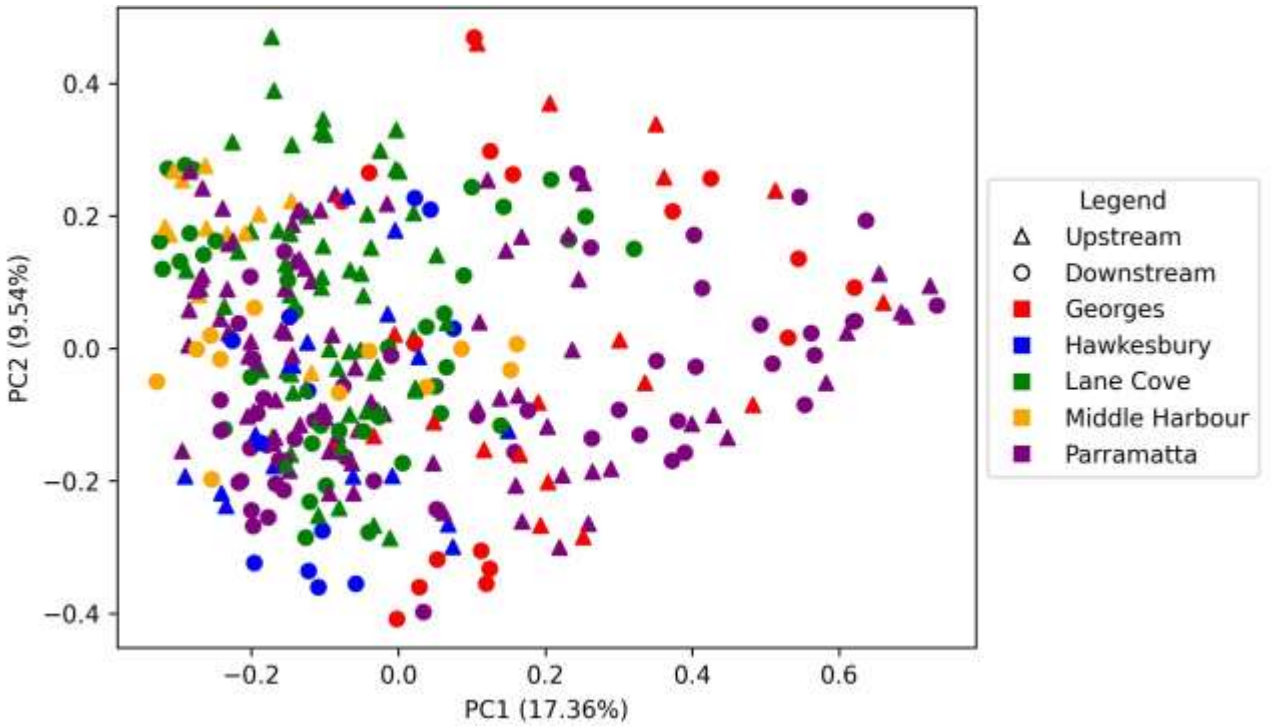


Figure 5-50: A PCA representing communities from all bacterial samples.
The closer the points are to each other, the more similar the communities are in composition. Triangles represent upstream sites and circles downstream sites

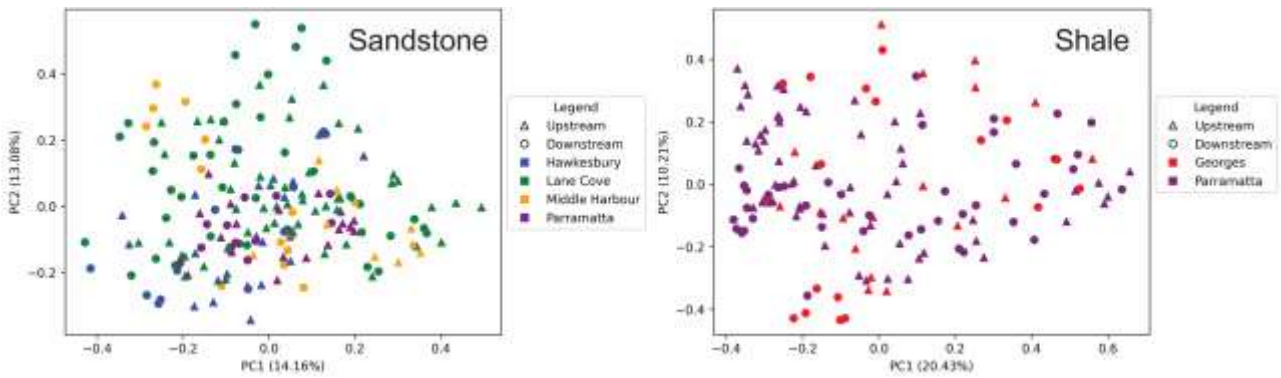


Figure 5-51: PCAs representing bacterial communities combined at the collection point (time) from catchments with sandstone and shale geologies.
Triangles represent upstream sites and circles downstream sites

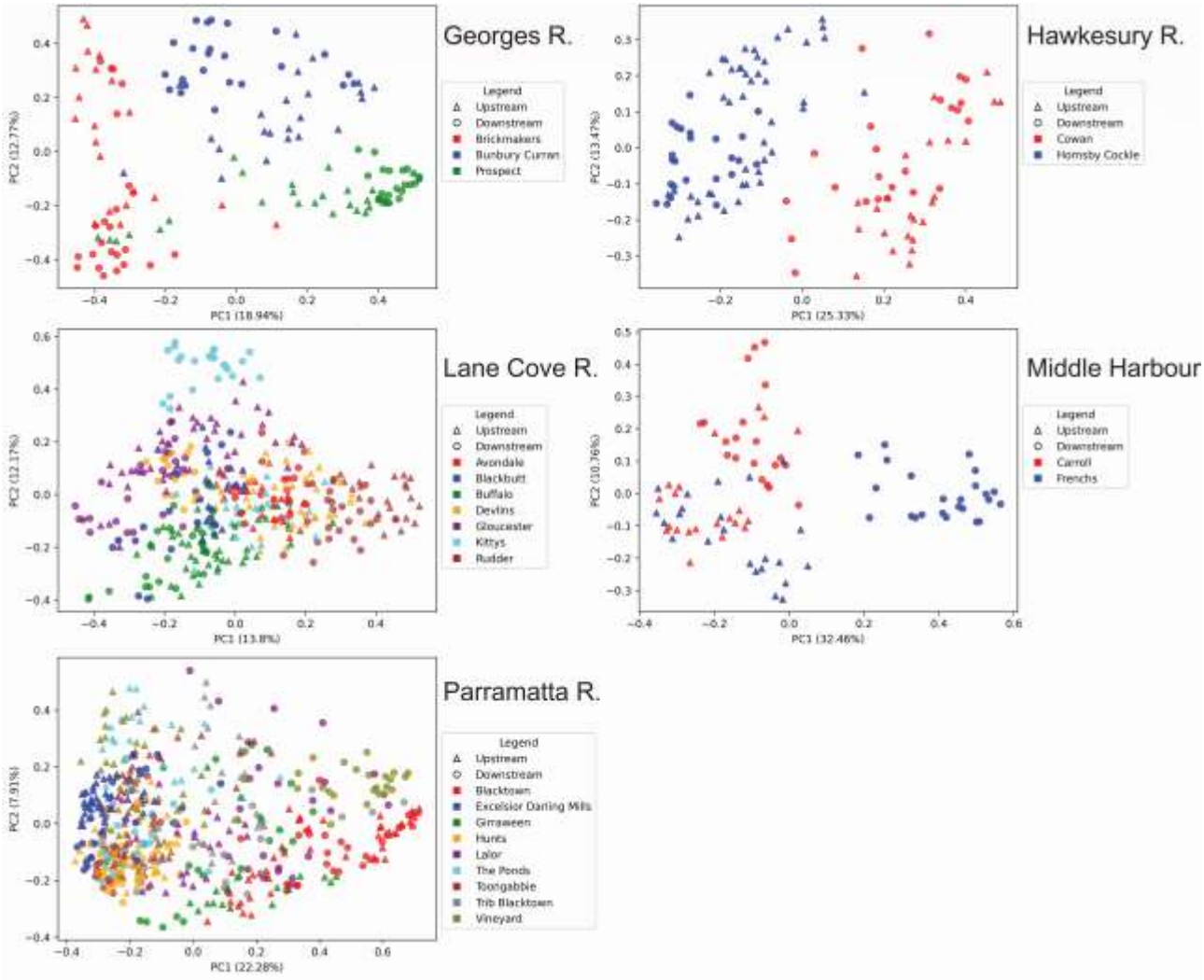


Figure 5-52: PCAs representing bacterial communities combined at the collection point (time) from each catchment.

Triangles represent upstream sites and circles downstream sites

Ordination plots for diatom communities (Figure 5-53), showed similar patterns to those observed with the bacterial community data. Diatom composition was shown to significantly differ between catchments (F=12.9, P<0.001), sites (F-3.66, P<0.001), collection period (F=8.41, P<0.001), and to a lesser extent, position in the river (F=1.61, P=0.043). This included the clear separation of diatom communities from sandstone and shale catchments (Figure 5-54). Differences in upstream and downstream diatom communities were most evident in the Middle Harbour sites for Carroll and Frenchs, as well as Prospect in Georges River (Figure 5-55). As in the case of the bacterial data, the drivers for difference in upstream and downstream sites cannot be elucidated without the inclusion of environmental data.

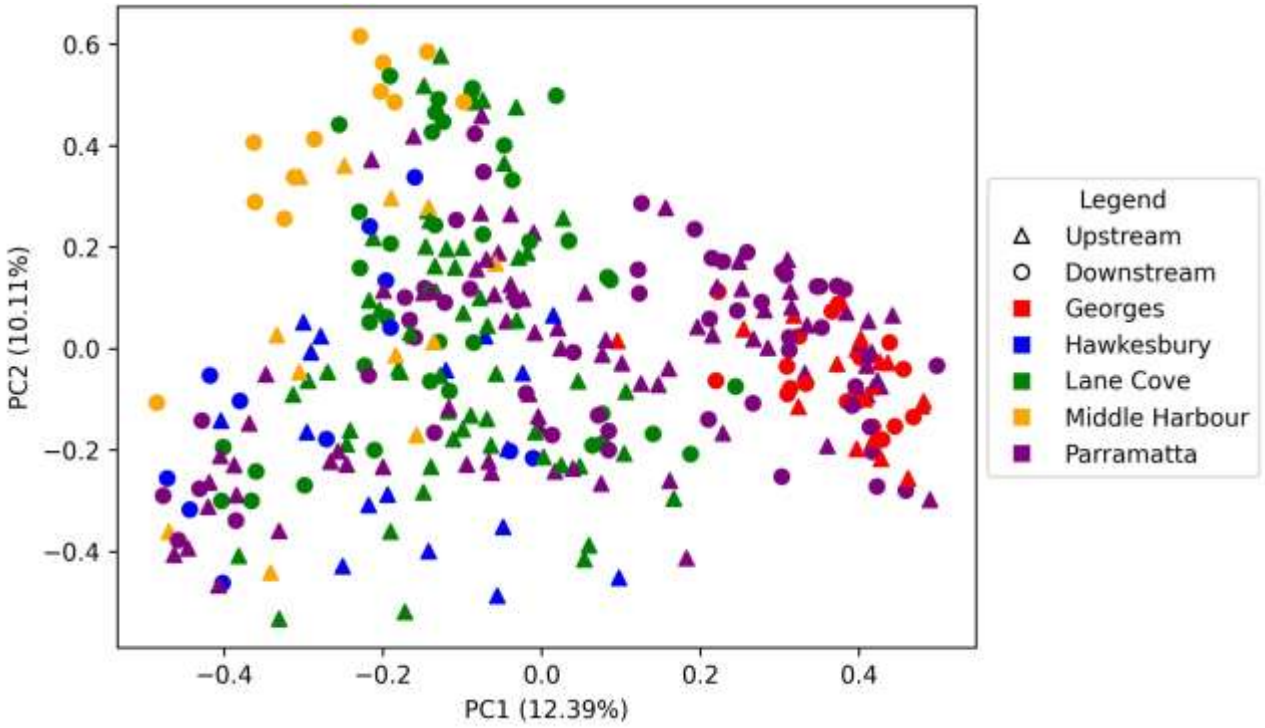


Figure 5-53: A PCA representing communities from all diatom samples combined at the collection point (time).

The closer the points are to each other, the more similar the communities are in composition. Triangles represent upstream sites and circles downstream sites

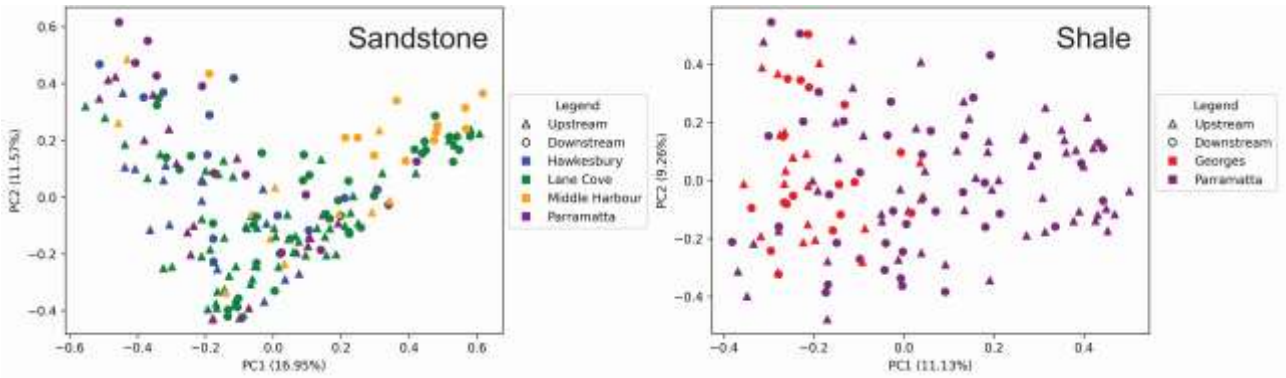


Figure 5-54: PCAs representing diatom communities from catchments with sandstone and shale geologies.

Triangles represent upstream sites and circles downstream sites

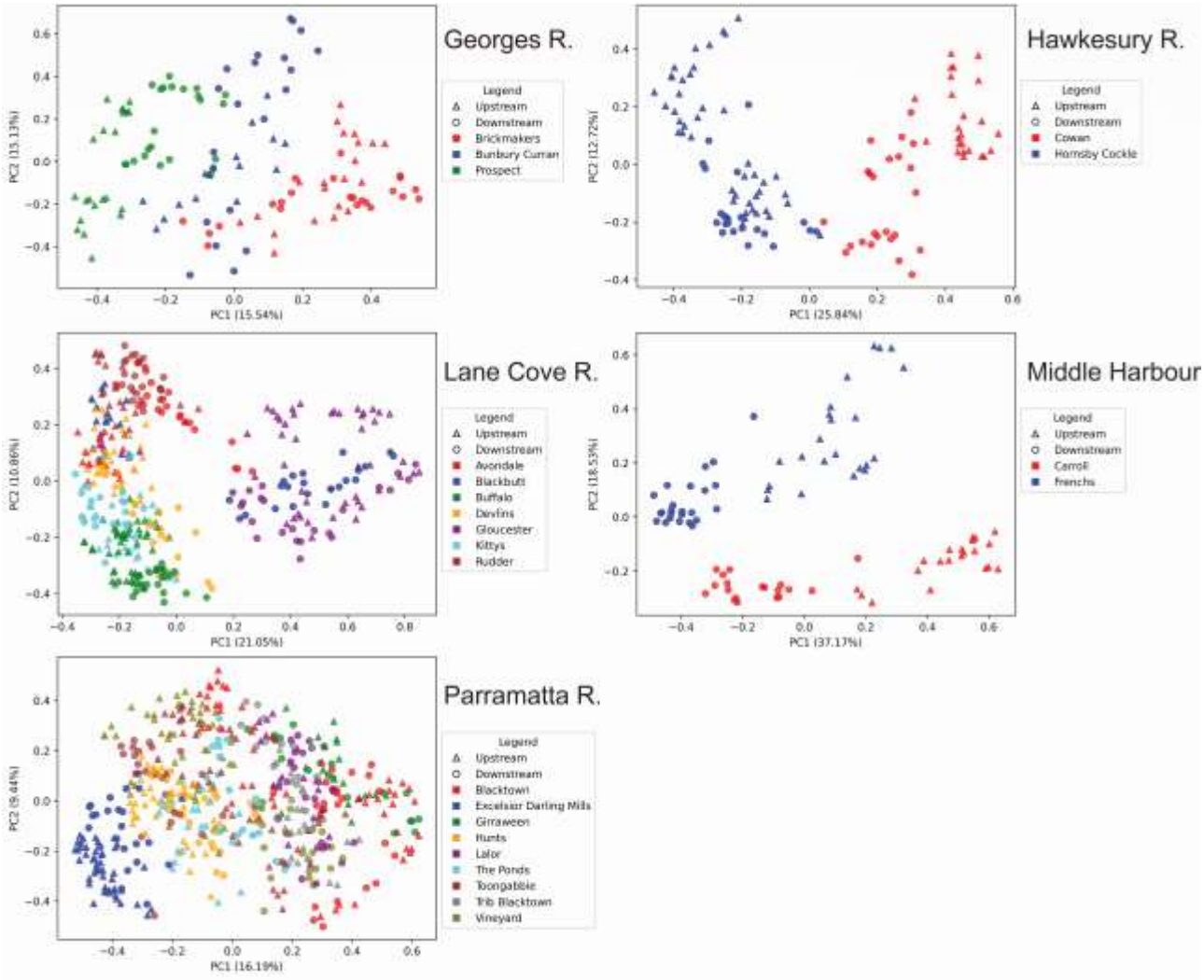


Figure 5-55: PCAs representing diatom communities from each catchment combined at the collection point (time).

Triangles represent upstream sites and circles downstream sites

Relationships between benthic communities (bacteria and diatoms) and environmental variables

The ordination plot derived from the CCA of the 16S rDNA data (bacteria) is presented in Figure 5-56. The plot clearly illustrates the separation of bacterial communities from sandstone and shale catchments, with the latter being strongly correlated with the ratio of mud to sand (lr_mud_sand). A number of downstream (circle) shale sites were also strongly correlated with concentrations of the microbial source tracker CPQ_056 and overflow contribution, which is the modelled overflow volume divided by the catchment area.

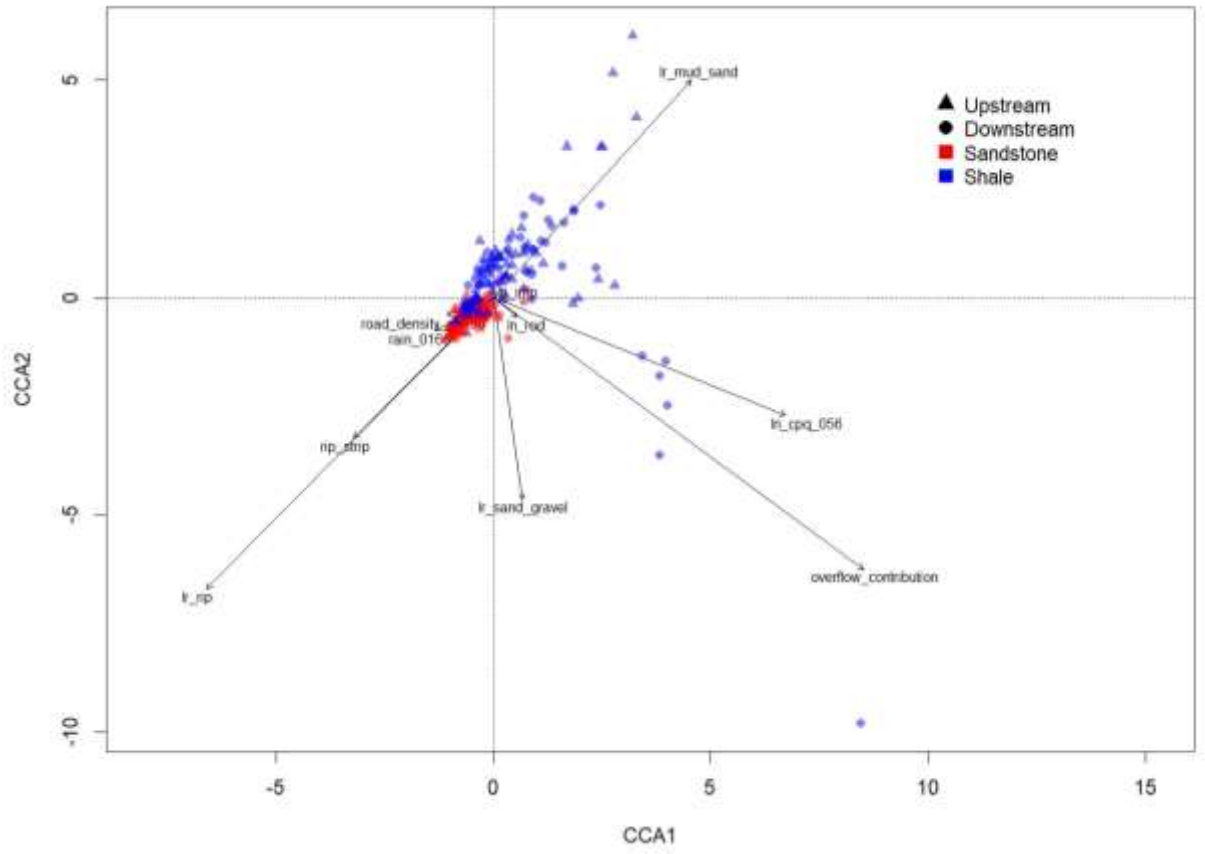


Figure 5-56: Canonical correspondence analysis (CCA) ordination biplots derived from the environmental data and bacterial composition.
The longer the vector the stronger its correlation with the 16S rDNA bacteria community data

Variables displayed in vectors are further described in Table 5-5.

The finding from the GGLVM for bacteria found both latent variables (LVs), which control for substrate quality and rainfall, were significantly correlated with all landscape variables at all levels of taxonomic resolution ($P < 0.05$). Spearman rank correlations ranged from 0.10 ($P = 0.02$) to 0.54 ($P < 0.001$). The highest correlation was found between road density and LV1 at the genus level (Figure 5-57). This highlights the over-riding influence of road density on bacterial communities, reducing the capacity to determine specifically how these communities respond to rainfall. However, the GLLVMs did identify the taxa most sensitive to the environmental gradients represented by the two latent variables, as depicted in Figure 5-58, potentially providing indicator taxa for these variables.

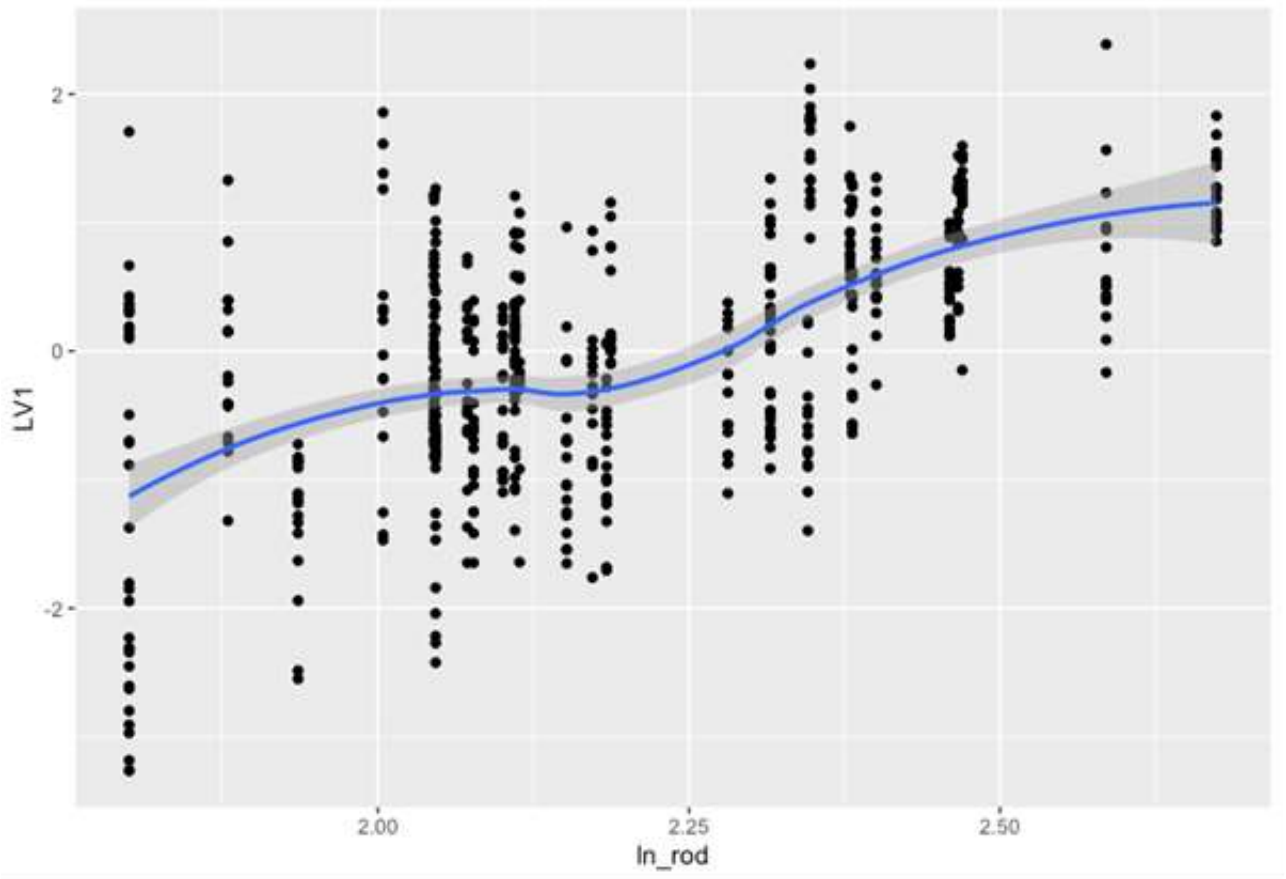


Figure 5-57: The relationship between the first latent variable (LV) for bacterial genera, derived from a GLLVM, and the logarithm of road density.

This LV represents a primary axis of covariation among bacteria genera after controlling for substrate effects using logarithm of the mud-gravel ratio and rainfall using a 3rd order polynomial

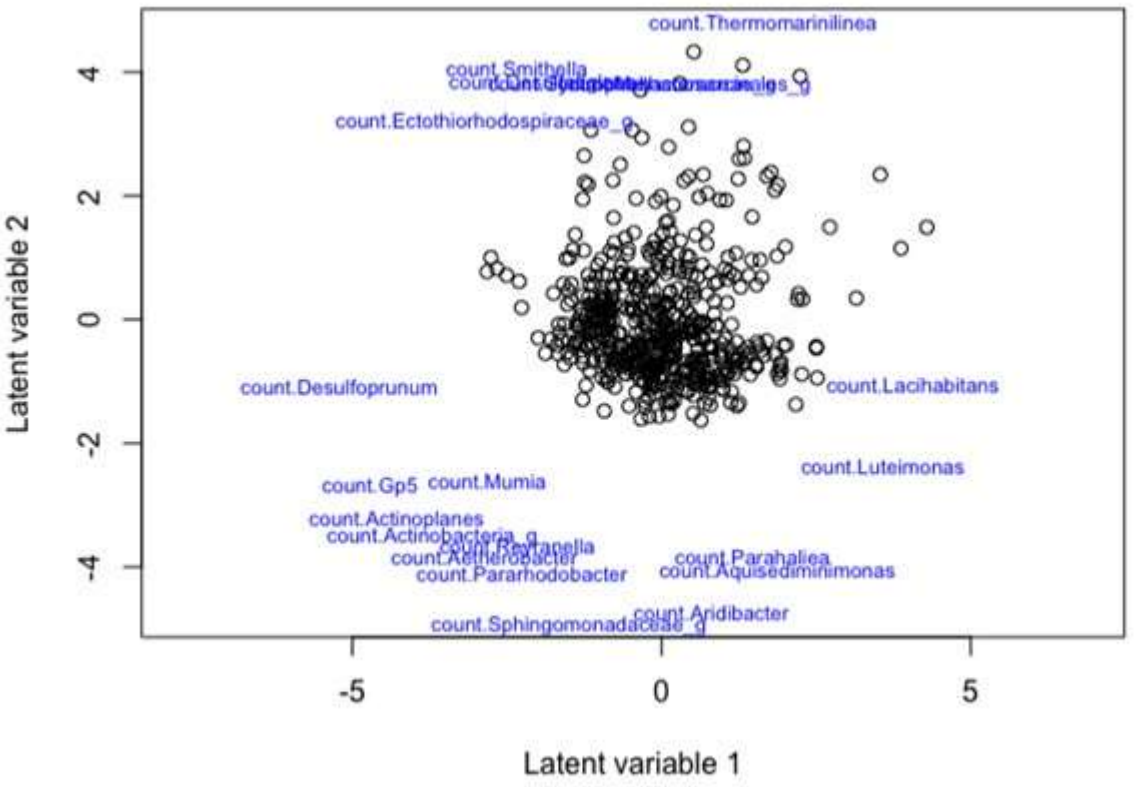


Figure 5-58: The bacterial taxa found to be the most sensitive to the environmental gradients represented by the two latent variables

The CCA for the diatoms is presented in Figure 5-59. To a marginally lesser extent than the bacteria data, diatom communities were still separated by the subsurface geology of the catchment. The environmental vectors explaining the variation in the diatom community data were less pronounced, and where more evident in the sandstone communities, which were predominately correlated with the ratio of sand to gravel (lr_sand_gravel), and riparian attributes- riparian length (rip_strip) and cover (lr_rip). The shale communities were most strongly correlated with the ratio of mud to sand (lr_mud_sand). For diatoms, Spearman rank correlations of the two LVs with the landscape variables ranged from 0.02 (n.s.) to 0.70 ($P < 0.001$). The highest correlation was between LV1 and the amount of impervious area at the genus level (Figure 5-60). Potential indicators associated with this relationship are shown in Figure 5-61.

As in the case of bacteria, the GLVV analysis again highlights the profound influence of static environment variables associated with the catchment (for example riparian vegetation and grain size) rather than wet-weather flows in shaping the sampled diatom communities.

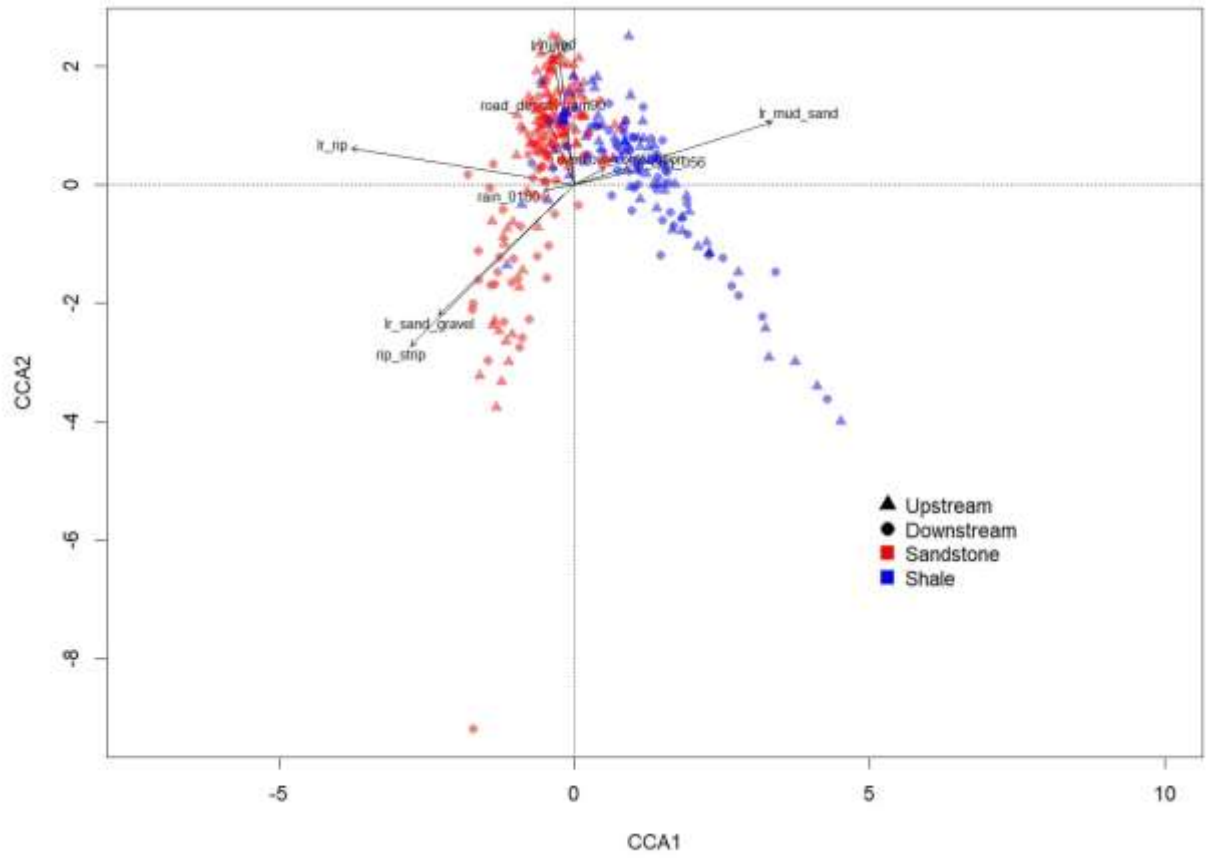


Figure 5-59: Canonical correspondence analysis (CCA) ordination biplots derived from the environmental data and diatom composition

The longer the vector the stronger its correlation with the 18S rDNA diatom community data

Variables displayed in vectors are further described in Table 5-5.

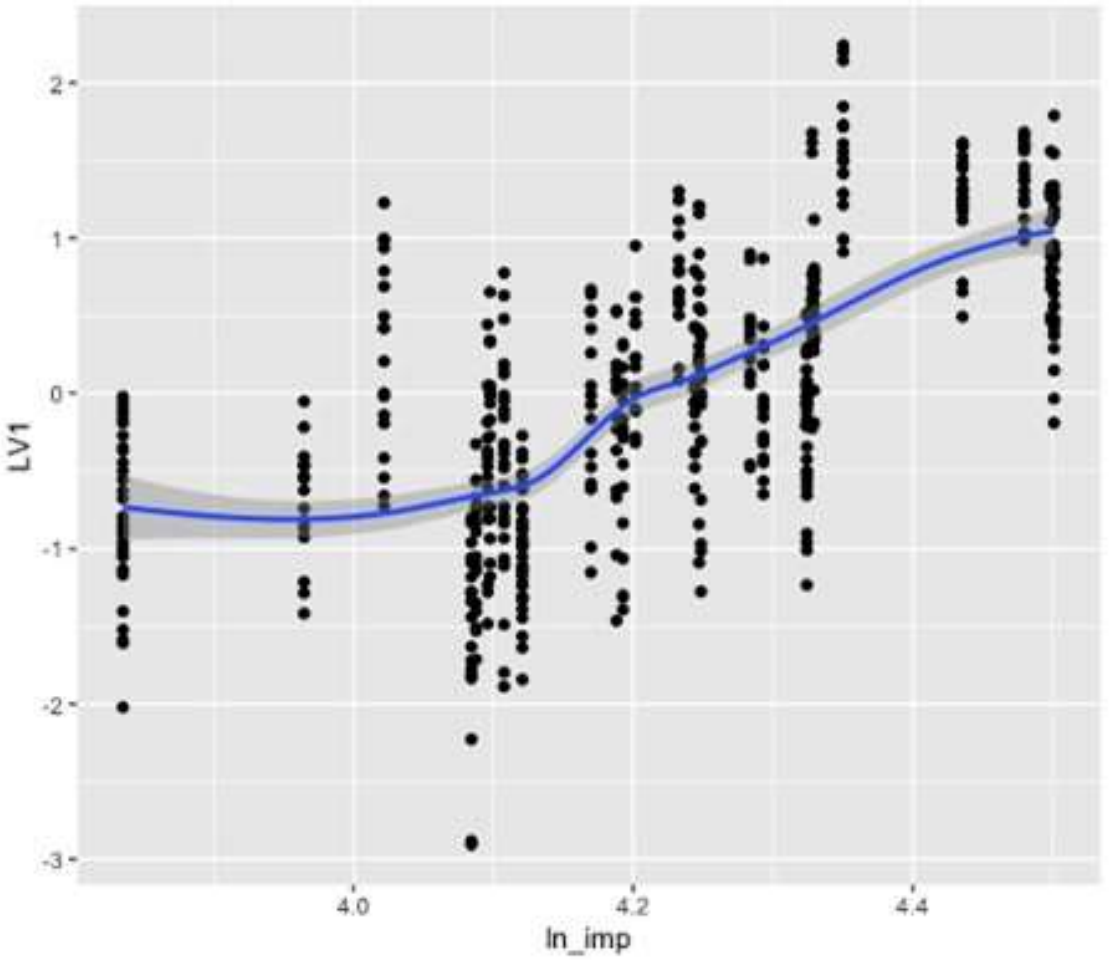


Figure 5-60: The relationship between the first latent variable (LV) for diatom genera, derived from a GLLVM, and the logarithm of impervious area.

This LV represents a primary axis of covariation among diatom genera after controlling for substrate effects using logarithm of the mud-gravel ratio and rainfall using a 3rd order polynomial. The smoothed curve in the figure is unrelated to the GLLVM and is included only to help visualise the trend

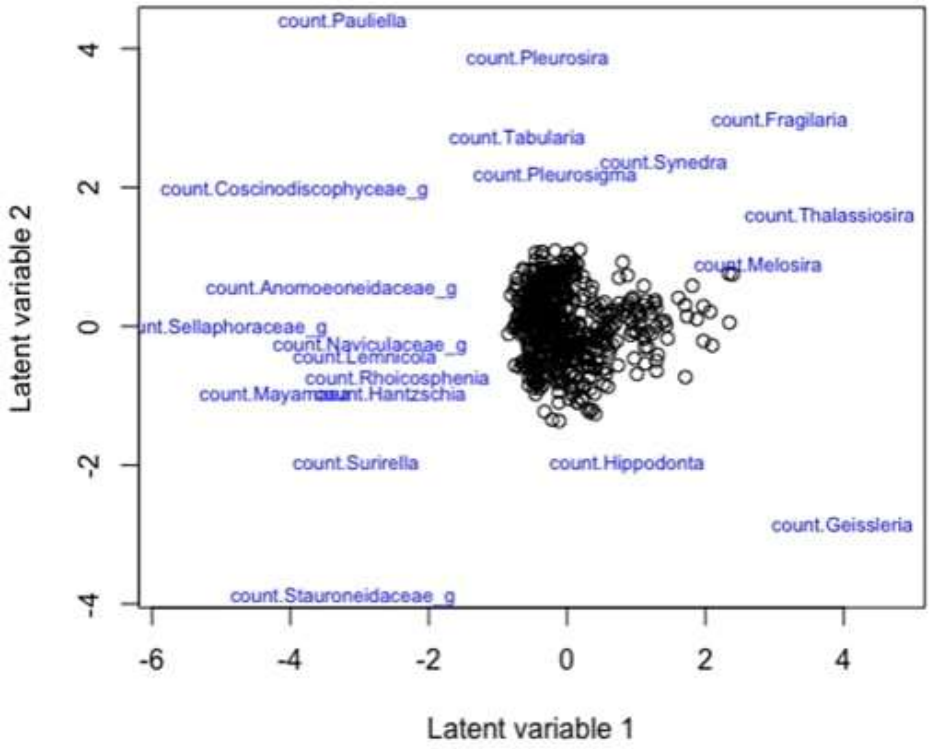


Figure 5-61: The diatom taxa found to be the most sensitive to the environmental gradients represented by the two latent variables

Neural network

The performance of Neural Network (NN) models is visually shown using prediction plots and loss curves. The prediction plot visually compares the predictions of the model (for example, bacterial composition data) against the target values in the dataset, in this case it is SIGNAL-SG. The x-axis represents the true SIGNAL-SG values of each sample derived from morphometric data of the macrobenthos, while the y-axis represents the predicted values generated by the NN model. Data points in the plot corresponds to a specific observation in the dataset. Ideally, the points should be close to the diagonal line which indicates that the model's predictions closely match the true values. Deviations from the diagonal line suggests discrepancies leading to an under-performing model. Loss curves are also shown which provide insights into the performance of the NN model during the training and validation phases. The loss function, mean squared error in this case, is plotted over successive epochs of the training process. The x-axis represents the number of epochs, while the y-axis represents the value of the loss function where a low value means a better performance.

Initially, the NN model was investigated using bacterial 16S rDNA data (bacteria) with no specific groupings of samples when the splitting between the training, test, and validation sets occurred, and SIGNAL-SG as the predicting label. The best performing NN model had a mean-squared error of 0.06 and 0.07 for the validation and test sets, respectively (Figure 5-62). Although the performance metrics showed some promise in the model being able to accurately predict SIGNAL-

SG, this approach was naïve as the model was overfitted due to the splitting of the data without grouping them together. In essence, the approach resampled itself using some of the samples in both the modelling and testing phases, providing an over optimistic view of the relationship between the true (SIGNAL-SG) and predicted responses (bacteria data).

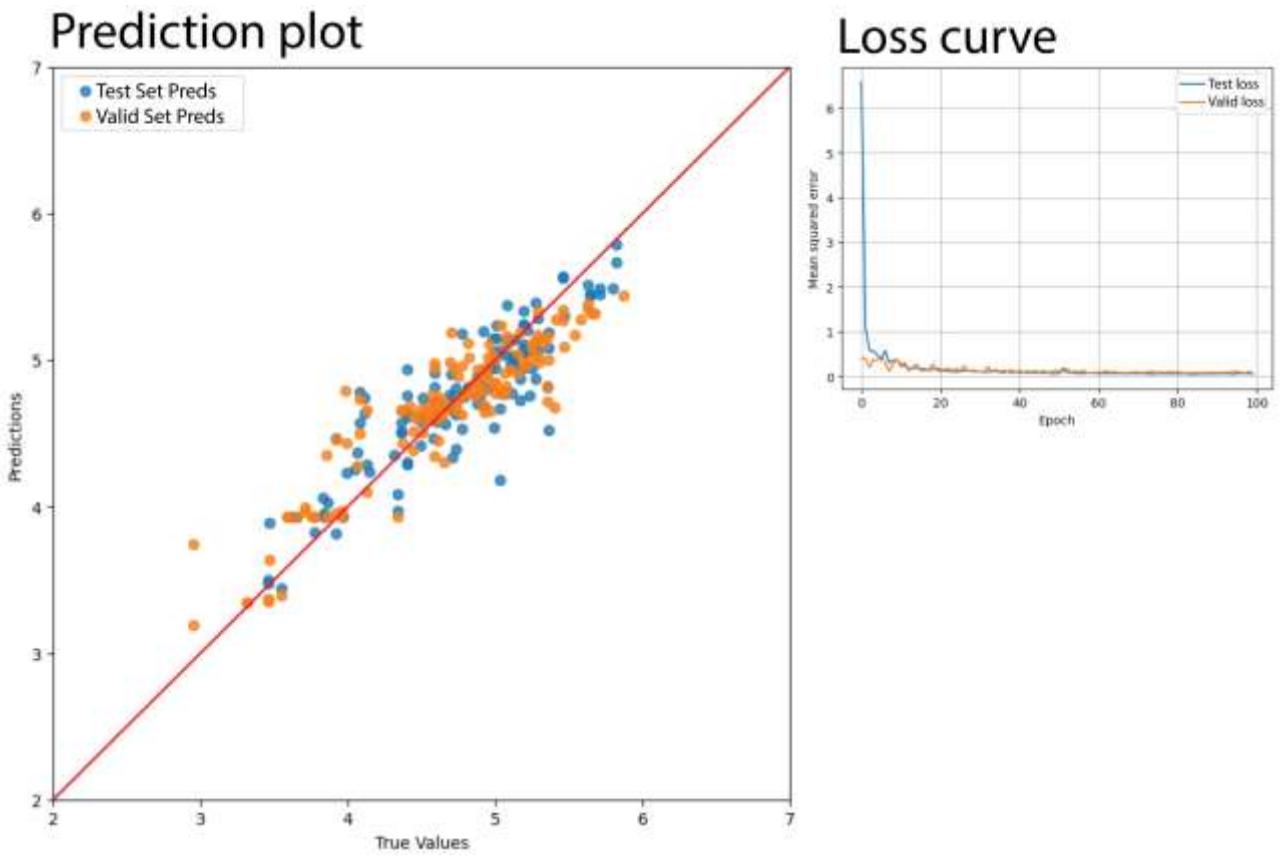


Figure 5-62: The bacterial 16S rDNA NN model (no grouping) prediction plot and loss curves using SIGNAL-SG as the label.
The blue-coloured data indicate data from the test set while the orange colour indicate data from the validation set

Grouping together samples which were collected from the same location is important to prevent overfitting of the model for several reasons, namely, preservation of spatial correlation and enhanced generalisation. By grouping the samples from the same location together, the model can better capture the spatial correlation or patterns to that location. Without the grouping, the model may instead be predicting the values based on a location which will lead to poor performance on unseen data. Taking this into consideration, the NN model was then adjusted to group together sites when the splitting of the data occurs, for example, when a sample is picked to be placed in the training set, all the samples from the same site are also placed into the training set. SIGNAL-SG was again used as the predicting label.

The best performing NN model only had a mean squared error of 0.33 (Figure 5-62) and 0.23 (Figure 5-63) for the validation and test sets, respectively. Hence, there was a notable decline between the NN model with no grouping and the one with grouping. This contrast serves to

highlight the inherent risks of employing a NN model without correct grouping of samples. The model with no grouping of samples exhibits signs of overfitting which suggests that it has memorised specific individual samples and is predicting based on location rather than grasping the broader patterns that may be hidden within the dataset. Furthermore, it is important to note that neural network models perform best when trained on vast quantities of data. Although in this study there was a large of amount of data, samples were collected within a relatively small temporal period (approx. 1 yr), restricting the capacity of NN to extract generalised patterns which are required to make accurate predictions. This emphasised the unsuitability of NN's for this study. Hence, additional machine learning modelling was performed using Random Forest, which is more applicable to correlative patterns and not restricted to a factorial base founded on pre-assigned states.

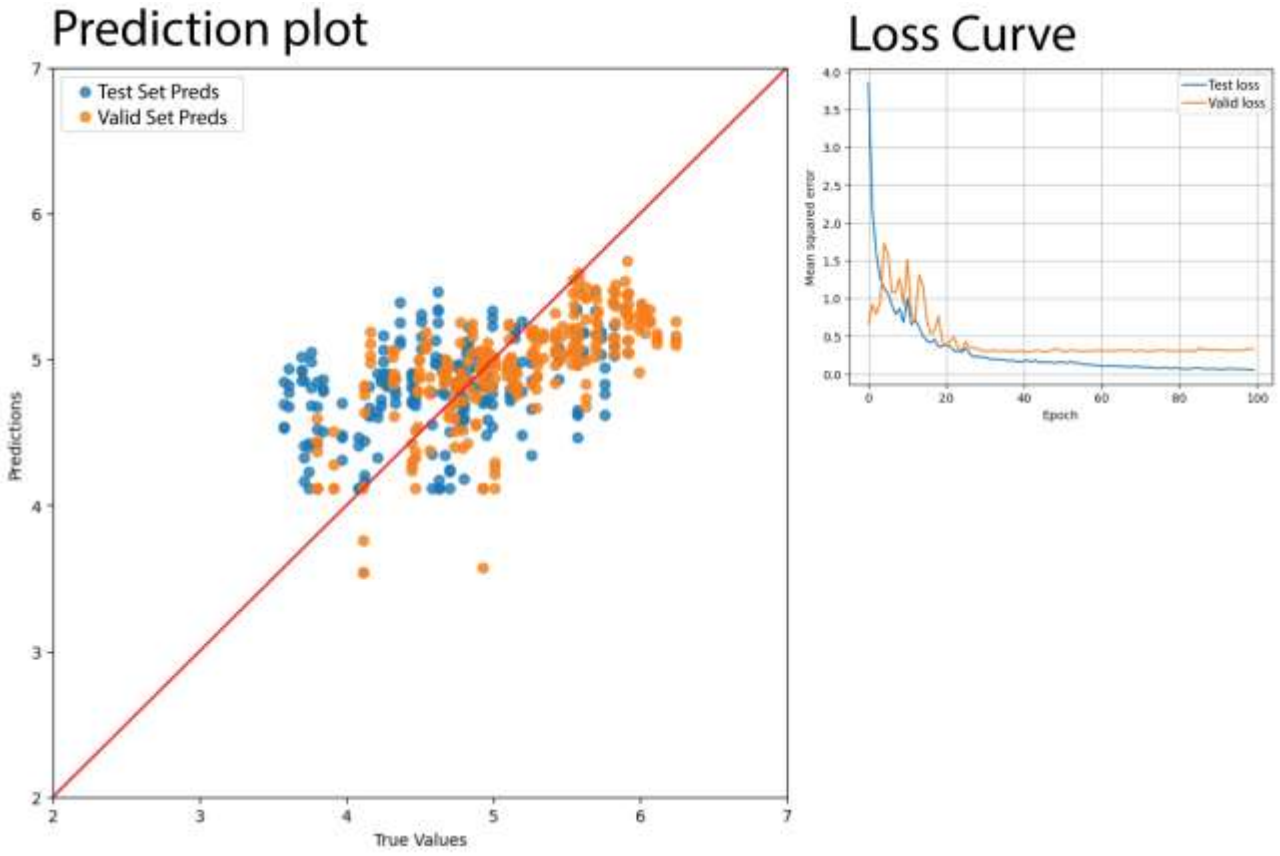


Figure 5-63: The bacterial 16S rDNA NN model (with site grouping) prediction plot and loss curves using SIGNAL-SG as the label.

The blue-coloured data indicates data from the test set while the orange colour indicates data from the validation set

Random forest (RF)

In an analogous manner as described above for the NN, biotic data (bacteria) was incorporated into random forest (RF) approach and compared to SIGNAL-SG scores. However, in this case, catchments were split into sandstone and shale, given the previously stated over-riding influence of subsurface geology on the metabarcoded communities. The RF performed best in shale catchments, with R^2 of 0.38 and mean squared error of 0.19 for both the phylogenetic resolutions (Table 5-8). As illustrated in Figure 5-64, predictions were less accurate for bacterial communities sampled from sandstone catchments with higher SIGNAL-SG scores, that is, higher scores, deviating from the line of prediction. Conversely, the predictions were poorer for shale sites with lower SIGNAL-SG scores (Figure 5-64).

Table 5-8: Performance metrics for the random forest analysis of the bacterial 16S rDNA (bacteria) for the sandstone and shale samples at species and genus levels using SIGNAL-SG as the label

Subsurface geology	Phylogenetic resolution	Mean squared error	Root mean squared error	R^2	Mean absolute error	Explained variance score
Sandstone	Genus	0.16	0.40	0.27	0.31	0.28
	Species	0.15	0.39	0.32	0.29	0.32
Shale	Genus	0.19	0.43	0.38	0.35	0.38
	Species	0.19	0.43	0.38	0.34	0.38

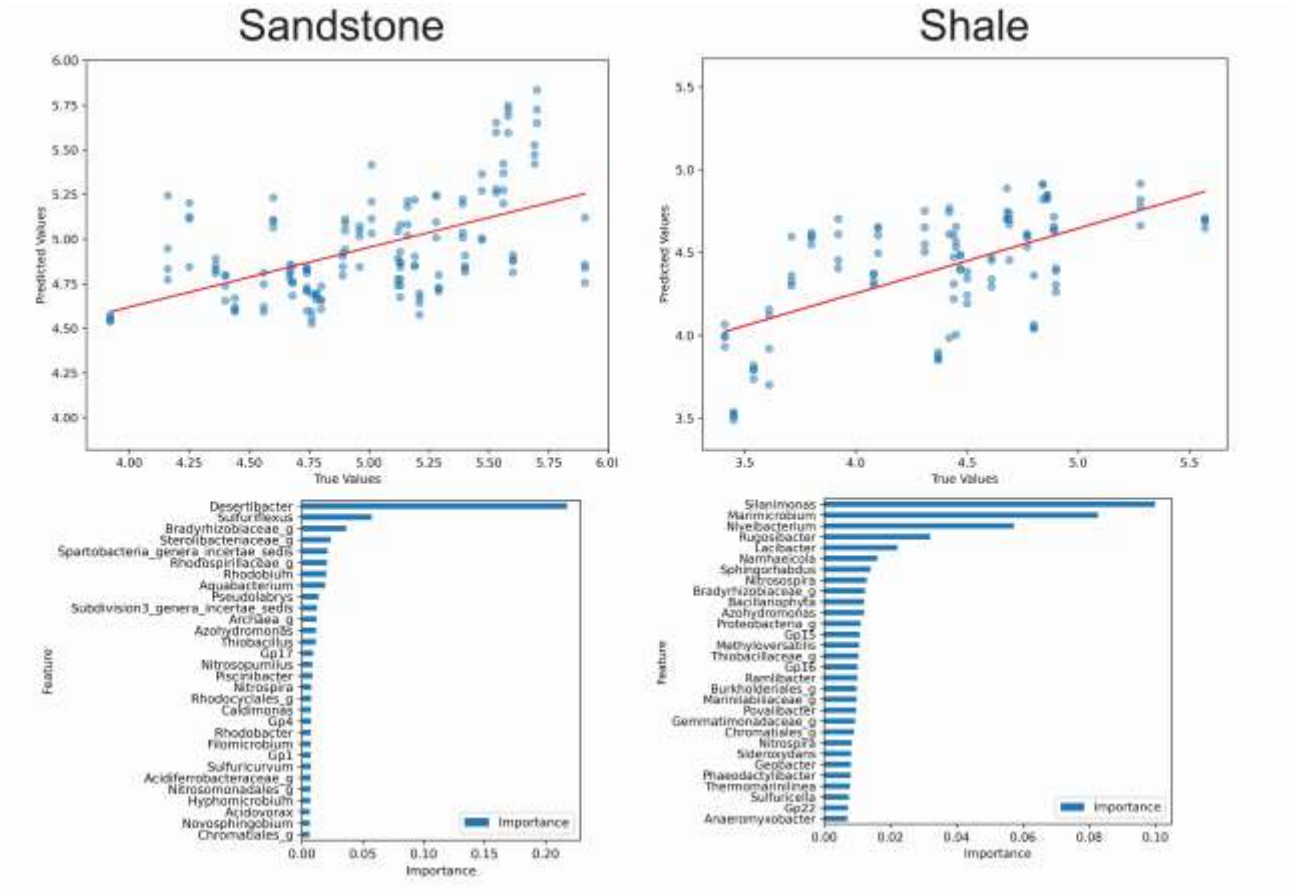


Figure 5-64: RF prediction plots and feature importance of bacterial 16S rDNA (y-axis) at the genus using SIGNAL-SG (x-axis) as the label.

Figures on the left are for sandstone catchments, with shale catchments presented on the right. Higher importance of taxa indicates their importance in predicting the model. Note that only the top 30 features are shown

In addition to using bacterial composition to predict SIGNAL-SG scores, inferred functions (for example, nitrate reduction, nitrogen respiration, nitrate respiration, and nitrification) derived from the bacterial data using the software Faprotaxa were also performed on both sandstone and shale samples. Both geologies had negative R^2 values at -0.16 and -0.08, respectively, with respective mean squared errors of 0.37 and 0.25. However, in both cases, the inferred functional data only explained 7 % of the variation in the SIGNAL-SG data. The features which were most important varied between the sandstone and shale samples (Figure 5-65). For the sandstone samples, the most important features included many nitrogen related metabolic pathways such as nitrate reduction, nitrogen respiration, nitrate respiration, and nitrification. In addition, human-associated and human pathogens were also of high importance when it comes to SIGNAL-SG predictions. This suggests that some nitrogen cycling functions and pathways, as well as pathogens, are negatively correlated with SIGNAL-SG scores, emphasising a link between sandstone site condition and function. In the communities from the shale samples, the highest ranked features were associated with photosynthesis, as features associated with chloroplasts, photoautotrophy, anoxygenic photoautotrophy S oxidising, and anoxygenic photoautotrophy were the most important

features. The human pathogens feature was also ranked high in shale samples. This further suggests there is some relationship in shale site condition (based on SIGNAL-SG) and functional processes, although it is not possible to accurately determine site condition purely by its inferred functionality. Furthermore, this relationship is less pronounced in communities from shale catchments than those from sandstone catchments.

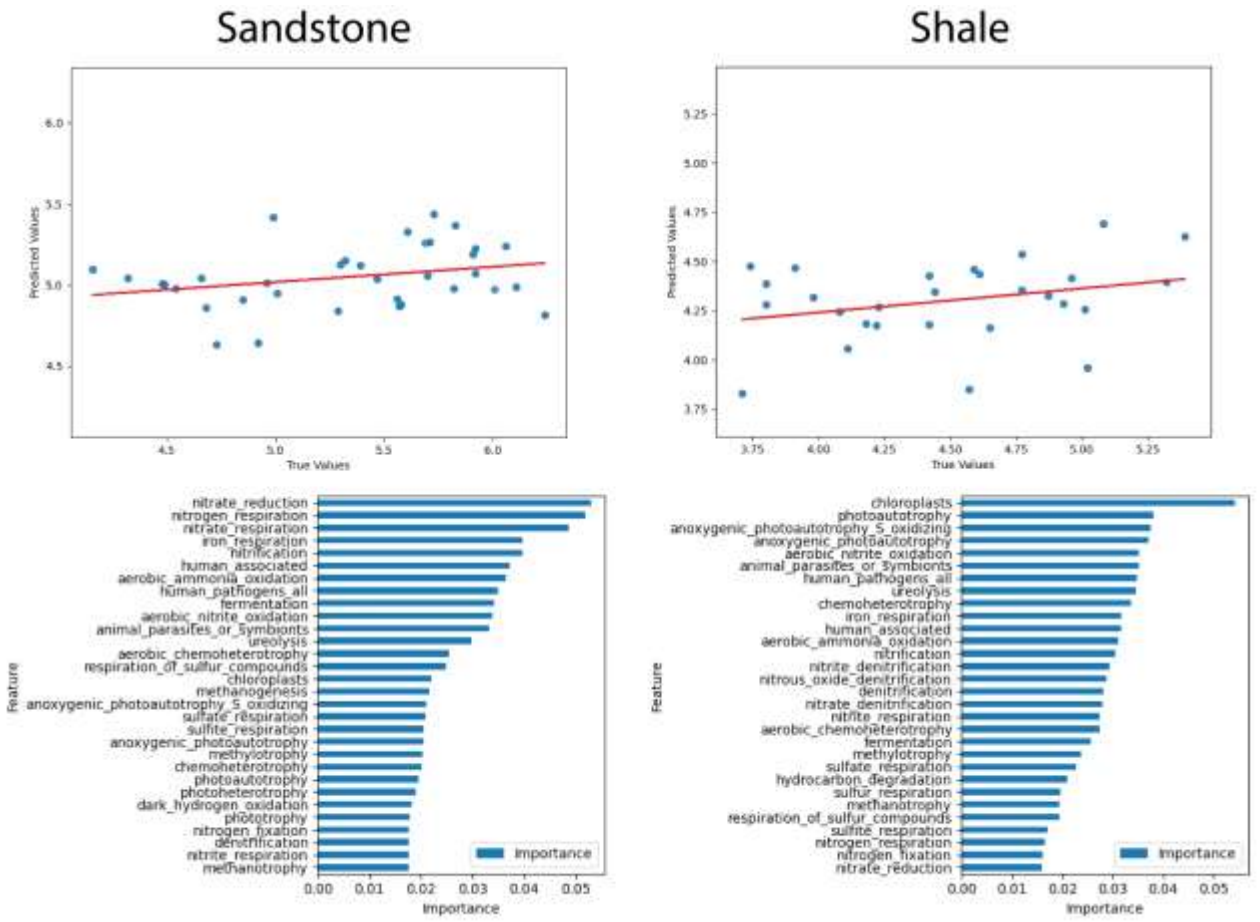


Figure 5-65: The RF prediction plots and feature importance for sandstone and shale derived from inferred functional analysis of the bacteria data (Faprotaxa) using SIGNAL-SG as the label.
Note that only the top 30 features are shown here out of a total of 59 features

Using an analogous approach to that performed with the bacterial data (16S rDNA), RF was also performed using the diatom data derived from sequencing a region of the 18S rDNA gene. The sandstone model had the highest R^2 value (0.46 at the species level), providing a better prediction than in the shale sampled communities (0.27 at the genus level) (Table 5-9). As illustrated in Figure 5-66 (right figure), predicted values for the diatom communities against the SIGNAL-SG scores were particularly inaccurate in shale sites of poor condition (low SIGNAL-SG scores), with RF often modelling a lower score than SIGNAL-SG.

Table 5-9: Performance metrics for the random forest analysis of diatom communities (using 18S rDNA) for the sandstone and shale samples at species and genus levels using SIGNAL-SG as the label

Subsurface geology	Phylogenetic resolution	Mean squared error	Root mean squared error	R^2	Mean absolute error	Explained variance score
Sandstone	Genus	0.22	0.46	0.34	0.36	0.34
	Species	0.18	0.42	0.46	0.34	0.46
Shale	Genus	0.33	0.58	0.20	0.4	0.22
	Species	0.31	0.55	0.27	0.46	0.27

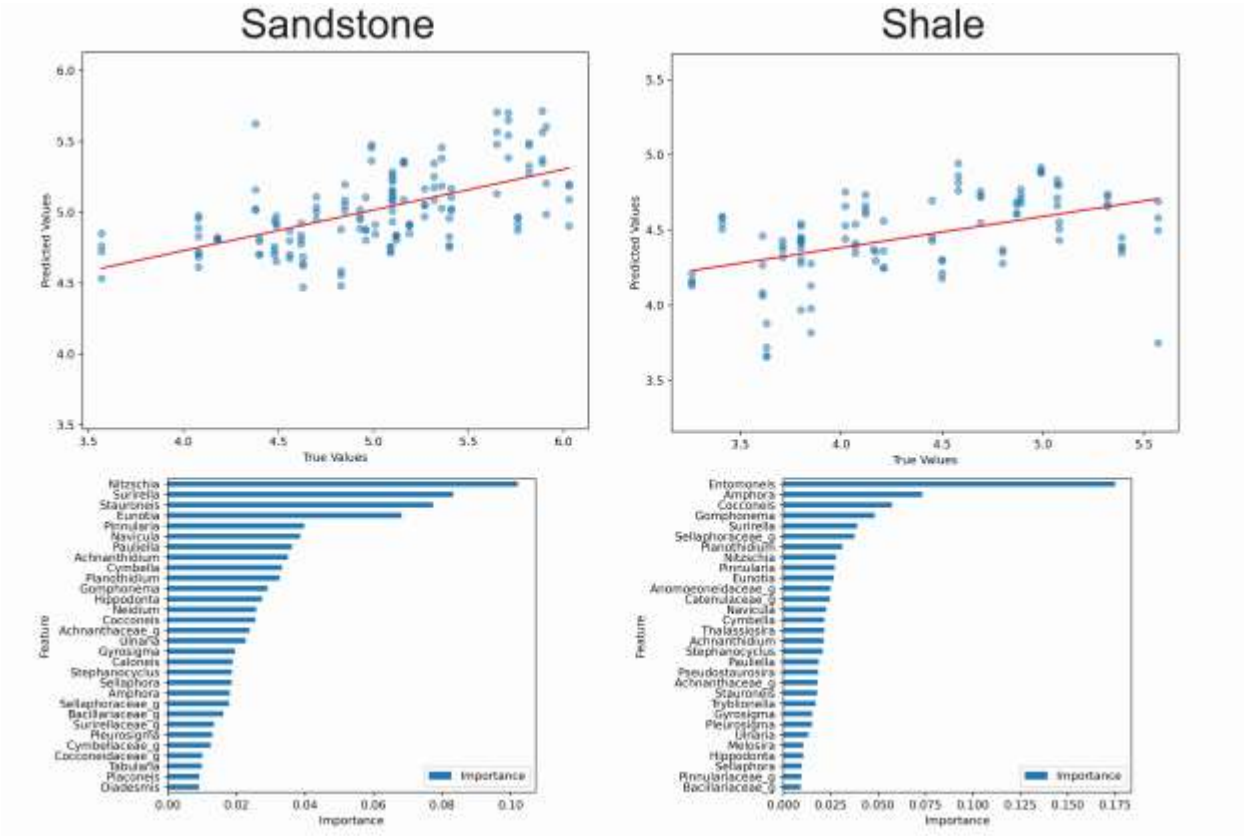


Figure 5-66: RF prediction plots and feature importance of diatoms (y-axis) using SIGNAL-SG (x-axis) as the label.

Figures on the left are for sandstone catchments, with shale catchments presented on the right. Higher importance of taxa indicates their importance in predicting the model. Note that only the top 30 features are shown here.

The RF was also run using the diatom EPID index. As previously stated, out of the 13 diatom indices tested, EPID had the highest R^2 value whilst having the lowest mean squared error, with 69% of the sequenced taxa able to be assigned to the relevant taxonomy used for this index. For both the sandstone and shale samples, the RF using the EPID had an R^2 of 0.28 and 0.22, and a mean squared error of 0.15 and 0.12, respectively. Hence, the predictability of the molecular EPID scores was low, and much less than that produced when comparing diatom communities against SIGNAL-SG. The sandstone sites were mostly influenced by the density of roads (ln_rod) and impervious road area (ln_imp), based on returned R^2 of 0.161 and 0.151, respectively (Figure 5-67). For the shale sites, the most important features were the concentrations of the MST CPQ056 (ln_cpq_056) and the ratio of mud to sand (lr_mud_s) (Figure 5-67). Whilst the relationships were not strong, these findings again emphasise the influence of roads, grain size and other static variables.

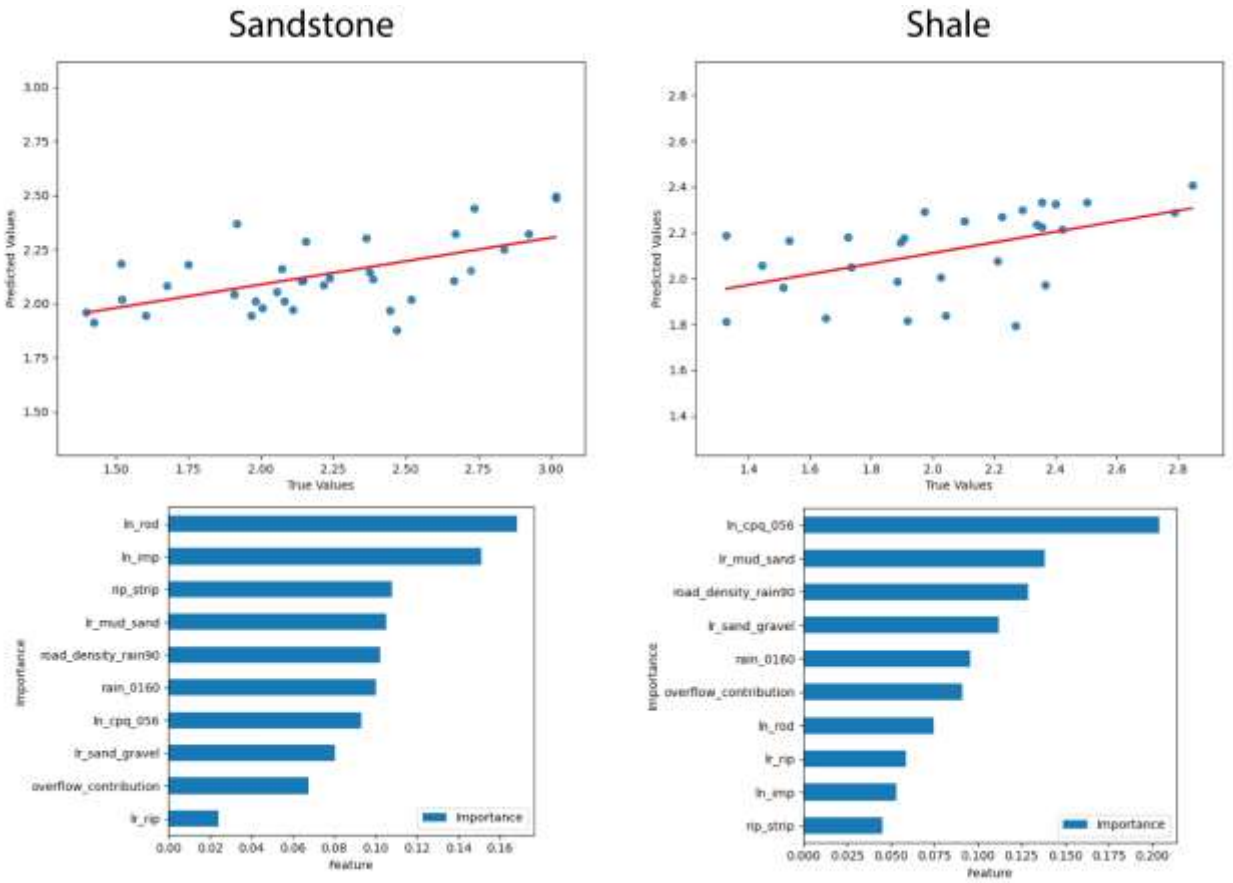


Figure 5-67: EPID index prediction plots and feature importance for sandstone and shale using environmental variables

Additional examination of the EPID showed that this diatom index was poorly correlated with SIGNAL-SG (R^2 =-0.15) (Figure 5-68). While a negative correlation was expected, as sites in better ecological condition have lower scores, with the reverse being true using SIGNAL-SG, the lack of congruence between the two methods was discerning. Furthermore, EPID classified every site as 'poor' (scores <5.5), which is clearly not the case. Consequently, the eDNA version of the EPID index was deemed unsuitable for this study.

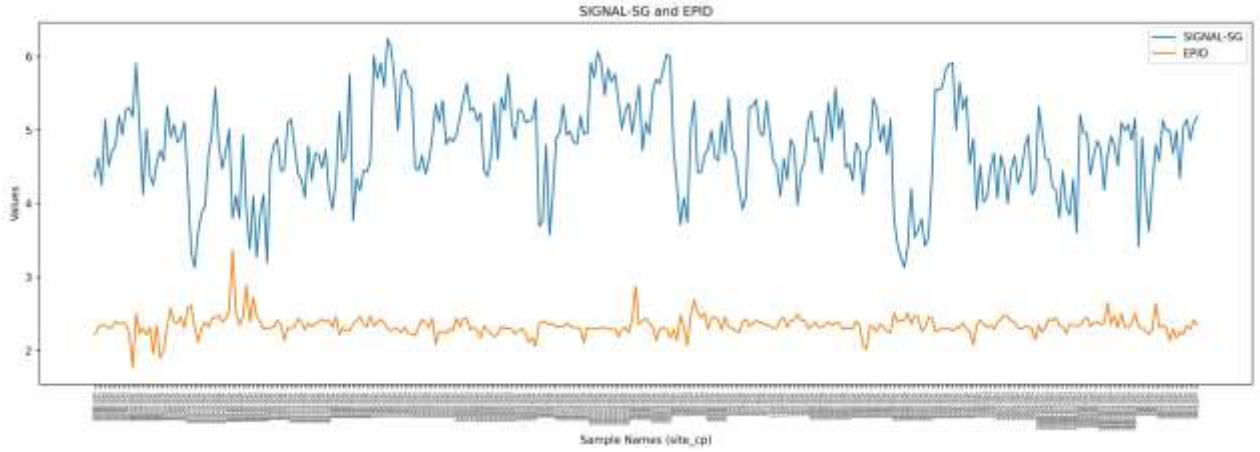


Figure 5-68: Comparison between metabarcoded derived EPID scores and SIGNAL-SG scores for each site.

Mean values are presented. To aid visualization samples were aggregated by collection points (time)

Summary

The sequenced data were spatially extensive (56 sites) and captured six time points. Examination of the data emphasised that the quality of the sequencing was excellent, with a vast majority of the variation in the ecological data being due to biological variability and not due to technical artefacts.

For both bacterial and diatom communities, catchment subsurface geology had a profound influence on composition, and hence, communities from sandstone and shale catchments were analysed separately. In some streams, clear differences between upstream and downstream sites were evident, but this was not universal. The molecular microbial source-tracking marker used as a surrogate for sewage contamination (HFMG CPQ_056) was found in most samples and was correlated with impervious area, suggesting that there were more instances of sewage exposure in more urbanised catchments.

The initial approach of assigning samples to various states (Table 5-4) proved unrealistic and a more correlative approach was taken which included a wider range of environmental variables (Table 5-5). When examined collectively, the findings from the different approaches (multivariate and machine learning) examining the relationships between metabarcoded communities (bacteria and diatoms) and the environmental attributes of the systems all highlighted the pronounced influence static variables had on shaping composition. This included the area of impervious surfaces, road length and riparian cover. Consequently, any signal due to variations in the dynamic variables of interest, for example rainfall, overflow contribution and HFMGs, were less pronounced, poorly correlated and often obscured. This made identifying any overall changes due to WWO spills challenging, constraining the predictive power of the machine learning models.

A number of issues also limited our capacity to discriminate communities exposed to different overflow events. Most notably, flow was modelled and normalised by catchment area, and therefore was applied to the whole site, meaning that the same values were applied to both upstream and downstream sites. This not only reduced our capacity to examine these two positions separately, but also, reduced the number of samples which could be used in the machine learning, as upstream and downstream samples were essentially aggregated. The ubiquitous presence of CPQ_056 suggested that most sites were exposed to some level of sewage, however, quantifying the level of exposure from the abundance of this marker gene was not possible. To further complicate the modelling, the sites were all very different and many were complex. That is, some small streams had many inputs, some large streams a few inputs. Hence, overflow contribution could not be ground-truthed, and in many cases, likely markedly deviated from the modelled data. In essence, this means that there was no true measure of overflow or potential sewage inputs for the sites. Without such information, it is difficult to develop a predictive model since the conditions used to establish the initial states remain ambiguous.

It is emphasised that the data were robust and provided an incredible amount of background and baseline information for each site. Consequently, this information could be used to monitor the trajectories of sites prior- and subsequent-to any remedial actions. This information is currently being drafted as several manuscripts, which will aid in its utility for further pilot studies under the Sydney Water Aquatic Monitoring program. Additional analysis is also being explored to potentially attenuate the influence of the static landscape variables, providing further insight into how communities respond to modelled flow conditions, and what taxa are likely to reflect different flow conditions.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C. 2016. TensorFlow: Large-scale machine learning on heterogeneous distributed systems. ArXiv:1603.04467.

Agarap, A.F. 2018. Deep Learning using Rectified Linear Units (ReLU). ArXiv: 1803.08375.

Aylagas, E., Borja, Á., Muxika, I., Rodríguez-Ezpeleta, N. 2018. Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. Ecol. Indic. 95, 194-202. doi:https://doi.org/10.1016/j.ecolind.2018.07.044

Berney, C., Henry, N., Mahé, F., Richter, D.J., de Vargas, C. 2022. EukRibo: a manually curated eukaryotic 18S rDNA reference database to facilitate identification of new diversity. bioRxiv 2022.11.03.515105.

Besley, C.H., Chessman, B.C. 2008. Rapid biological assessment charts the recovery of stream macroinvertebrate assemblages after sewage discharges cease. Ecol. Indic. 8: 625–638.

Borja, A. 2018. Testing the efficiency of a bacterial community-based index (microgAMBI) to assess distinct impact sources in six locations around the world. Ecol. Indic. 85, 594-602. doi:https://doi.org/10.1016/j.ecolind.2017.11.018

Burnham, K.P., Anderson, D.R. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer, New York.

Chariton, A.A., Court, L.N., Hartley, D.M., Colloff, M.J., Hardy, C.M. 2010. Ecological assessment of estuarine sediments by pyrosequencing eukaryotic ribosomal DNA. Front. Ecol. Environ. 8(5), 233-238. doi:10.1890/090115

Chariton, A.A., Stephenson, S., Morgan, M.J., Steven, A.D.L., Colloff, M.J., Court, L.N., Hardy, C.M. 2015. Metabarcoding of benthic eukaryote communities predicts the ecological condition of estuaries. Environ Pollut. 203, 165-174. doi:10.1016/j.envpol.2015.03.047

Chessman, B.C., Bate, N., Gell, P.A., Newall, P. 2007. A diatom species index for bioassessment of Australian rivers. Mar. Freshw. Res. 58, 542–557.

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., et al. (2014) Ribosomal

Cordier, T., Alonso-Sáez, L., Apothéloz-Perret-Gentil, L., Aylagas, E., Bohan, D. A., Bouchez, A., Lanzén, A. 2021. Ecosystems monitoring powered by environmental genomics: A review of current strategies with an implementation roadmap. Mol. Ecol. 30(13), 2937-2958. doi:https://doi.org/10.1111/mec.15472

Cordier, T., Esling, P., Lejzerowicz, F., Visco, J., Ouadahi, A., Martins, C., Pawlowski, J. 2017. Predicting the Ecological Quality Status of Marine Environments from eDNA Metabarcoding Data Using Supervised Machine Learning. Environ. Sci. Technol. 51(16), 9118-9126. doi:10.1021/acs.est.7b01518

Cordier, T., Forster, D., Dufresne, Y., Martins, C. I. M., Stoeck, T., Pawlowski, J. 2018. Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. Mol. Ecol. Resour. 18(6), 1381-1391. doi:https://doi.org/10.1111/1755-0998.12926

Cordier, T., Lanzén, A., Apothéloz-Perret-Gentil, L., Stoeck, T., Pawlowski, J. 2019. Embracing Environmental Genomics and Machine Learning for Routine Biomonitoring. Trends Microbiol. 27(5), 387-397. doi:10.1016/j.tim.2018.10.012

Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 42: D633–42.

Dell'uomo, A. Torrisi, M. 2011. The Eutrophication/Pollution Index-Diatom based (EPI-D) and three new related indices for monitoring rivers: The case study of the river Potenza (the Marches, Italy). Plant Biosyst. 145: 331–341.

Dowle, E.J., Pochon, X. , Banks, J. C., Shearer, K., Wood, S.A. 2016. Targeted gene enrichment and high-throughput sequencing for environmental biomonitoring: a case study using freshwater macroinvertebrates. Mol. Ecol. Resour. 16, 1240–125.

Edgar, R.C. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26: 2460–2461.

Graham, S. E., Chariton, A. A., Landis, W. G. 2019. Using Bayesian networks to predict risk to estuary water quality and patterns of benthic environmental DNA in Queensland. Integr Environ Assess Manag. 15(1), 93-111. doi:10.1002/ieam.4091

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. Nucleic Acids Res. 41: D597–604.

Hose, G., Turak, E., Waddell, N. 2004. Reproducibility of AUSRIVAS rapid bioassessments using macroinvertebrates. J. North Am. Benthol. Soc. 23(1), 126-139. doi:10.1899/0887-3593(2004)023<0126:Roarbu>2.0.Co;2

Hudson, M. E. 2008. Sequencing breakthroughs for genomic ecology and evolutionary biology. Mol. Ecol. Resour. 8(1), 3-17. doi:https://doi.org/10.1111/j.1471-8286.2007.02019.x

Kingma, D.P. and Ba, J. 2014. Adam: A Method for Stochastic Optimization. ArXiv:1412.6980.

Louca, S., Parfrey, L.W., Doebeli, M. 2016. Decoupling function and taxonomy in the global ocean microbiome. Science 353: 1272–1277.

Meyer, A., Todt, C., Mikkelsen, N.T., Lieb, B. 2010. Fast evolving 18S rRNA sequences from Solenogastres (Mollusca) resist standard PCR amplification and give new insights into mollusk substitution rate heterogeneity. BMC Evol. Biol. 10: 70.

Nichols, S.J., Kefford, B.J., Campbell, C.D., Bylemans, J., Chandler, E., Bray, J.P., Furlan, E.M. 2020. Towards routine DNA metabarcoding of macroinvertebrates using bulk samples for freshwater bioassessment: Effects of debris and storage conditions on the recovery of target taxa. Freshw. Res. 65(4), 607-620. doi:https://doi.org/10.1111/fwb.13443

Nicolosi Gelis, M.M., Sathicq, M.B., Jupke, J., Cochero, J. 2022. DiaThor: R package for computing diatom metrics and biotic indices. Ecol. Model. 465: 109859.

Pentinsaari, M., Salmela, H., Mutanen, M., and Roslin, T. 2016. Molecular evolution of a widely-adopted taxonomic marker (COI) across the animal tree of life. Sci. Rep. 6: 35275.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41: D590–6.

Sutcliffe, B., Hose, G.C., Harford, A.J., Midgley, D.J., Greenfield, P., Paulsen, I.T., & Chariton, A.A. 2019. Microbial communities are sensitive indicators for freshwater sediment copper contamination. Environ Pollut. 247, 1028-1038.

Taberlet, P., Bonin, A., Zinger, L., Coissac, E. 2018. Environmental DNA: For biodiversity research and monitoring: Oxford University Press.

Ter Braak, C.J., Verdonschot, P.F. 1995. Canonical correspondence analysis and related multivariate methods in aquatic ecology. Aquat. Sci. 57, 255-289.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D. 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods. 17: 261–272.

Vivien, R., Wyler, S., Lafont, M., & Pawlowski, J. (2015). Molecular Barcoding of Aquatic Oligochaetes: Implications for Biomonitoring. PLoS One, 10(4), e0125485. doi:10.1371/journal.pone.0125485.

Yang, L., Tan, Z., Wang, D., Xue, L., Guan, M.X., Huang, T., Li, R. 2014. Species identification through mitochondrial rRNA genetic analysis. Sci. Rep. 4: 1–11.

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.H. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat. Rev. Microbiol. 12: 635–645.

5.7 Environmental DNA assemblages in estuarine inshore waters

The initial draft of text within this section of the synthesis report was kindly provided by the Macquarie University project team.

5.7.1 Overview of the estuarine biomonitoring component

The eDNA biomonitoring component for benthic communities described in Section 5.6.2 was also applied to estuarine environments within the Sydney catchment. Specifically, the estuaries of the: Cooks River arm of Botany Bay; Georges River arm of Botany Bay; Lane Cove River; Middle Harbour; Outer Sydney Harbour; Parramatta River; and Port Hacking. From these estuarine systems, sediments were sampled from 26 sites across five collection periods from October 2019 to April 2021.

In contrast to freshwater environments, there are no regional or national biotic indices (equivalent to for example, SIGNAL-SG and diatom indices) for estuarine benthic communities. To address this, we trialled two approaches commonly used in Europe: the AMBI index (Borja et al., 2000), based on soft-bottom sediment benthic macroinvertebrate communities; and the microgAMBI index (Aylagas et al., 2017), which utilises 16S rDNA bacterial data. These biotic indices quantify the collective tolerance of communities to contaminants (metals, pesticides and nutrients) by assigning each taxa with a sensitivity value. Sensitivity values are assigned based on expert knowledge gained from species distributions along known contaminant gradients in Spanish estuaries. For both indices, sensitive species receive low tolerance values, while tolerant species are assigned higher values. As such, for both AMBI and microgAMBI, lower biotic index scores indicate a higher proportion of sensitive species, inferring the systems is of greater ecological integrity. In contrast, higher biotic scores indicate a greater abundance of tolerant species, reflecting impaired benthic communities. For the estuarine component, Random Forest (RF) was the only supervised machine-learning algorithm used as the neural network model was found to be unsuitable.

Objectives of the estuarine biomonitoring study

In this study, our objective was to investigate the use of eDNA biomonitoring to assess benthic communities across different estuary systems within Sydney at differing intensities of wet weather events and overflows. Those events encompass dry periods, minor rainfall occurrences, and periods following substantial rainfalls during which overflow spills infiltrate into the estuaries. However, it is emphasised, due to a lack of spatial and temporal flow data, no exposure data were available during each collection period (point of time). The objectives of this study were to:

1. Examine spatial and temporal patterns in the bacterial and metazoan biotic communities across sites and estuaries
2. Identify the environmental variables, both static and dynamic, which were correlated with eDNA metabarcoded communities
3. Investigate the use of RF models and its potential for outlining important factors (biotic, functional and environmental) that attribute to the microgAMBI biotic index

5.7.2 Sampling, sequencing, bioinformatics and statistical approaches for estuarine samples

Experimental design for estuarine samples

Sediment samples were taken from 26 sites across five collection periods from October 2019 to April 2021. Sediment samples were collected using a clean core, with the surficial layer (approx. 1.5-2 cm) removed for DNA analysis. Samples were collected over a range of rainfall conditions and potential flow wet weather flow rankings (Table 5-10). At each site, conductivity was measured at the bottom of the water column. At each site, five sediment samples were taken for eDNA which was extracted and amplified for five loci/genes: 16S rDNA (bacteria); 18S rDNA (eukaryotes); 18S rDNA (diatoms); 18S rDNA (amphipods); and COI (metazoans). For each sample, 3 separate polymerase chain reactions (PCRs) were performed. In addition, 240 synthetic sequences (positive controls) and 720 negative controls (blank samples using DNA-free water) were sequenced for quality assurance. Samples were sequenced on an Illumina NextSeq800 at the Ramaciotti Centre for Genomics (UNSW).

Table 5-10: Ranking across five collection periods based on rainfall

Collection period	Date	Conditions	Rank
7	October 2019	Return of rainfall 40 to 50 mm after 4 intensely dry months	Low
10	March 2020	Sediment collections post east coast low major event in February 2020 300+ mm	Major
13	August 2020	Sediment collection post moderate weather event 150+ mm over 3 days	Moderate
16	February 2021	Many rainy days accumulated 150+ mm	Moderate
17	April 2021	Sediment collections post east coast low major weather event March 2021 300+ mm	Major

Bioinformatics and statistical analyses

The bioinformatics, PCA and CCA statistical analyses methods used for the estuarine data were the same as that described previously (see Section 5.6.2). The suitability of two indices was examined for the Random Forest (RF) models: AMBI (Borja et al., 2000) and microgAMBI (Aylagas et al., 2017). The AMBI index, was produced using COI metazoan data, and calculated using the AMBI software v6.0 (Borja et al., 2000). Specifically, the input file consisted of an OTU table containing COI amplicons filtered to include only metazoans. Similarly, the microgAMBI index was computed utilising the 16S rDNA bacterial data with the microgAMBI bacterial taxonomic list comprising genus and species classifications, along with their corresponding sensitivity values. It is important to note that AMBI is generally calculated using macrofauna data (not eDNA data), while microAMBI is specifically designed for eDNA-derived bacterial data. It is emphasised that the samples collected in this study were not designed to capture macrofauna, as this would have required a far larger volume of sediment.

Environmental variables

The environmental variables used in the estuarine component are provided in Table 5-11. For a subset of the sites, additional information on overflow volume (Jan 2020 to Apr 2021) and the number of stormwater events with overflow spills was also available (Table 5-12). It is important to note that when using these variables there was only complete environmental data for 18 sites across four collection periods from March 2020 to April 2021. Furthermore, total overflow volumes and spill events with overflow spills are based on measurements which capture the whole temporal scale of this study and were not available for each collection point (sampling occasion).

Table 5-11: List of environmental variables used for assessing the relationships with the metabarcoded derived estuarine communities

Variable	Abbreviation	Description
Natural log impervious road area	ln_imp	Log of impervious road area
Natural log road density	ln_rod	Log of road density
Natural log crAssphage	ln_cpq_056	Log of the microbial source tracking marker CPQ_056
Road density by rainfall 90 days	road_density_rain90	Road density multiplied by 90 days rainfall period
Log mud sand ratio	lr_mud_sand	Log to the ratio of mud (<63 µm) and sand (>=63 µm to <2.0 mm)
Log sand gravel ratio	lr_sand_gravel	Log to the ratio of sand and (>=63 µm to <2.0 mm) and gravel (>=2.0mm)
Conductivity	Conductivity	Conductivity of lower water column (µS/cm)

Table 5-12: List of sites with complete environmental data and modelled overflow exposure categories

Estuary area	Estuary site label	Site name	Total overflow volume (ML)	Storm events with overflow spills
Cooks	CRSED01	Cooks SWSOOS	1605	24
Georges	GRSED01	Kogarah	3	11
	GRSED02	Lime Kiln	18	2
	GRSED03	Edith	40	3
	GRSED04	Salt Pan	1089	13
Lane Cove	LCSED01	Lane Cove NSOOS	109	13
Middle Harbour	MHSED01	Clontarf	1	4
	MHSED02	Quakers Hat	4	6
	MHSED03	Davidson Park	245	9
Outer Sydney	SHSED01	Chowder	1	5
	SHSED03	Mosman	0	5
Parramatta	PRSED02	Iron Cove Inner	51	31
	PRSED03	Tarban	118	2
	PRSED04	Hen and Chicken	33	26
	PRSED06	Meadowbank	16	5
Port Hacking	PHSED01	Yowie Bay	139	8
	PHSED02	Gymea Bay	14	12
	PHSED04	North West Arm	96	5

5.7.3 Key results and findings of the estuarine biomonitoring component

Spatial and temporal patterns in eDNA metabarcoded bacterial and metazoan communities in an estuarine environment

The bacterial communities from the estuarine samples are presented in the PCA plot (Figure 5-69). PERMANOVA analysis (based on 9,999 permutations) are presented in Table 5-13, and show that there was a significant interaction between sites and collection period, among sites and among estuaries. The largest proportion of explained variability was at the site scale (R^2= 0.22). When viewed collectively, all estuaries have different bacterial assemblages (P<0.05) Consequently, for most estuaries, clearer visualisation patterns were possible when the ordinations were presented at the estuary scale (Figure 5-70). Post-hoc analysis found that within each estuary, all sites had different benthic bacterial communities (P<0.05). The exception to this was Glades Bay, which had similar bacterial communities to two other Parramatta River sites (Meadowbank and Hen and

Chicken Bay). In some cases, sites contained bacterial communities that were markedly different to those within the same region/estuary, for example, Mosman in Outer Sydney Harbour to Chowder and Taylors Bays (Figure 5-70).

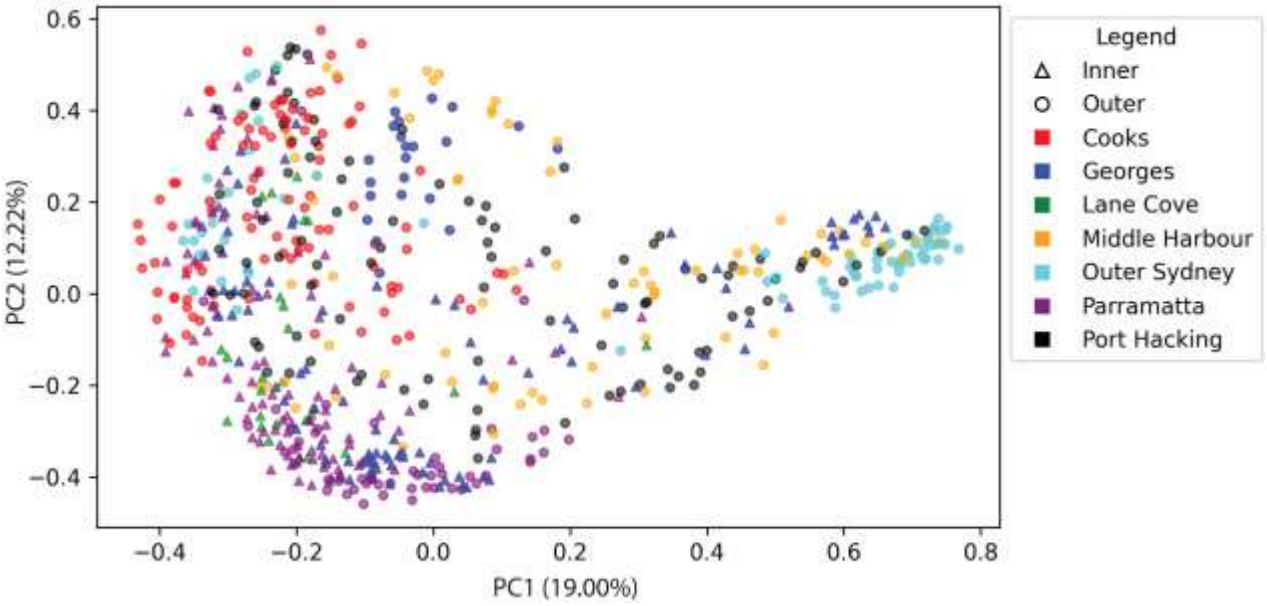


Figure 5-69: A PCA representing communities from all bacterial samples combined at the collection point (time)

The closer the points are to each other, the more similar the communities are in composition. Triangles represent inner (triangle) salinity zone sites and circles outer (circle) salinity sites

Table 5-13: PERMANOVA analysis determining differences in bacterial communities (16S rDNA) between treatments and their interactions.

Treatment	Degrees freedom	Sum of squares	R^2	F	P-value
estuary	6	29.6	0.16	33.7	0.001
site	19	44	0.24	15.8	0.001
site:cp	104	34.6	0.19	2.27	0.001
Residual	515	75.4	0.41		
Total	644	183.5	1		

The symbol ':' denotes an interaction term. cp=collection point (temporal)

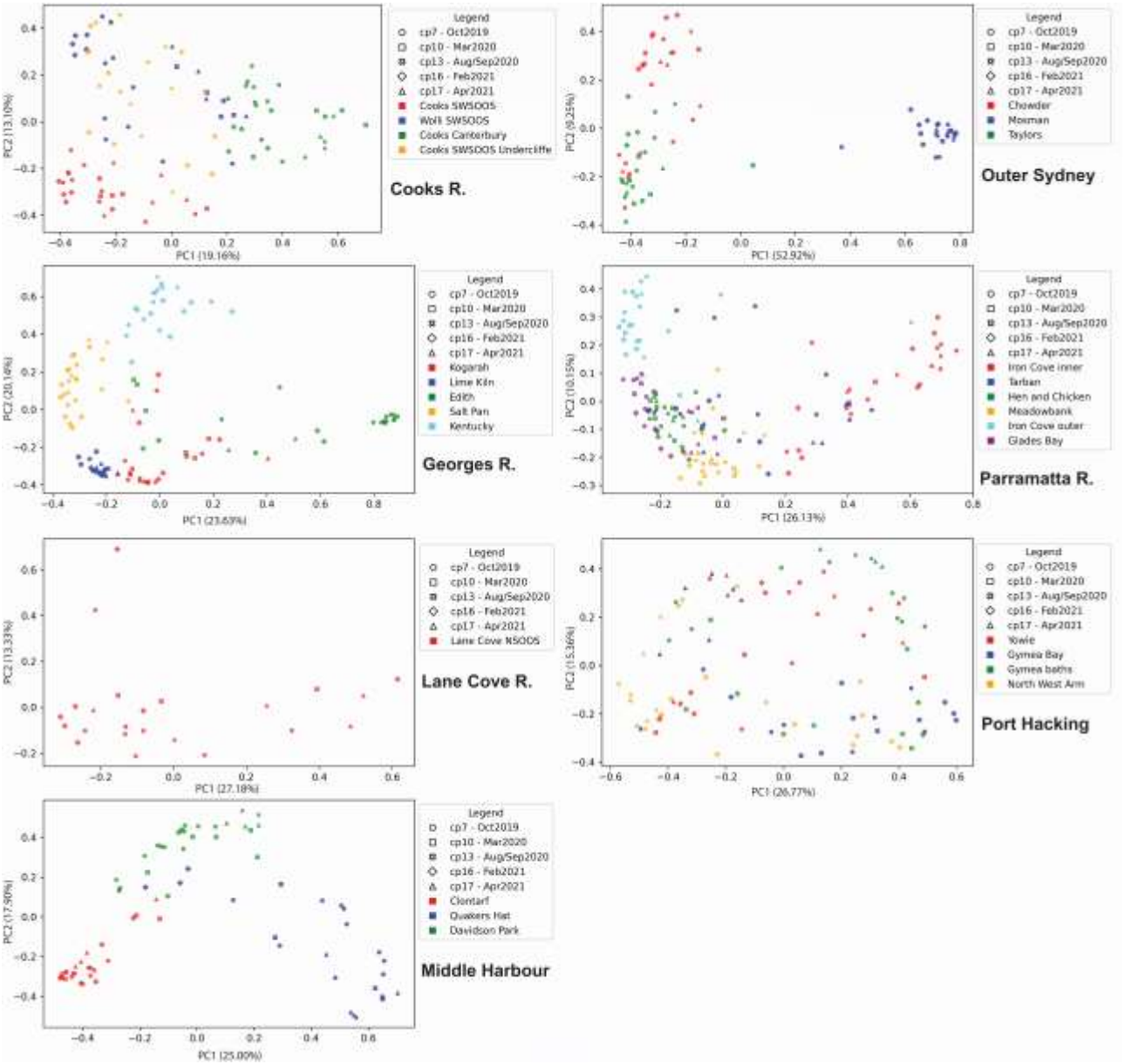


Figure 5-70: PCAs representing bacterial communities from each estuary combined at the collection point (time)

Various markers (shown in the legend) are used to represent collection periods

Post-hoc comparisons of the interactions between site and collection point (time) are summarised in Table 5-14. Red cells are where differences in communities were found between the two time points, and yellow cells are where no differences were found. The final column indicates any collection points where the benthic communities returned to a similar composition to those sampled at the commencement of the study (CP7) where rainfall was low. The benthic bacterial communities from two sites within the Parramatta River (Hen and Chicken Bay and Iron Cove outer) and one site within the Georges River (Kogarah) remained similar across all collection periods. In contrast, a number of sites from the Georges River (Lime Kiln and Salt Pan), Middle Harbour (Davidson Park and Quakers Hat), outer Sydney (Mosman) and Port Hacking (North West Arm), had communities which varied significantly between each collection period. In the majority of sites, benthic communities returned to similar compositions as found at the commencement of the study (CP7). This predominately occurred at CP13, indicating recovery from the major rainfall event at CP10. Furthermore, in a number of sites (for example, Lane Cove), the communities remained similar to those from CP7 for the duration of the study. Hence, in these cases, the benthic communities were only significantly modified from the major rainfall event, returning to the initial state even when subjected to moderate rainfall events. Collectively, this suggests that in some sites, the bacterial communities may reflect pronounced rainfall and run-off events, and thus, be a potential indicator for wet-weather events.

Table 5-14: Post hoc pairwise analysis on the collection periods for the estuarine 16S rDNA bacterial communities

Estuary	Site name	Collection points and conditional rank				CPs which resemble the initial communities sampled at CP7
		CP7 vs CP10 Low vs Maj.	CP10 vs CP13 Maj. vs Mod.	CP13 vs CP16 Mod. vs Mod.	CP16 vs CP17 Mod. vs Maj.	
Cooks	Cooks Undercliffe					
	Cooks SWSOOS					
	Wolli					13
	Cooks Canterbury					13
Georges	Lime Kiln					
	Kentucky					13
	Salt Pan					
	Kogarah					13, 16, 17
	Edith					
Lane Cove	Lane Cove					13, 16, 17
Middle Harbour	Davidson Park					
	Quakers Hat					13
	Clontarf					
Outer Sydney	Chowder					13
	Mosman					13
	Taylors					13, 16, 17
Parramatta	Meadowbank					16
	Hen and Chicken					13, 16, 17
	Iron Cove outer					
Parramatta	Iron Cove inner					13
	Glades Bay					13
	Tarban					
Port Hacking	Yowie					13, 16, 17
	Gymea Bay					13, 16
	North West Arm					13
	Gymea Baths					13

Orange cells indicate significant differences between the two collection points, blue cells indicated not significant difference. The final column indicates any collection points where the bacterial communities were similar ($p>0.05$) to the communities initially sampled (CP7) during a low rainfall period. The abbreviation cp = collection period. Low, Mod (moderate) and Maj (Major) are as defined in Table 5-10

The metazoan communities from the estuarine samples are presented in the PCA plot (Figure 5-71). PERMANOVA analysis found a significant interaction between sites and collection period, with significant differences also occurring among sites and estuaries (Table 5-15). PCAs of the metazoan communities for individual estuaries/systems are presented in Figure 5-72. Post-hoc analysis found the benthic metazoan communities differed (P<0.05) between all sites within each estuary.

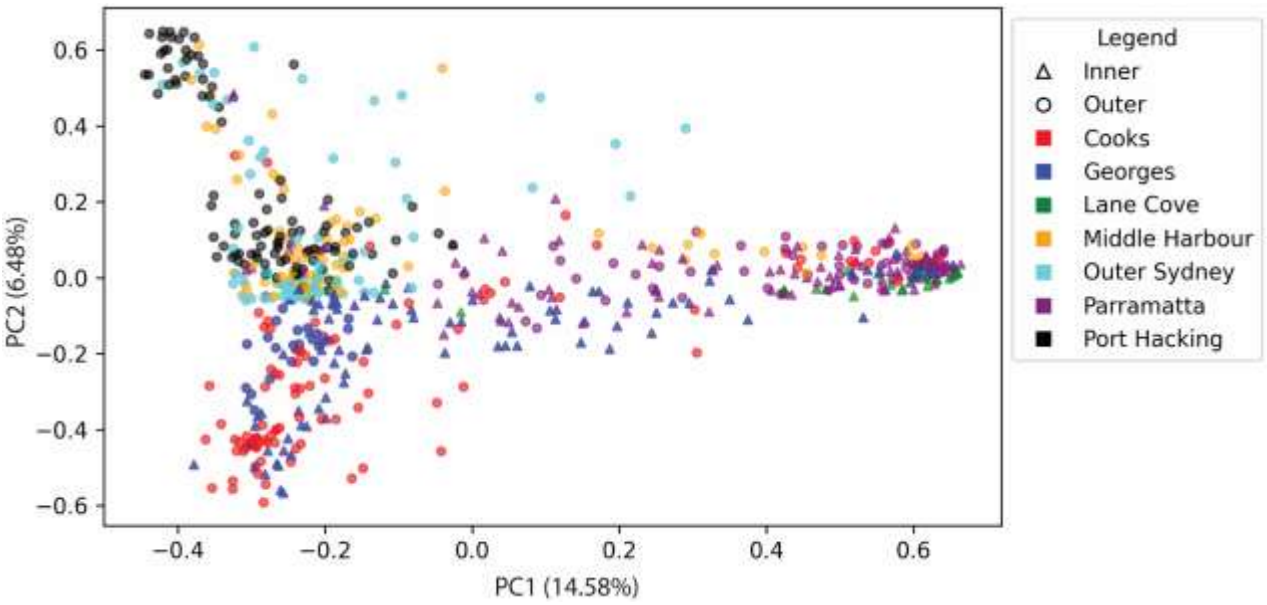


Figure 5-71: A PCA representing communities from all metazoan samples combined at the collection point (time)

The closer the points are to each other, the more similar the communities are in composition. Triangles represent inner salinity zone sites and circles outer salinity zone sites

Table 5-15: PERMANOVA analysis determining differences in metazoan communities (COI mtDNA) between treatments and their interactions

Treatment	Degrees freedom	Sum of squares	R^2	F	P-value
Estuary	6	43.4	0.15	28.8	0.001
Site	19	42.7	0.15	8.95	0.001
site:cp	104	68.1	0.24	2.6	0.001
Residual	511	128	0.45		
Total	640	283	1		

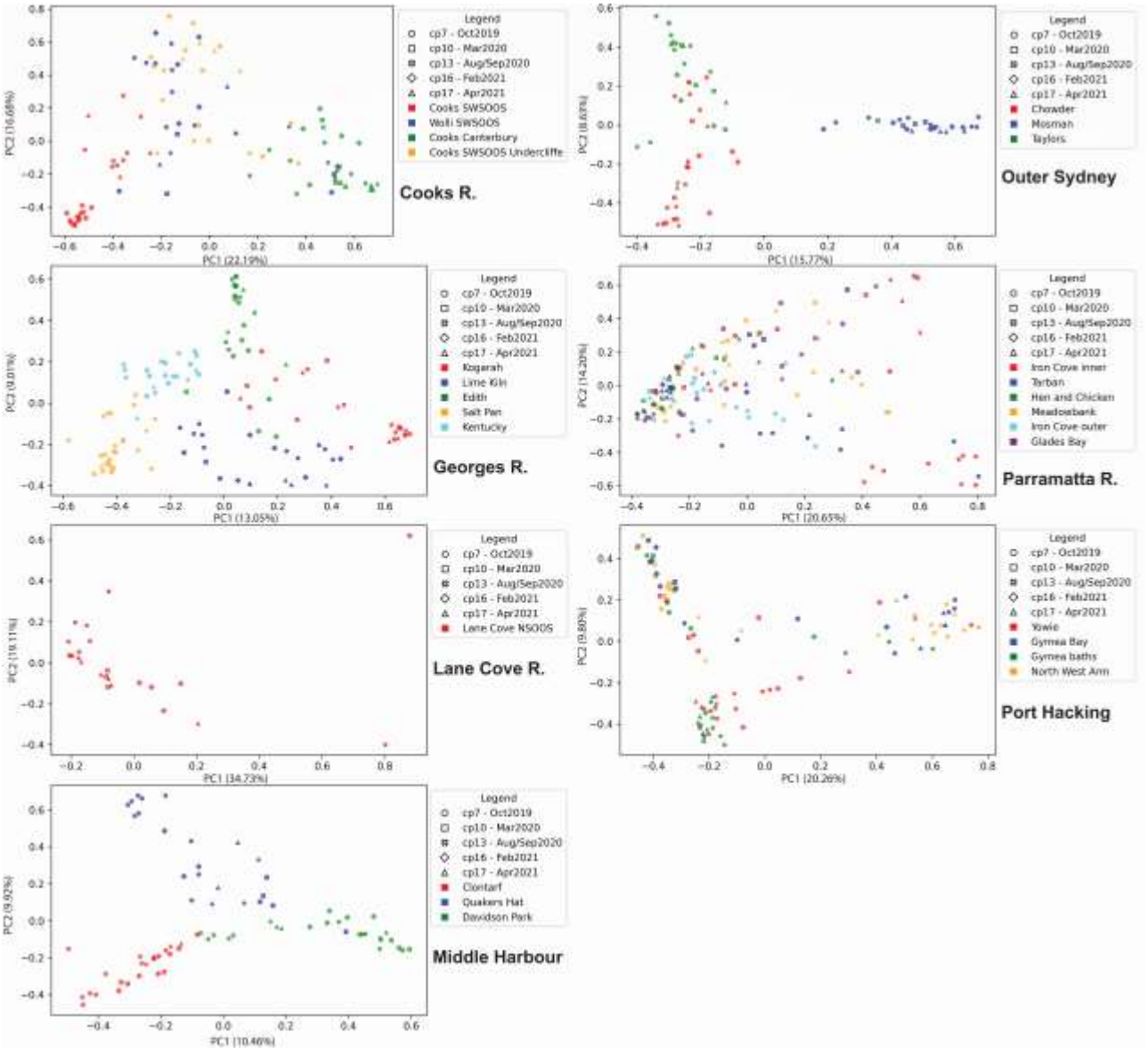


Figure 5-72: PCAs representing metazoan communities from each estuary combined at the collection point (time)

Various markers (shown in the legend) are used to represent collection periods

Post-hoc comparisons of the interactions between sites and collection points (time) for the metazoan benthic communities are summarised in Table 5-16. In contrast to the bacterial communities, in the vast majority of sites the communities remained stable (that is, not significantly different) across the duration of the study. Only one site (Kentucky) appeared to change between the low and subsequent major rainfall event (CP7- CP10), with the community returning to its initial state in subsequent collection periods (CP13-17). Temporal differences occurred at some sites where rainfall states changed from moderate to major (for example, North West Arm and Gymea Baths), however, these sites did not change between low and major rainfall events. Three sites (Salt Pan, Edith and North West Arm) had different metazoan communities at the end of the study following a major rainfall event when compared to the initial low rainfall event captured in CP7. In summary, our findings suggest that metazoan communities are site specific and in generally vary little to marked rainfall events. This suggests metazoan communities might not be ideally suited for monitoring wet weather events. We emphasise that in this study metazoans are predominately micro- and meio-fauna as the samples were not designed to capture macrofauna. Hence, it is not possible to state whether the same patterns would occur in macrofaunal communities.

Table 5-16: Post hoc pairwise analysis on the collection periods for the estuarine metazoan communities

		Collection points and conditional rank			
Estuary	**Site name**	**CP7 vs CP10**	**CP10 vs CP13**	**CP13 vs CP16**	**CP16 vs CP17**
		Low vs Maj.	**Maj. vs Mod.**	**Mod. vs Mod.**	**Mod. vs Maj.**
Cooks	Cooks Undercliffe				
	Cooks SWSOOS				
	Wolli				
	Cooks Canterbury				
Georges	Lime Kiln				
	Kentucky				
	Salt Pan				
	Kogarah				
	Edith				
Lane Cove	Lane Cove				
Middle Harbour	Davidson Park				
	Quakers Hat				
	Clontarf				
Outer Sydney	Chowder				
	Mosman				
	Taylors				
Parramatta	Meadowbank				
	Hen and Chicken				
	Iron Cove outer				
Parramatta	Iron Cove inner				
	Glades Bay				
	Tarban				
Port Hacking	Yowie				
	Gymea Bay				
	North West Arm				
	Gymea Baths				

Orange cells indicate significant differences between the two collection points, blue cells indicated not significant difference. The abbreviation cp = collection period. Low, Mod (moderate) and Maj (Major) are as defined in Table 5-10

Relationships between benthic communities (bacteria and metazoans) and environmental variables

A CCA ordination plot (Figure 5-73) was generated using the bacterial 16S rDNA estuarine data using the estuarine environmental variables (Table 5-11 and Table 5-12). The ordination highlights several gradients, most notable was granulometry (mud:sand and sand:gravel) as well as road density by 90 day rainfall.

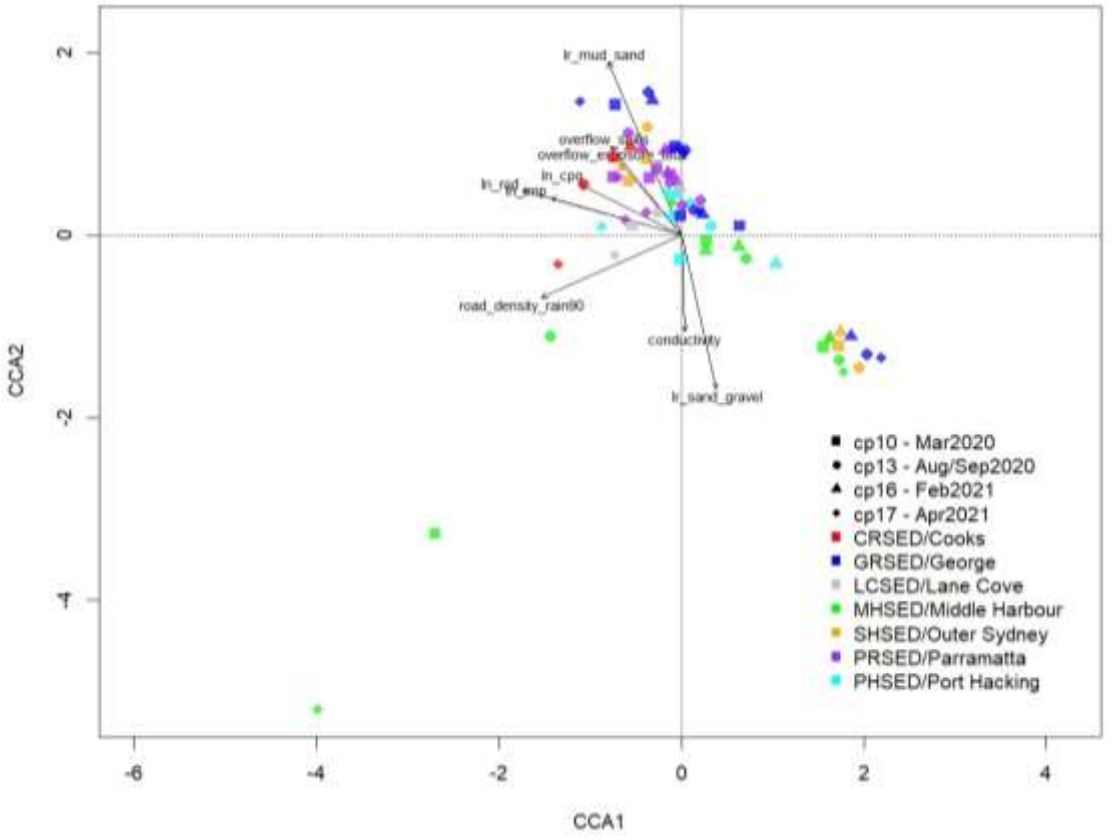


Figure 5-73: Canonical correspondence analysis (CCA) ordination biplots derived from the environmental data (Table 5-11 and Table 5-12) and bacterial community estuarine data
The longer the vector the stronger its correlation with the 16S rDNA bacteria community estuarine data

In total, the measured environmental variables explained 30.14 % of the variation in bacterial community data (Table 5-17). The variable road density (natural log) explained the largest proportion of variation in the bacteria data (≈ 6 %). Collectively, granulometry (ratios of mud:sand and sand:gravel) explained 9.5 % of the variation in the bacterial data. The other variables: craAssphage, overflow exposure, road density by rainfall, conductivity and impervious road all explained similar amounts of variation (≈ 3% each). Collectively, the results suggest that the dominant measured variables driving bacterial communities are associated with granulometry and habitat modification (road density). While it is difficult to quantify, catchments with a high road density are likely to have increased levels of urbanisation and run-off, which maybe a proxy for higher levels of wastewater run-off into the system. This also adds credence to the temporal dynamics in the bacterial communities (Table 5-14).

Table 5-17: Total explained variation and the variation explained by significantly correlated environment variables from the CCA of the 16S rDNA bacterial data

Variable	F	P-value	Contribution (%)
Road density	5.10	0.001	6.01
mud:sand	4.79	0.001	5.65
sand:gravel	3.30	0.001	3.89
crAssphage CPQ_056	2.74	0.002	3.23
Road density by rainfall (90 days)	2.44	0.001	2.87
Overflow exposure	2.46	0.001	2.84
Conductivity	2.40	0.001	2.83
Impervious road area	2.38	0.008	2.82
Total amount of explained variation			30.14

A CCA ordination plot generated from the metazoan data (COI gene) is presented in Figure 5-74. A qualitative assessment of the plot suggests that the environmental gradients are less pronounced than those observed in the bacteria data (Figure 5-73). As expected, metazoans from the outer Harbour were generally correlated with high conductivity. Assemblages from sites within some estuaries were very dispersed and appeared to not to be significantly influenced by the variables measured in this study, for example, the Middle Harbour sites in the lower right quadrant (Figure 5-74).

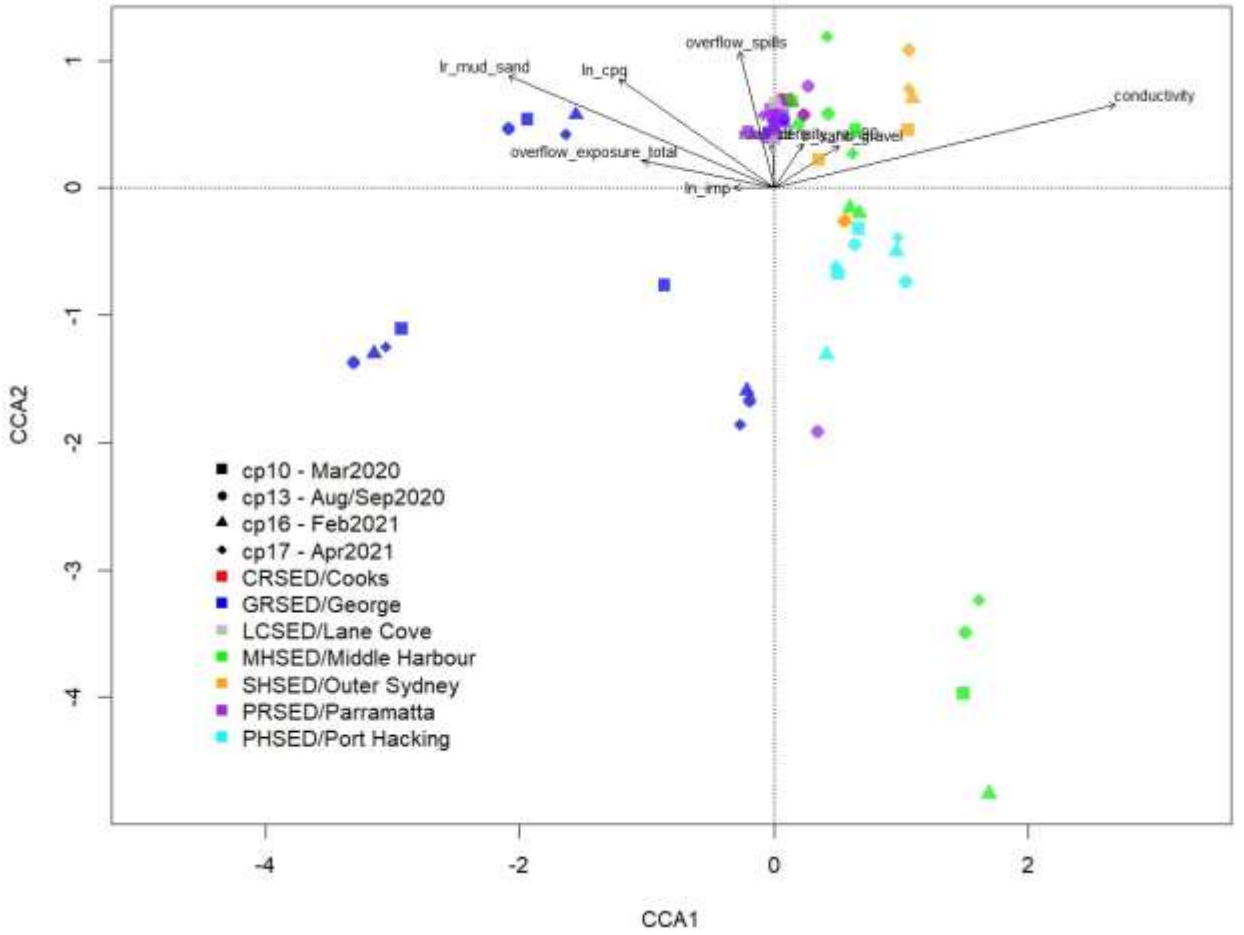


Figure 5-74: Canonical correspondence analysis (CCA) ordination biplots derived from the environmental data (Table 5-11) and metazoan composition

The longer the vector the stronger its correlation with the COI metazoan community estuarine data

Only 19.8 % of the variation in the metazoan data could be explained by the measured variables (Table 5-18). The ratio of mud to sand explained the largest amount of variation (3.7%) in the metazoan community data. Salinity explained a similar amount of variation in the metazoan data (2.8%) as it did in the bacterial community data (2.8%). Individually, variables associated with urbanisation (that is, road density, impervious road modification) explained similar amounts of variation (≈ 2 %) as those associated with wet weather overflow (crAssphage, total overflow exposure and overflow spills).

Table 5-18: Total explained variation and the variation explained by significantly correlated environment variables from the CCA of the 16S rDNA bacterial data

Variable	F	P-value	Contribution (%)
mud:sand	2.69	0.001	3.67
Conductivity	2.04	0.001	2.78
Overflow exposure	1.79	0.001	2.44
crAssphage CPQ_056	1.69	0.002	2.30
Road density	1.69	0.001	2.31
Sand:gravel	1.60	0.001	2.18
Impervious road area	1.59	0.004	2.17
Total overflow exposure	1.75	0.001	2.38
Overflow spills	1.50	0.008	2.05
Total amount of explained variation			19.82

Random forest (RF) models using estuarine data

Prior to performing any machine learning modelling, AMBI (metazoan-based index) and microgAMBI (bacteria-based index) scores were calculated for each site, shown in Figure 5-75 and Figure 5-76, respectively. Our findings suggest that AMBI was not suitable for Sydney's estuaries, producing results which were contrary to our understanding and extensive experience in these systems. For example, Parramatta River sites were classified as having an undisturbed ecological status and were deemed to be of greater ecological integrity to those sites sampled in the outer Harbour (Figure 5-75). In contrast, microgAMBI data showed that a majority of the Parramatta River sites to be categorised as moderate (Figure 5-76). In general, microgAMBI showed a better fit for what was expected for these systems, however, it also appeared to underestimate the degree of disturbance in many sites. Comparisons between the AMBI and microgAMBI scores for the sites showed that the two indices were poorly correlated (R^2=0.21) (Figure 5-77). Given this and the lack of consensus between the known relative conditions of many of the sites and AMBI scores, no further analysis was performed using AMBI, although microgAMBI was used in the Random Forest modelling.

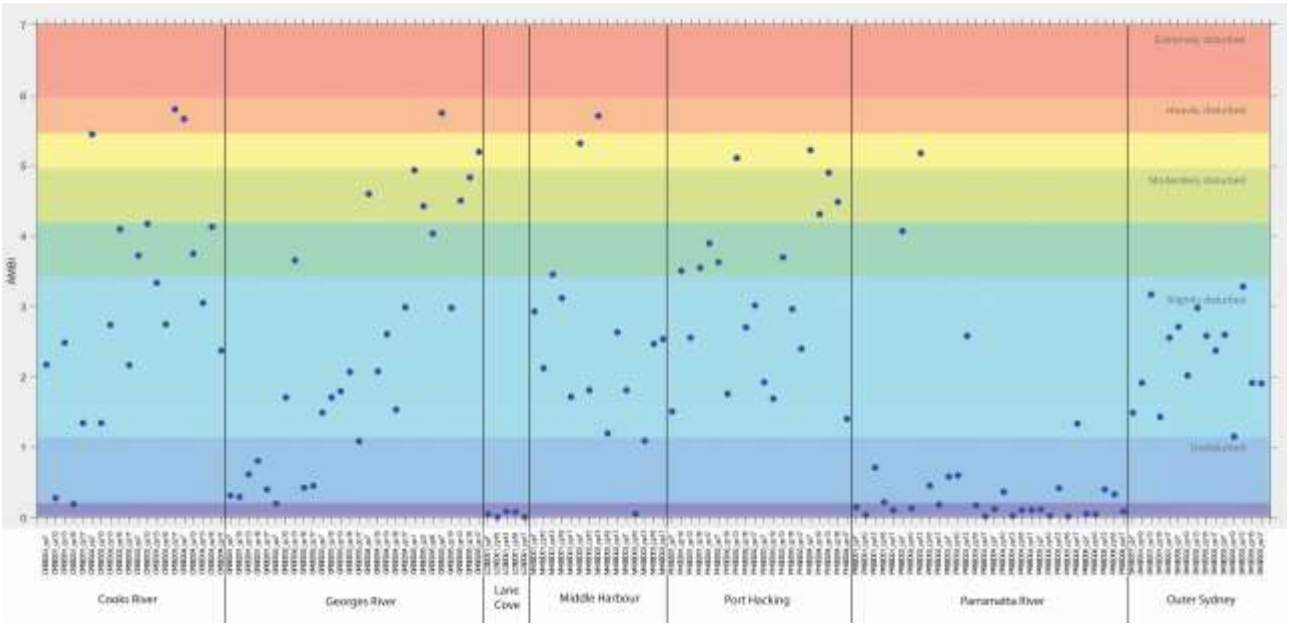


Figure 5-75: The distribution of the AMBI index using the COI metazoan samples across sites and collection periods

The various levels of ecological status, increasing in biotic index, are shown from undisturbed, slightly disturbed, moderately disturbed, heavily disturbed, and extremely disturbed

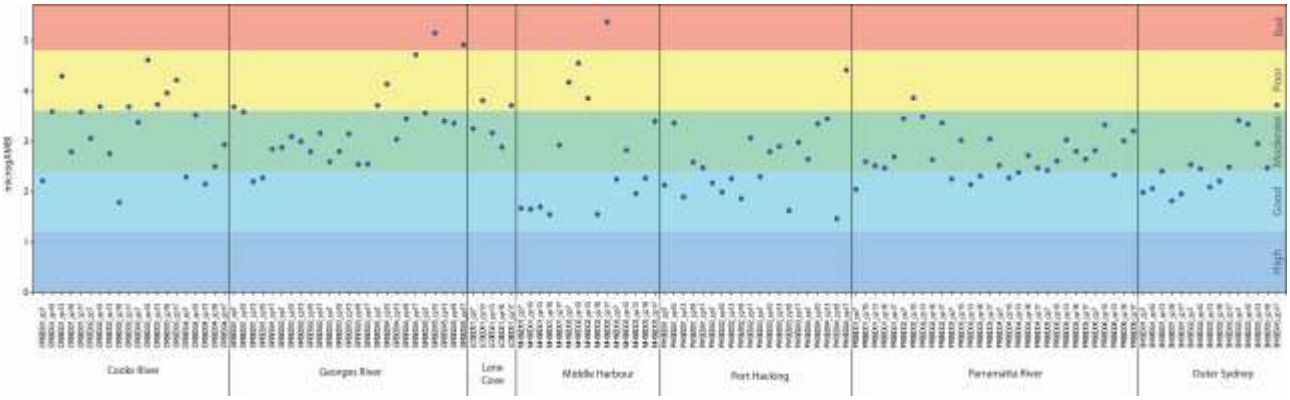


Figure 5-76: The distribution of the microgAMBI index using the 16S rDNA bacterial samples across sites and collection periods

The various levels of ecological status, increasing in biotic index, are shown from high, good, moderate, poor, and bad

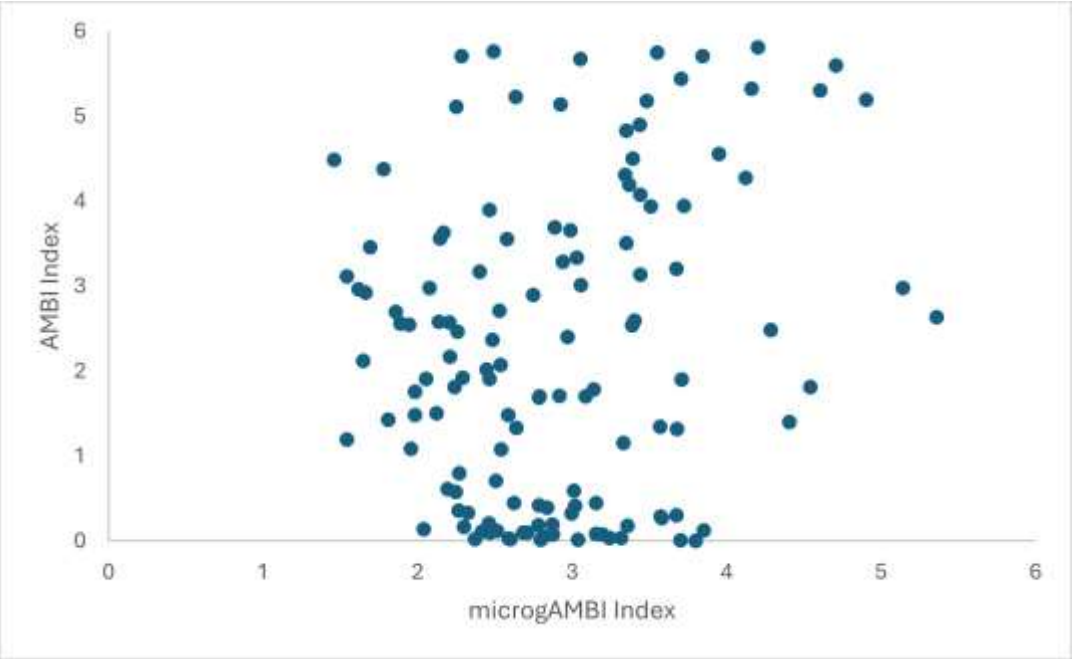


Figure 5-77: Biplot of the AMBI and microgAMBI scores for each sampled site

The same computational architecture for the RF models follows that previously described in Section 5.6.2. Datasets derived from the five genes/loci were used to capture ecological communities consisting of: bacteria, eukaryotes, metazoans, diatoms and amphipods. In addition, the environmental variables used in the CCAs and inferred functional data derived from the analysis of the bacterial data using FAPROTAX (Section 5.6.2) were included as part of the feature set. This approach was chosen to ensure comprehensive biotic coverage, functional processes, as well as the influence of static and dynamic environmental variables when modelling the estuarine systems. To prevent redundancy in the predictive features, the bacterial data were excluded from the microgAMBI RF models, as these data were already used to develop the microgAMBI scores. Finally, the biotic data were run at two levels of phylogenetic resolution, genus and species.

Table 5-19: The performance metrics for the random forest analyses of the bacteria, eukaryotes, diatoms, amphipods, and metazoans at species and genus levels, and the environmental data using microgAMBI as the label

Data type	Phylogenetic resolution	R^2	Mean squared error
Eukaryotes	Species	0.48	0.32
	Genus	0.43	0.35
Diatoms	Species	0.20	0.50
	Genus	0.34	0.41
Amphipods	Species	0.03	0.41
	Genus	0.31	0.29
Metazoans	Species	0.35	0.28
	Genus	0.37	0.27
Environmental	Not applicable	-0.68	0.21
FAPROTAX	Not applicable	-0.01	0.13

The results for the RF models using microgAMBI as the label are shown in Table 5-19. It should be noted that Psuedo-R-squared (here on referred to as R^2), estimates how well the model would do at predicting data that were not used to fit the model. Ideally, models should have a high R^2 with a low mean squared error value (variability).

The best RF models were those which used the eukaryotic data derived from the sequencing of the 18S rDNA gene. This model performed best at the species level (R^2=0.48, mean squared error = 0.32), with the RF model from the eukaryotic species data explaining approximately 50% of the variation in the microgAMBI data. The negative R^2 for the environmental data indicated that the analysis was not predictive, possibly due to inadequate data (Chicco et al., 2021). The value near zero for the RF model using the FAPROTAX data suggested that none of the variability in the microgAMBI data could be explained by the inferred functional attributes of the bacterial data (Chicco et al., 2021). Consequently, no further analysis was run using the FAPROTAX derived functional data.

The eukaryotic taxa which were most important for the microgAMBI RF models were extracted at both the species (and higher) and genus level (and higher) (Figure 5-78). However, it is emphasised that the taxonomy and sequence databases for most aquatic eukaryotes, especially micro- and meio-fauna, was poor, and hence the taxonomy was notional and based on the nearest match to the database. At the species level, the most important taxa were the: gastotrich *Heterolepidoderma* spp., the ciliate *Urosoma salmastra* and the ostrocod *Paracypria* spp. As expected, these taxa were also identified at the genus-level analysis, with the primary difference being the inclusion of an OTU associated with a nematode, however, it could only be assigned to the order Chromadorida.

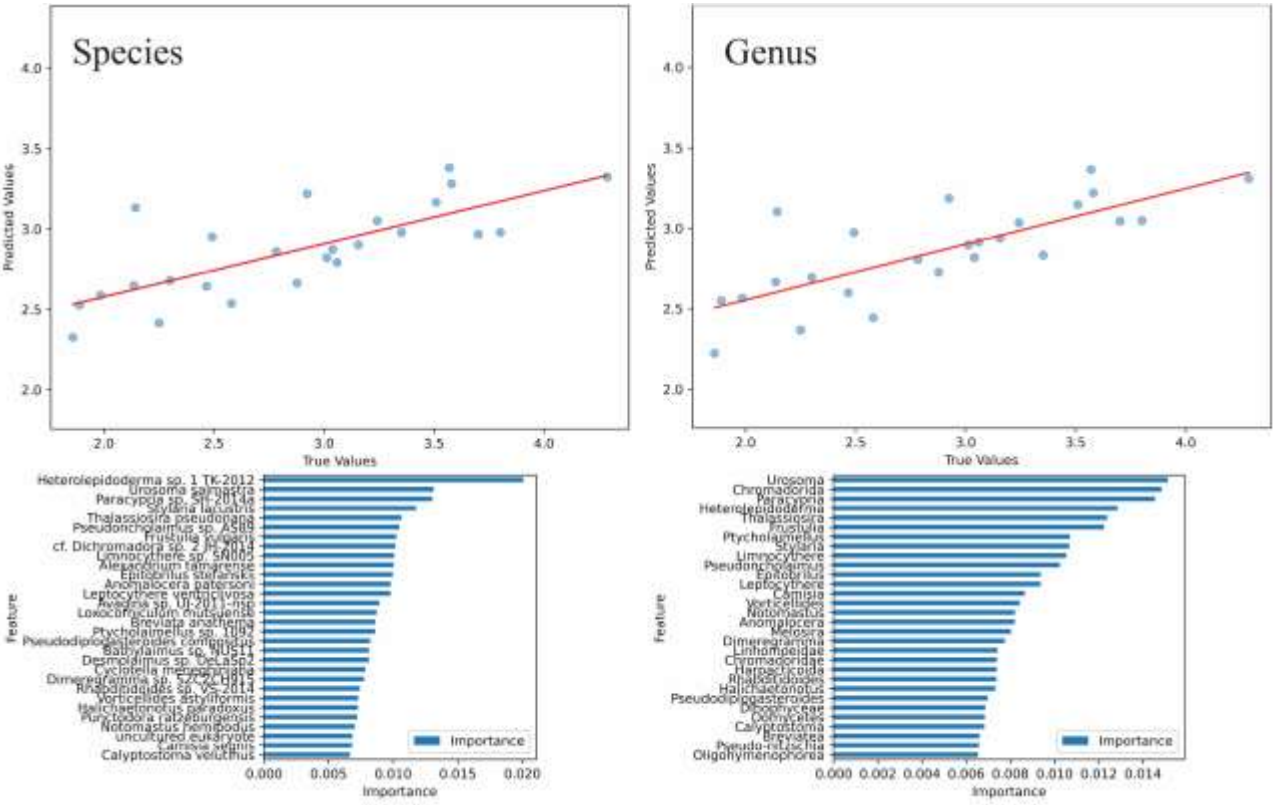


Figure 5-78: RF prediction plots and feature importance of eukaryotic 18S rDNA (y-axis) using microgAMBI (x-axis) as the label

Figures on the left are at species level, with genus level presented on the right. Taxa with a higher importance contribute more in predicting the label than those of lower importance

As previously indicated, the negative R-squared value for the RF using the environmental data implies a poor model fit or lacked adequate data (Table 5-19). While the overall model was a poor fit, one variable had very high importance, road density by rainfall 90 days (Figure 5-79).

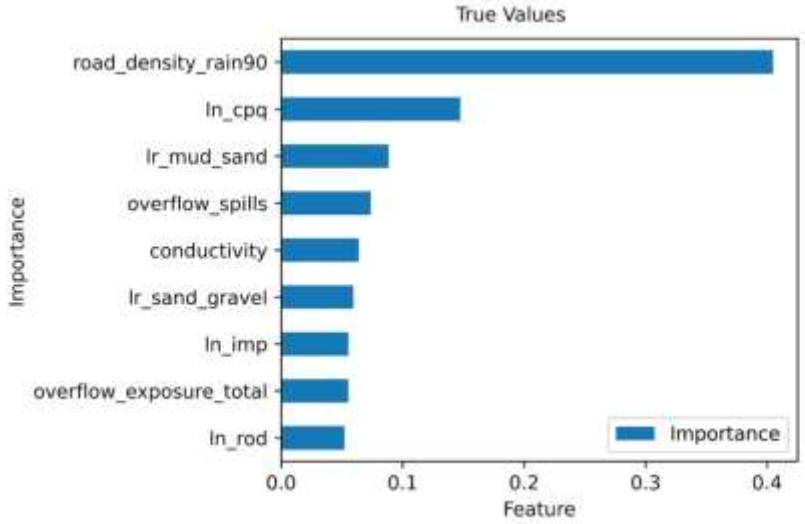


Figure 5-79: RF prediction plots and feature importance of estuarine environmental data (y-axis) using microgAMBI as the label

Environmental variables with a higher importance contribute more in predicting the label than those of lower importance

Summary

Environmental-DNA derived compositional data targeting benthic estuarine bacteria, eukaryotes, metazoans, diatoms and amphipods was obtained from 26 sites across five collection periods. The collection periods captured a range of rainfall conditions ranging from low to major (Table 5-10). Environmental variables were predominantly static, the exceptions being MST data for crAssphage (CPQ056) and conductivity (Table 5-11). Modelled total flow volume and the number of storm events with overflow spills was only available for 18 of the 26 sites (Table 5-12). Furthermore, these data were temporally aggregated, and therefore static across the length of the study.

Bacteria communities were estuary and site dependent and frequently varied over the collection points (Table 5-13). Importantly, most bacterial communities responded to the pronounced rainfall events (major) and generally returned to their initial state (low rainfall) (Table 5-14). This suggests that bacterial communities maybe a responsive indicator of overflow events in estuarine systems. Metazoans communities also varied among and within estuaries and over time (Table 5-15), however, responses to specific rainfall events and their recoveries to the initial state were unclear at most sites, the exception being Kentucky within the Georges River (Table 5-16). As such, metazoans did not appear to be ideally suited to catch the temporal responses associated with wet-weather overflow events.

The amount of variation explained by the measured and modelled environmental variables was much greater in the bacterial data (Table 5-17) than that in the metazoan data (Table

5-18). Grain size and road density were the best predictors of microgAMBI scores. To a lesser degree, but still significant, were a range of variables associated with catchment modification and potential overflow volumes and occurrences (Table 5-17).

The RF modelling showed that the eukaryote data (18S rDNA) best fitted the microg-AMBI scores (Table 5-19). With this approach also identifying potential indicator taxa. Given this, additional analysis is currently being run, but not presented, using the eukaryote dataset to examine its spatial and temporal attributes and its responses to wet-weather events.

Collectively, the available data suggests that there are no regional predictive patterns in estuarine biota to rainfall and overflow events. This may be attributed to a number of reasons: 1) the natural variability in the biotic communities, including their marked differences within and between estuaries; 2) the communities do not respond or respond in an unpredictable manner to wet-weather events; and 3) a lack of suitable environmental data, including temporal overflow data for each site. The latter not only limits the number of sites which could be analysed, but also, hinders linking the machine learning to temporal responses, that is, direct responses to wet-weather events. As in the case of the freshwater component, without a true or accurate measure of overflow and wastewater inputs at site scale, it is not possible to model a biological response. Importantly, the bacteria (16S) data show promise as a site-specific indicator, being responsive to rainfall events, as well as recovering to an initial state. Further analysis is currently being completed to determine the suitability of the eukaryote data if similar data patterns are afforded like those observed from the bacterial data.

References

Aylagas, E., Borja, Á., Tangherlini, M., Dell'Anno, A., Corinaldesi, C., Michell, C.T., Irigoien, X., Danovaro, R., Rodríguez-Ezpeleta, N. (2017) A bacterial community-based index to assess the ecological status of estuarine and coastal environments. Marine Pollution Bulletin 114: 679–688.

Borja, A., Franco, J., Pérez, V. (2000) A Marine Biotic Index to Establish the Ecological Quality of Soft-Bottom Benthos Within European Estuarine and Coastal Environments. Marine Pollution Bulletin 40: 1100–1114.

Chicco, D., Warrens, M.J., Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science e623: doi:10.7717/peerj-cs.623.

Rovira, L., Trobajo, R., Ibáñez, C. (2012) The use of diatom assemblages as ecological indicators in highly stratified estuaries and evaluation of existing diatom indices. Marine Pollution Bulletin 64: 500–511.

5.8 Key findings from the evaluation of adverse ecological effects

Summary

A paired-site type approach was taken to assess morphometric macroinvertebrate data, with comparisons of an upstream site and a site situated downstream of an ERS/s, or from comparisons of two upstream sites with a site situated downstream of an ERS/s. The closer sites could be situated around an ERS minimised the sources of natural variation from mesohabitat structure differences (Section 5.2.3). This approach also minimised the influence of sources of natural variation, such as the effect of altitude on water temperature and the effect of subsurface geology (Section 5.2.2). Streams with sandstone subsurface geology generally have coarser (metal-poor) sediments compared to streams with shale subsurface geology that have finer particle sized sediments. Limiting statistical analyses to collection periods 10 to 17, provided samples collected under more typical flowing water conditions when WWOs were more active and removed the effect of reduced or ceased flows along with stream drying that occurred during the severe drought period in 2019 prior to collection periods 10 to 17.

The tracking contaminants study (Section 4.4) in concert with the toxicity testing pilot studies (Section 4.5) documented ammonia to be a contaminant of potential concern under conditions of insufficient dilution of WWO spilled influent in Vineyard Creek. This learning enabled interpretation of other paired site assessments of morphometric macroinvertebrate data together with companion field inspection information (Section 4.7) to identify discernible ongoing (press) adverse ecological effects at downstream sites of Blacktown, Kittys, Girraween and Frenchs creeks. These five instances of adverse ecological effect represent 22% of the broader suite of freshwater urban stream sites studied under the WWOM. Among these instances three low dilution settings were evident where a lack of sufficient dilution WWO spilled volume occurred in the receiving waters.

1) too much overflow volume from multiple ERSs discharging to the same point of a stream reach as in the case of Vineyard Creek
2) too much combined overflow volume from a number of ERSs that discharge at spatially differing points of a small urban stream as in the case of Girraween Creek
3) too much overflow volume from a single ERS is discharged into a small urban stream reach as is the case in Kittys Creek

These instances are pictorially displayed in Figure 8-7.

Multivariate regression modelling of the morphometric macroinvertebrate dataset against assembled metadata variables did not encapsulate a predictive capacity to enable a viable modelling approach as a future option to assist in prioritising ERS sites for ranking within the risk prioritisation methodology based upon morphometric macroinvertebrate taxonomic assemblage data. The same outcome was apparent from more complex statistical analyses employed in analyses in Sections 5.6 and 5.7 of eDNA metabarcoding datasets. As predictive capability based upon variables related to WWOs was unable to be developed with machine learning models for either freshwater or estuarine sites. Predictive modelling is

not proposed as an option to inform the risk prioritisation methodology going forward based upon metadata assembled under this study.

Implementing a paired-site approach, as demonstrated with the morphometric indicator, would be cost-prohibitive for assessing about 660 ERS (from the broader 3000 ERS) based on extrapolation above morphometric findings. Another consideration against this approach was illustrated under the study of the Darling Mills Creek system where upstream and downstream sites could not be positioned spatially close (within a few hundred metres) around the target ERS, as the ERS was situated at a stream junction with 10 km of additional stream length on one branch (Section 5.3.7). The same issue would be faced if a paired site approach was attempted based upon DNA derived assemblage data. A desktop assessment together with walks as outlined in Sections 4.8 and 8.2.3 would enable potential low ammonia dilution stream settings to be identified. Then ERSs situated in those settings could be progressively assessed with deployed instrumentation to measure ammonia as outlined in Sections 4.8 and 8.2.3.

It would be more cost-effective to measure ammonia using deployed instrumentation and evaluate those results against the ANZG (2018) default guideline value (DGV), to establish ERS locations where there is potential for adverse ecological effects. A direct measure of ammonia could then become an input into a revised risk prioritisation methodology.

5.9 Recommendations to explore application of DNA-based taxonomy with paired-site assessments under the SWAM program

In 2023, a review of the Sydney Water Aquatic Monitoring (SWAM) program endorsed the biotic index SIGNAL-SG as fit for purpose and suitable for assessments under a paired-site design. That review also recommended exploring genomic approaches. An external peer-panel conducted this review based on input from both Sydney Water and the NSW EPA.

The paired-site design is currently employed under SWAM program to assess continuous discharges from 14 wastewater treatment plants (also described as Water Resource Recovery Facilities). This design has been enacted since 1995 under the SWAM and preceding programs (Environmental Indicator Monitoring Program, Sewage Treatment System Impact Monitoring Program) to fulfil Environment Protection Licence requirement M5.1. Under the SWAM program 40 sites are situated upstream and downstream of the continuous discharge points on receiving streams along with paired-sites upstream and downstream of the confluence of these streams with the Nepean River.

The following discussion and recommendations outline potential pilot studies to explore genomic approaches to evaluate if equivalent or more sensitive assessments may be provided by taxonomic assemblages obtained with DNA methods. The below proposed pilot studies build upon WWOM learnings to address the 2023 SWAM review recommendation.

5.9.1 'Renovate' approach to calculate SIGNAL-SG sample scores

As outlined in Section 5.5, biological indices derived from taxonomic lists provide convenient ways to summarise assemblage data, as illustrated by the application of the biotic index SIGNAL-SG based on morphometric data from freshwater urban streams (Section 5.3). Inputting taxonomic data from metabarcoding of DNA samples provides an alternative basis for raising SIGNAL-SG scores from a sample. This approach was applied with community-DNA (the identification of species from mixtures of whole animals in bulk-samples) (Section 5.5.1), which yielded SIGNAL-SG scores that correlated strongly with morphometrically derived scores. This represented a 'renovate' approach as described by Pawlowski et al. (2021).

Whereas, comparison of morphometric data with metabarcoding of eDNA from environmental sediment samples (Section 5.5.3) returned low yields (18% and 30%, respectively) of genera with assigned SIGNAL-SG biotic index grades under taxonomy assigned from eDNA amplicons of mitochondrial 16S and COI. This low yield potentially influenced the lack of detection of impaired downstream assemblages in ecological control charts (Section 5.5.3).This outcome clearly indicates that at present, for Sydney Water, this is not a viable option to implement a 'renovate' approach without an improvement in yield of genus level taxa that have been assigned SIGNAL-SG grades. To improve this yield, there are two options.

1) an enhanced primer pair under metabarcode sequencing
2) metagenomics also known as shotgun sequencing.

These are discussed further below.

Bioinformatic assignment of taxonomy to sequenced eDNA data (Sections 5.5.3 and 5.6) was not done for those taxa within the Aquatic Invertebrates of Australia (AIA) DNA database hosted in the

Barcode Of Life (BOLD) database. The WWOM invested to infill taxonomy gaps in the AIA DNA database as described in Section 5.5.2 for COI and mitochondrial 16S, as this gap issue has been commented upon frequently in scientific literature. Given this work, another validation option would be to reconduct bioinformatic assignment of taxonomy using the AIA DNA database of sequencing amplicons (primer pairs) COI and mitochondrial 16S. After this bioinformatic taxonomy assignment step, the assessment conducted in Section 5.5.3 could then revisit raising SIGNAL-SG sample scores from metabarcoded eDNA dataset.

To understand remaining gaps in molecular taxonomy, a gap analysis of DNA barcoded taxa hosted in the AIA and the common morphometric taxa of the SWAM program should be undertaken to enable a further round of DNA barcoding. This would then allow a more effective implementation of the 'renovate' approach into the long-term regulatory SWAM program.

A further pilot study is recommended to evaluate more optimised COI primer-pair (amplicons) that better target macroinvertebrate taxa under a 'renovate' approach. Section 5.5.4 outlines this recommended pilot study with an optimised primer-pair for evaluation with both community-DNA and eDNA. This recommendation is further outlined in the '*Renovate approach community-DNA and water/sediment eDNA studies*' recommendations box below.

5.9.2 'Rebuild' approach

As highlighted by Pawlowski et al. (2021), a 'rebuild' approach would involve constructing new bioindicators. This would include a biotic index based on DNA. The bioinformatic processed eDNA datasets created as part of the Macquarie University collaboration, from six amplicons (primer pairs) and metabarcoding (as outlined in Sections 5.5.4 and 5.6), provided taxonomic groups beyond the metazoan macroinvertebrates to commence exploring a 'rebuild' approach upon for freshwater taxa of the Sydney region.

Exploration of these datasets can be conducted with minimal cost, as only statistical analysis time is required. This statistical exploration could take two possible approaches. The BVStep routine of PRIMER (Clarke and Warwick, 1998) and the R package mvabund (Wang et al., 2021). This latter routine may better handle the large number of rows in DNA datasets, especially when there are many absences, that is, taxa which were not detected in a sample. An effective search strategy within these datasets maybe to constrain the set of DNA samples at urban streams with identified adverse ecological effects observed with the morphometric macroinvertebrate data (Sections 5.3.1 Vineyard Creek, 5.3.4 Kittys Creek, 5.3.6 Frenchs Creek and 5.3.5 Girraween Creek). These streams represent type 1, 2 or 3 examples of low dilution situations where adverse ecological effects were apparent (Figure 8-7). The adverse ecological effect detected in morphometric data at the Vineyard downstream site has also been confirmed under toxicity testing (Section 4.5.1). This represents what could be described as 'a gold standard' to compare data patterns from statistical analysis of sets of DNA taxa.

Viewing outcomes across these paired-site assessments under these analysis methods may establish particular taxonomic groups from the primer pairs (amplicons) across the bacteria, diatoms and eukaryotes to base paired-site assessments under the SWAM program.

BVStep search approach

Once the BVStep routine has identified a subset of taxa, those taxa can then be used as the basis for further multivariate assessment to determine if they illustrate similar/equivalent patterns to the

morphometric macroinvertebrate data from the above suggested four urban streams (Section 5.3). Classification, nMDS ordination and shade plots from multivariate analysis would then be viewed against the same plots of the morphometric data to inspect for similar data patterns. To assess for differences in taxonomic assemblages between upstream and downstream sites of each creek, ANOSIM R values of pairwise comparisons would be a useful statistic, as large R values (close to unity) are indicative of complete separation of the groups, whereas small R values (close to 0) imply little or no segregation (Clarke et al., 2014). Differences in sample number by collection period existed with two per site for the morphometric dataset. While in the DNA datasets, four replicates were available, however these different replicate numbers would not influence the R value. As Clarke et al. (2014) describes, the R value itself is not unduly affected by the number of replicates in the two groups being compared; but it is in stark contrast to its statistical significance, which is dominated by the group sizes (for larger number numbers of replicates, R values near zero could still be deemed significant, and conversely, few replicated could lead to R values to unity being classified as significant). If ANOSIM analysis of sets of DNA taxa return R values that are larger than those from the morphometric pairwise comparisons of upstream and downstream sites, this would provide confidence that an adequate subset of taxa had been identified.

mvabund proposed search approach

The mvabund approach would use the same experimental approach as described above for the for the BVStep, the difference being in statistical analysis. In contrast to the BVStep approach, mvabund does not use a similarity matrix derived from a distance-based measure (for example, Bray-Curtis). Such metrics can be problematic with e-DNA data (metabarcoding and metagenomic) due to the large number of taxa (tens of thousands in the case of bacteria), and the naturally high variability. Hence, traditional multivariate approaches (for example, PERMANOVA, ANOSIM and CCA) are less likely to detect a treatment effect (for example, upstream/downstream differences). In addition, many taxa in eDNA data are rare, and hence have little variance, while common taxa may have large variance. This mean-variance property, which reflects the natural distribution of the taxa, is not accounted for under traditional distance-based approaches. Indeed, both approaches were applied in Section 5.6, however, in this case, the analysis was performed on all the data collected. Here we propose a more focused approach, as described above, which would test for differences in communities between treatments (for example, upstream/downstream and overtime), as indicated by a significant ($P<0.05$) Likelihood Ratio Statistic (LRT) (Wu et al., 2012). Furthermore, the key taxa which are driving these differences, as well as those associated with any measured environment gradients, can be extracted, identifying potential indicators associated with particular states (for example, upstream and downstream)

5.9.3 Metagenomics

Under morphological identification count data are obtained from detected taxa, which provides an abundance component within raised SIGNAL-SG scores for site samples. Abundance information was incorporated within analyses of the demonstration paper (Besley and Chessman, 2008) of application of the Sydney version of SIGNAL (Chessman et al., 2007). This approach has been the basis of analysis under the SWAM and preceding programs since 1995.

An advantage of DNA metabarcoding is removal of microscopic identification, however, metabarcoding has limitations. Bell et al. (2021) outlines some of these issues such as species that produce different DNA isolation yields and vary in organellar or ribosomal genome copy

number, which could lead to biases in DNA quantity going into PCRs (Kembel et al., 2012; Lamb et al., 2019; Pawluczyk et al., 2015). Primers for PCR may differ in their binding efficiencies to different species (Pompanon et al., 2012), or polymerases may be biased toward different nucleotide composition (Nichols et al., 2018), leading to PCR biases. To nullify these influences in statistical analyses metabarcoded data are analysed on a presence-absence or relative abundance.

An alternative approach is metagenomics, which allows an inclusion of a measure of quantification into both community-DNA and environmental DNA. It should be noted that this measure of quantification is not directly comparable to abundance counts, as the copy neighbour of genes and the number of cells varies greatly among taxa but offers potential improvement for more sensitive assessment than from presence-absence data which is a limitation of metabarcoding. Metagenomic approaches circumvent PCR and do not target particular gene regions, eliminating amplification bias and potentially reducing copy number bias (Bista et al., 2018). With the targeted genes being extracted subsequent to sequencing and bioinformatics. In essence, one can 'fish out' the genes and loci of interest from the sequenced pool of DNA. The quantitative accuracy (for example, Morgan et al., 2010) and species detection ability (Ranjan et al., 2016) of metagenomics has been investigated for prokaryote communities and more recently eukaryotic communities (Bista et al., 2018; Garrido-Sanz et al., 2020; Gómez-Rodríguez et al., 2015; Ji et al., 2019; Tang et al., 2014, 2015) (as outlined in Bell et al., 2021). Recently, the power of this approach has been demonstrated with the bulk extraction and metagenomic sequencing of terrestrial vertebrates.

Hence, metagenomics has the potential to provide quantitative taxonomic data, attenuating many of the issues associated with PCR, which forms the basis of metabarcoding. Metagenomics provides an alternative pathway under the 'renovate' approach to derive SIGNAL-SG sample scores with a component of quantification from DNA. Metagenomics has recently been demonstrated using terrestrial invertebrates in a manner analogous to the metagenomic analysis of aquatic benthos (Li et al., 2024).

5.9.4 Potential journal papers from WWOM ecology studies

Three journal papers are envisaged. The timeline of the WWOM project precluded completion of these papers before finalisation of this report.

1. Technical paper on high quality of sequencing outputs, benefit of reduced cost in future metabarcoding. This paper looks at the variability associated with different components of an eDNA biomonitoring program, specifically, technical variability (PCRs and sequencing runs) and natural variability, how biological communities vary at site, catchments and temporal scales. The paper will not only highlight the quality and reproducibility to the eDNA collected from a technical perspective, but also, helps assist in making informed future decisions about sample size and the spatial requirements of future studies
2. Freshwater paired-site comparison of morphometric macroinvertebrate data (Sections 5.3.1 Vineyard Creek, 5.3.4 Kittys Creek, 5.3.6 Frenchs Creek and 5.3.5 Girraween Creek) with eDNA data. This study will use mvabund to demonstrate how eDNA derived data (from multiple genes) can be used to examine upstream and downstream difference in communities across time

3. Estuarine paper. This paper is framed as an ecological assessment of Sydney's estuaries. It will encompass a multi-gene approach which not only compares the estuaries, but also, will show how different taxonomic communities respond to environmental variables and over time to rainfall events. Importantly, the paper will highlight the influence of catchment modification (for example, density of roads) on residing estuarine communities. The study will provide a template for estuarine eDNA studies within an urbanised landscape, as it highlights the need to not only capture traditional variables (for example, salinity and grainsize) but also, catchment variables, to better understand what drives benthic community structure

Recommendations

Outcomes of the WWOM preclude endorsement for adoption of Biomonitoring 2.0 under the SWAM, without further comparative pilot studies that illustrate at least equivalent outcomes to those from morphometric identification-based assessments. To explore application of DNA under 'renovate' and 'rebuild' approaches with a pair-site design the following recommendations are made.

Infilling taxonomic gaps within DNA databases

A gap analysis should be conducted between DNA barcoded taxa hosted in the AIA and the common taxa of the SWAM program for macroinvertebrate size class taxa with SIGNAL-SG grades. A further round of DNA barcoding is proposed to infill those defined gaps to best explore implementation of the 'renovate' approach under this long-term regulatory project with either COI or mitochondrial 16S.

Metagenomics can be applied to samples of community-DNA and eDNA under both the 'renovate' and 'rebuild' approaches to afford a measure of quantitation.

'Renovate' approach community-DNA and water/sediment environmental-DNA studies

The following two options are recommended for future pilot studies to evaluate the COI primer pair raised by Leese et al. (2020) against Sydney region samples for the 'renovate' approach as outlined under Section 5.5.4.

1. Conduct sequencing of reserve aliquots of previously extracted eDNA from previously collected WWOM sediment samples and compare to community-DNA extracted from previously collected WWOM bulk samples of hand-picked macroinvertebrates. This would allow evaluation of the morphometric taxonomy dataset with molecular taxonomy datasets from sediment and community-DNA
2. Collect new water and sediment samples at the same time as hand-picked macroinvertebrates are collected for the Sydney Water Aquatic Monitoring program. This would allow evaluation of the morphometric taxonomy dataset with eDNA datasets from water and sediment, and with community-DNA

From a cost perspective, option 1 would be relatively cheaper, as both traditional hand-picked macroinvertebrate samples along with companion sediment samples have already been collected, and e-DNA has been extracted from sediment samples. Comparison to the

existing WWOM sequenced COI primer dataset is also available. Whereas that would have an additional cost under option 2.

‘Rebuild’ approach environmental-DNA

To explore a ‘rebuild’ approach, taxonomic groups beyond the metazoan macroinvertebrate size class could be explored for sets of taxa with the BVSTEP routine of PRIMER from paired-site assessments of molecular metabarcoding datasets. Those sets of taxa could then be compared against morphometric data of the sites of the six creeks (Sections: 5.3.1 Vineyard Creek, 5.3.3 Avondale Creek, 5.3.4 Kittys Creek, 5.3.5 Girraween Creek, 5.3.6 Frenchs Creek, and 5.3.8 Blacktown Creek) with apparent adverse ecological effects. Viewing outcomes across these six paired-site assessments may establish particular taxonomic groups from the primer pairs (amplicons) across the bacteria, diatoms and eukaryotes to base paired site assessments under the SWAM program.

Measure of quantification from metagenomics in community-DNA and environmental-DNA

Metagenomic sequenced data could be obtained from an additional sequencing run as part of pilot studies recommended above under the ‘renovate’ approach. This would enable an assessment of quantification information returned from the metagenomic data, removing the biases associated with PCR, and to that of morphometric identification to establish if this assists in raising better correlated SIGNAL-SG scores from site samples to implement Biomonitoring 2.0 under the SWAM.

References

Bell, K.L., Petit III, R.A., Cutler, A., Dobbs, E.K., Macpherson, J.M., Read, T.D., Burgess, K.S., Brosi, B.J. 2021. Comparing whole-genome shotgun sequencing and DNA metabarcoding approaches for species identification and quantification of pollen species mixtures. Ecology and Evolution. 11, ece3.8281. https://doi.org/10.1002/ece3.8281

Besley, C.H., Chessman, B.C. 2008. Rapid biological assessment charts the recovery of stream macroinvertebrate assemblages after sewage discharges cease. Ecol. Indic. 8: 625–638.

Bista, I., Carvalho, G.R., Tang, M., Walsh, K., Zhou, X., Hajibabaei, M., Shokralla, S., Seymour, M, Bradley, D., Liu, S., Christmas, M., Creer, S. 2018. Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. Mol. Ecol. Resour. 18(5), 1020–1034. 10.1111/1755-0998.12888

Chessman, B.C., Williams, S.A., Besley, C.H. 2007. Bioassessment of streams with macroinvertebrates: effect of sampled habitat and taxonomic resolution. J. North Am. Benthol. Soc. 26, 546-565

Clarke, K.R., Warwick, R.M., 1998. Quantifying structural redundancy in ecological communities. Oecologia 113, 278–289.

Clarke, K.R., Gorley, R.N., Somerfield, P.J., Warwick, R.M., 2014. Change in Marine Communities: An Approach to Statistical Analysis and Interpretation, 3rd ed. PRIMER-E, Plymouth, U.K.

Garrido-Sanz, L., Senar, M.À., Piñol, J. 2020. Estimation of the relative abundance of species in artificial mixtures of insects using low-coverage shotgun metagenomics. Metabarcoding Metagenomics 4, e48281. 10.3897/mbmg.4.48281

Gómez-Rodríguez, C., Crampton-Platt, A., Timmermans, M.J.T.N., Baselga, A., Vogler, A.P. 2015. Validating the power of mitochondrial metagenomics for community ecology and phylogenetics of complex assemblages. Methods in Ecol. Evol. 6(8), 883–894. 10.1111/2041-210X.12376

Ji, Y., Huotari, T., Roslin, T., Schmidt, N.M., Wang, J., Yu, D.W., Ovaskainen, O.2019. SPIKEPIPE: A metagenomic pipeline for the accurate quantification of eukaryotic species occurrences and intraspecific abundance change using DNA barcodes or mitogenomes. Mol. Ecol. Resour. 20(1), 256–267. 10.1111/1755-0998.13057

Kembel, S.W., Wu, M., Eisen, J.A., Green, J.L. 2012. Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. PLoS Comput. Biol. 8(10), e1002743. 10.1371/journal.pcbi.1002743

Lamb, P.D., Hunter, E., Pinnegar, J.K., Creer, S., Davies, R.G., Taylor, M.I. 2019. How quantitative is metabarcoding: A meta-analytical approach. Mol. Ecol. 28(2), 420–430. 10.1111/mec.14920

Li, Y., Devenish, C., Tosa, M., Luo, M., Bell, D., Lesmeister, D., Greenfield, P., Pichler, M., Levi, T., Yu. D.S. 2024. Combining environmental DNA and remote sensing for efficient, fine-scale mapping of arthropod biodiversity. Philos. Trans R Soc. Lond. B. Biol. Sci. 379, 1904: 20230123

Morgan, J.L., Darling, A.E., Eisen, J.A. 2010. Metagenomic sequencing of an in vitro-simulated microbial community. PLoS One. 5(4), e10209. 10.1371/journal.pone.0010209

Nichols, R.V., Vollmers, C., Newsom, L.A., Wang, Y., Heintzman, P.D., Leighton, M.K., Green, R.E., Shapiro, B. 2018. Minimizing polymerase biases in metabarcoding. Mol. Ecol. Resour. 18(5), 927–939. 10.1111/1755-0998.12895

Pawlowski, J. Bonin, A., Boyer, F., Cordier, T., Taberlet, P. 2021. Environmental DNA for biomonitoring. Mol. Ecol. 30, 1-6. https://doi.org/10.1111/mec.16023

Pawluczyk, M., Weiss, J., Links, M.G., Egaña Aranguren, M., Wilkinson, M.D., & Egea-Cortines, M. 2015. Quantitative evaluation of bias in PCR amplification and next-generation sequencing derived from metabarcoding samples. Anal. Bioanal. Chem. 407(7), 1841–1848. 10.1007/s00216-014-8435-y

Pompanon, F., Deagle, B.E., Symondson, W.O.C., Brown, D.S., Jarman, S.N., Taberlet, P. 2012. Who is eating what: diet assessment using next generation sequencing. Mol. Ecol. 21(8), 1931–1950. 10.1111/j.1365-294X.2011.05403.x

Ranjan, R., Rani, A., Metwally, A., McGee, H.S., Perkins, D.L. 2016. Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. Biochem. Biophy. Res. Commun. 469(4), 967–977. 10.1016/j.bbrc.2015.12.083

Tang, M., Tan, M., Meng, G., Yang, S., Su, X.U., Liu, S., Song, W., Li, Y., Wu, Q., Zhang, A., Zhou, X. 2014. Multiplex sequencing of pooled mitochondrial genomes-a crucial step toward biodiversity analysis using mito-metagenomics. Nucleic Acids Res. 42(22), e166. 10.1093/nar/gku917

6 WWOM outcomes relevant to dry- weather sewer overflow investigations

Environmental performance of dry-weather sewer overflows (DWSOs) is managed under the Sydney Water Enterprise Initiative T2.4 Improving Wastewater Quality via effective source control. DWSOs include sewer influent spills from ERSs due to a choke in the sewerage network or from breaks within this network.

In response to DWSOs, Sydney Water utilises a 24-hour rapid response environmental assessment team within the Field Sampling and Testing (FST) team of Laboratory Services. Total ammonia is currently used as the primary indicator for sewage contamination in DWSO investigations. To understand spatial extent of a DWSO into a receiving waterway both visual and olfactory observations are used in conjunction with field readings for dissolved oxygen (DO), conductivity (Cond) and pH at several sites located progressively distant from the DWSO. Water samples are also collected and analysed for faecal coliforms and enterococci. HFMGs of the MST approach are employed for post clean-up assessments when bacterial indicators remain elevated to confirm if a human faecal source remains present.

Findings of the WWOM provide support for the inclusion of ammonia, chemical tracers, and HFMGs as appropriate indicator tools within dry-weather leakage investigations as outlined in the following Sections 6.1, 6.2 and 6.3, respectively.

6.1 WWOM findings support measuring ammonia to determine risk of adverse ecological effects from DWSOs

The use of ammonia as an appropriate indicator in DWSO investigations was supported by the WWOM studies. Our studies found that analysis of influent within the Vineyard Creek sewer carrier during dry weather, measured ammonia as exceeding the ANZG (2018) guideline value for 95% species protection by 84x (Section 4.4). This indicated that DWSOs have the potential for adverse ecological risk. Under wet-weather rainwater ingress conditions, water samples collected from the influent of four sewer carriers still exceeded the ANZG (2018) guideline values but with decreased ammonia measurements of 1.5 to 13x the guideline values (Section 4.4). Further support was provided under companion toxicity testing (as outlined in Section 4.5.1) where ammonia concentration results evaluated against the ANZG (2018) guideline value showed reducing risk of adverse ecological effect under increasing dilution.

Key learnings relevant to DWSO investigations

Adopting the current dry-weather overflow abatement program approach to measure ammonia as a primary line of evidence seems well justified by these documented observations of the WWOM.

In addition, development of in-house laboratory capability with animal associated faecal marker genes (Section 3.6) would also help interrogate the enterococci concentrations detected in post clean-up samples for potential animal faecal contamination sources. For example, these tests could be employed when waterfowl are observed within the receiving waterbody, or when farming activities with cattle, horses, chickens, or pigs are observed nearby the waterway.

Development of in-house laboratory capability with animal associated faecal marker genes (Section 3.6) has commenced to support DWSO investigations.

6.2 Chemical tracers and application to dry-weather leakage investigations

As outlined in Section 4.2, Birch et al. (2015) investigated estuarine waters from 30 sites adjacent to stormwater outlets across the entire Sydney estuary after a dry period. They found the presence of eight (out of 59 tested) pharmaceuticals and the artificial sweetener acesulfame in water from all parts of the Sydney estuary. Birch et al. (2015) concluded that the presence of these chemicals indicated that untreated sewer influent is leaking from the sewerage system into the stormwater network that flows to the estuaries.

Outcomes of the pilot study outlined in Section 4.3, enabled identification of the pharmaceuticals and personal care products that were applied to the tracking contaminants study outlined in Section 4.4. The tracking contaminants study recommended that the suite of eight organic chemicals (acetaminophen, ibuprofen, metformin, sucralose, theobromine and three benzotriazoles) should be used together as a line of evidence in future WWOA investigations as markers (tracers) for the presence of sewage contamination up to two days after an ERS spill, based on half-life considerations. This recommendation was influenced by the lack of uniformity of detection of these chemicals across receiving water sites after WWO spills. This in turn suggests that under much less diluted dry-weather stream baseflow conditions, these chemical tracers of sewage contamination may occur in the receiving waters at relatively higher concentrations (compared to wet-weather observations) to allow detection of dry-weather leakage. If these chemical tracers are detected this would indicate very recent, if not ongoing, sewage contamination due to the short half-life of these tracers.

Dry-weather detections in the water column of these chemical tracers is supported by the outcomes of the presence-absence water column study using passive samplers (Table 6-1). That study analysed the same eight chemical tracers. Passive samplers were deployed on a four-weekly cycle across 32 urban freshwater sites, upstream of ERS, during the severe drought period April to early September 2019 with five retrievals (for the months of April, May, June, July and

August) from each stream. A total of 152 passive samplers were retrieved from the 160 deployed. The eight passive samplers not recovered were due to being dry (out of the water) or insufficiently covered in water due to the stream drying as illustrated by Figure 5-2B. Dry or partly dried passive samplers were invalid for laboratory processing. Average rainfall at Sydney Olympic Park between April to August is 367 mm, while during the study period for these same months in 2019 a total of less than half of this (156 mm) was recorded for the nearby Vineyard Creek catchment. A similar trend was observed in catchments further west in Sydney. For example, 116 mm of rainfall was recorded for the Blacktown Creek catchment against an average of 328 mm for nearby Prospect Reservoir. During this drought period with reduced rainfall, all three benzotriazoles were detected in all 152 passive samplers, with sucralose detected in 99% of passive samples (Table 6-1). While theobromine, ibuprofen and metformin were detected in 63%, 49%, and 22% of passive samplers respectively. Acetaminophen was detected in 80% of passive samplers across these five-months (Table 6-1) and provides support for Birch et al. (2015) widespread study detections of this chemical in estuarine waters during a dry-weather period. Results for individual 32 upstream creek sites are displayed in Table 6-1 and were relatively consistent across sites that suggested relatively persistent dry-weather leakage of sewage contamination. This leakage can come from both the private and Sydney Water systems.

The documented passive sampler results from a period of below average rainfall period provided a counterfactual argument that the use of passive sampling to detect overflows from ungauged ERSs was in fact a false premise. However, successful detection in the water column of this suite of eight chemical tracers (Section 4.4, Besley et al., 2023) illustrates the applicability for future dry-weather leakage investigations when active leakage sources are suspected.

Recommendation

The suite of eight organic chemicals (acetaminophen, ibuprofen, metformin, sucralose, theobromine and three benzotriazoles) provide a line of evidence as markers (tracers) for the presence of recent sewage contamination as these chemicals have short half-lives of up to two days in the water column.

Hence detection of these chemicals in the water column would indicate recent/active sewage contamination and illustrates the applicability of this suite of chemical markers (tracers) for DWSO investigations.

Table 6-1: Percentage detection of chemical tracers from contiguous four weekly deployments of POCIS Oasis HLB passive samplers deployed at 32 upstream creek receiving water sites from April to August 2019

Upstream creek sites	# passive samplers	Acetaminophen	Ibuprofen	Metformin	Sucralose	Theobromine	1H-Benzotriazole	4-Methyl1H-benzotriazole	5-Methyl1H-benzotriazole
	% detection								
Cowan	5	40	40	0	100	40	100	100	100
Cockle	5	80	40	20	100	60	100	100	100
Hornsby	5	100	80	20	100	80	100	100	100
Carroll	5	80	100	40	100	80	100	100	100
Frenchs	4	50	50	0	100	25	100	100	100
Devlins 1	5	80	40	0	100	60	100	100	100
Devlins 2	4	75	25	0	100	50	100	100	100
Avondale	5	80	40	0	100	60	100	100	100
Congo	3	100	0	0	100	33	100	100	100
Rudder	5	40	20	0	100	20	100	100	100
Blackbutt	5	100	60	20	100	80	100	100	100
Kittys	5	80	100	60	100	80	100	100	100
Buffalo 1	5	100	60	0	100	80	100	100	100
Buffalo 2	5	40	0	0	100	40	100	100	100
Darling Mills system 1	5	80	80	40	100	80	100	100	100
Darling Mills system 2	5	80	80	60	100	60	100	100	100
Hunts 1	5	80	20	0	100	80	100	100	100
Hunts 2	5	80	40	20	100	80	100	100	100
The Ponds 2	5	100	100	80	100	100	100	100	100
The Ponds	5	80	80	0	100	40	100	100	100
Vineyard 1	5	100	40	40	100	80	100	100	100
Vineyard 2	5	80	20	20	100	40	100	100	100
Blacktown 1	5	100	60	80	100	80	100	100	100
Blacktown 2	5	100	20	20	100	60	100	100	100

Upstream creek sites	# passive samplers	Acetaminophen	Ibuprofen	Metformin	Sucralose	Theobromine	1H-Benzotriazole	4-Methyl1H-benzotriazole	5-Methyl1H-benzotriazole
Tributary Blacktown	5	100	0	20	100	40	100	100	100
Lalor	5	100	60	20	100	80	100	100	100
Girraween	5	80	40	0	100	60	100	100	100
Toongabbie 1	5	60	20	0	100	40	100	100	100
Toongabbie 2	2	100	0	50	100	50	100	100	100
Bunbury Curran	4	25	25	0	75	75	100	100	100
Brickmakers	5	60	80	20	100	40	100	100	100
Prospect	5	100	80	60	100	100	100	100	100
32 upstream sites	152	80	49	22	99	63	100	100	100

6.3 Using HFMG concentrations in sediment as a tool to assist with dry-weather leakage investigations

Water column samples are regarded as capturing the most recent event, whereas sediment samples are considered to incorporate a longer record of events. This assertion is supported by our recent study that reported sewage-associated marker genes may persist in the sediment longer than in the water column (Ahmed et al., 2019d cited in Ahmed et al., 2020a).

Results from pilot sub-study 5, (Section 3.2.5) showed that urban estuarine sediment can act as a reservoir of faecal indicator bacteria (FIB), HFMGs, animal marker genes (particularly avian-birds), and human adenovirus. Significant associations were determined between the HF183 human faecal-associated marker gene (HFMG) with WWOs occurrence in the preceding 6 and 12 weeks, while the HFMG CPQ_056 was also significant when tested against the preceding 12 weeks. This suggests CPQ_056 persists longer in estuarine sediment and may have a different decay rate to HF183, which may reflect that these two HFMGs belong to different groups of microorganisms. Although the lack of MST marker gene decay data in sediment makes it difficult to assess potential illness risks, the apparent differential persistence of these two HFMGs in sediment may be beneficial in distinguishing between sites that are subject to more recent or longer-term exposure to sewage contamination. HF183 appears to indicate more recent sewage contamination. The question arises, what threshold concentration represents dry-weather leakage from the Sydney Water system?

HFMG concentrations analysed from surface sediment collected from the Ponds Creek upstream site 2 (PONU2) were found to be above 100,000 gene copies per gram. This potentially illustrates concentration where it is likely that dry-weather leakage is from a Sydney Water source. This threshold is advocated to initiate investigations in urban streams. Analysis described below further supports the use of HFMG concentrations in the dry-weather investigation program.

Case Study: The Ponds Creek investigation (2020-21)

HFMGs were assessed in sediment samples from six collections in March, May, September November and December 2020 and April 2021 with three replicate samples taken at each site collection event. Results from the microbial source tracking HFMG duplex assay from Bacteroides HF183 and crAssphage CPQ_056 assessment of all freshwater sediment samples are presented in Figure 6-1.

Sections 4.4, 4.5.1, 4.5.2 and 5.3.1 document the atypical agglomeration of ERS and sewage exposure / contamination at the Vineyard Creek site. Hence, an expected outcome illustrated for these two HFMGs, was their elevated concentrations at the downstream site within Vineyard Creek (VIND) (Figure 6-1). The marginally elevated HFMG levels at the Vineyard upstream site 1 (VINU1) were also expected as repetitive network issues occurred throughout the timespan of the project. However, the similar elevated concentrations to those at the downstream site of Vineyard Creek were observed for The Ponds Creek upstream site 2 (PONU2) (Figure 6-1). This was surprising, as the PONU2 site was situated just above an ERS (Figure 6-2) with no other ERS higher up in this sewer catchment. Without sewage loading to this site (PONU2), this warranted further examination.

A field investigation initiated in late April 2021 identified a strong sewage smell and elevated ammonia concentrations within the reach of The Ponds Creek in question (PONU2). Leakage from a nearby sewer pipe was identified and repairs were conducted along with normal protocols of pump outs and flushing of the waterway. Near the breakage was a pile of wet wipes and toilet paper, which may suggest this leakage had gone undetected for some time as illustrated by the HFMG sediment results (Figure 6-1).

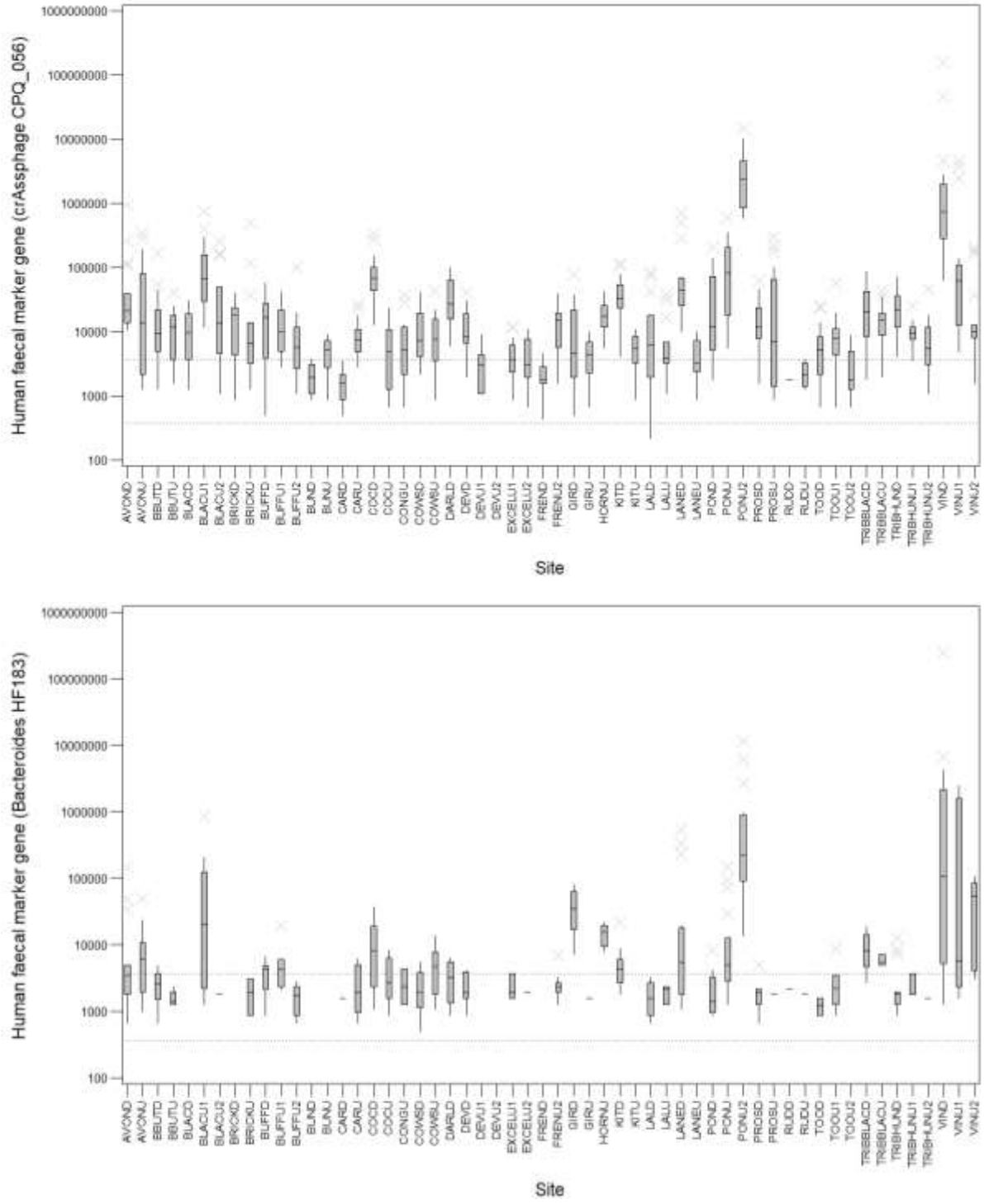


Figure 6-1: Microbial source tracking HFMG concentrations (gene copies per gram) in freshwater sediment of 56 urban stream sites

Based on six collections in March, May, September November and December 2020 and April 2021 with three replicate samples taken at each site collection event. Upper dotted line represents LOQ (3800 gene copies per gram) and lower dotted line LOD (380 gene copies per gram)

Follow up study at The Ponds Creek in 2023

A follow up study was commissioned in The Ponds Creek system between May and July 2023 to explore if:

- repair works had been successful
- decay in HFMGs in PONU2 sediment had occurred over this two-year gap in sample collection, as part of revisiting longer-term persistence of MST HFMGs in sediment
- similar water column HFMG concentrations occurred at three sites in The Ponds Creek

Figure 6-2 illustrates The Ponds Creek site location for this follow-up study. The site marked in the middle of this map was upstream site PONU2, while the site at the bottom of the map was the other upstream site PONU from the original WWOM study. A third upstream site PONU3 was added for this follow-up study.

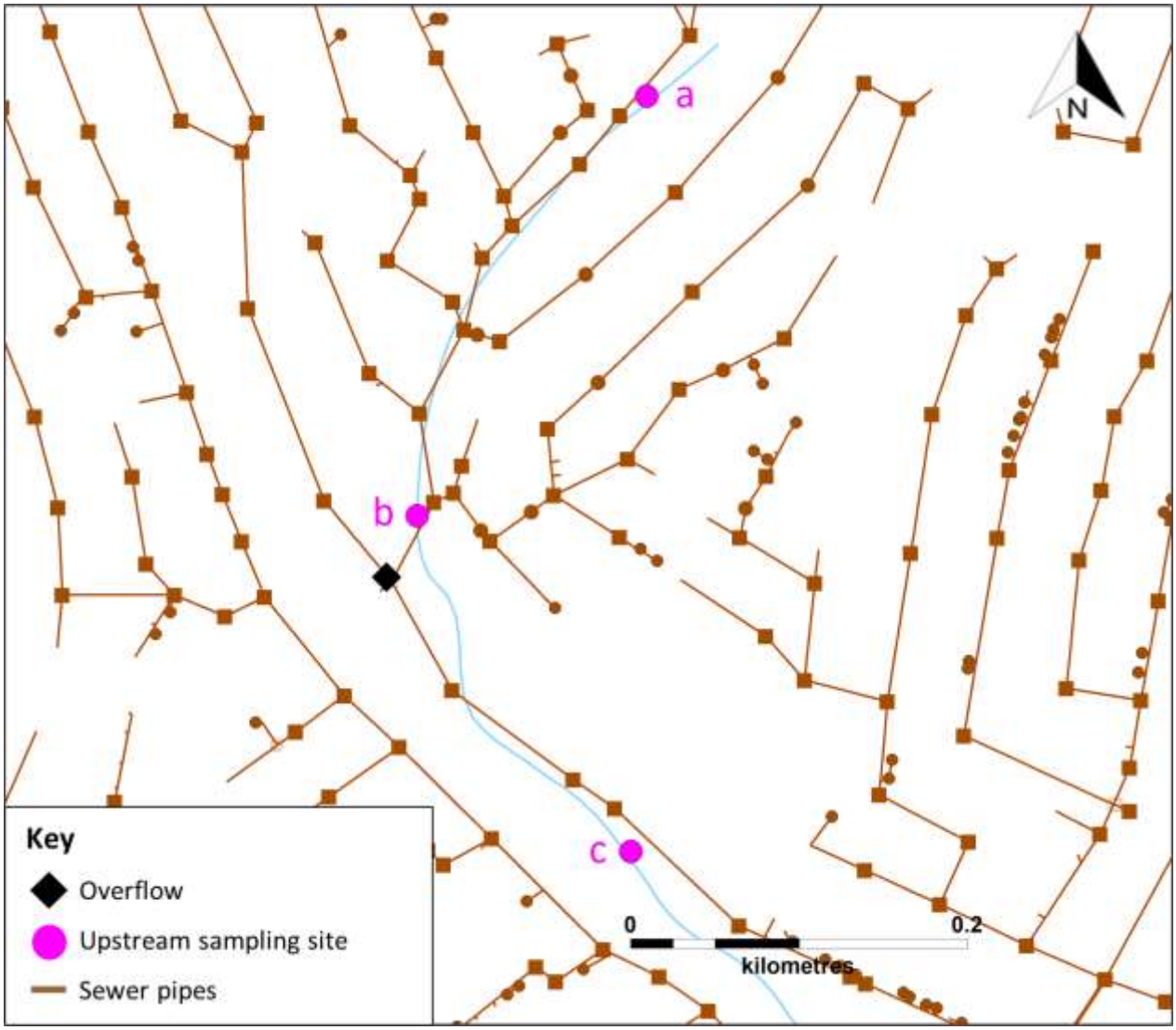


Figure 6-2: Ponds Creek case study location map
Sites: a = PONU3; b = PONU2; c =PONU

Both water and sediment samples were collected on each of four sample collection events (May 3, May 22, June 14, and July 27, 2023). Rainfall of note occurred on April 30, 2023, with 23 mm in the 24 h to 9 am and another 10 mm in the 24 h to 9 am on July 24, 2023; both these rainfalls events occurred three days before the first and fourth collection events (Figure 6-3). Prior to the first collection event more frequent and generally higher total daily rainfall occurred through April 2023 compared to each of the three following collection events (Figure 6-3).

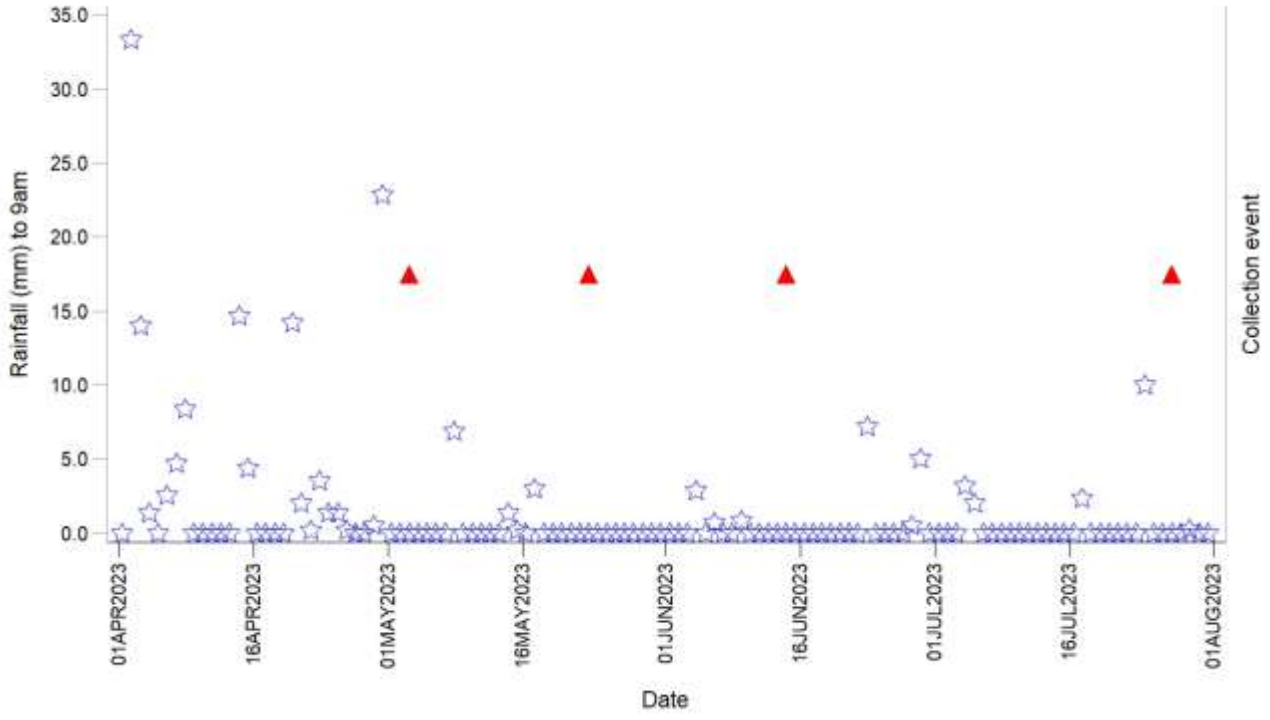


Figure 6-3: The Ponds Creek daily rainfall and four collection events

Stars = daily rainfall 24 h total (mm) to 9 am; Diamonds = collection events. Rainfall averaged from three nearest rain gauges (566037 - Ryde Pumping Station, 566081 - Carlingford Bowling Club, 567112 - North Parramatta).

Were repair works successful?

A visual comparison of the plots of HFMGs from the early sediment collection (March, May, September November and December 2020 and April 2021) (Figure 6-1) and of sediment samples collected for the 2023 study (Figure 6-4) illustrated an order of magnitude change in concentrations. The 2023 concentrations were generally less than 100,000 gene copies per gram (Figure 6-4) which were more typical of those detected at other urban sites in the early sediment collection (March, May, September November and December 2020 and April 2021) (Figure 6-1). This comparison of concentrations between these two plots clearly illustrated a change at PONU2 that suggested the repair had been successful.

Persistence of MST HFMGs in sediment

The dramatic drop in sediment concentrations of HFMGs at PONU2 described above also suggests HFMG concentrations quantified in 2023 were not from persistence of HF183 and CPQ_056 from 2021.

In the 2023 study, higher concentrations for both HFMGs from sediment samples were observed at the most downstream of these three sites (PONU) that was situated below an ERS (AGN 1372788, Figure 6-2). This ERS had a low modelled overflow spill frequency of 6 events and volume of 0.1 ML in 10 years. Curiously, these hydraulic modelled outcomes may suggest higher rainfall totals are required to detect activity from this ERS or the modelled outcomes may not be accurate.

Concentrations of HF183 were highest at each respective site for collection event four, that occurred three days after 10 mm of rainfall. A similar trend was evident for CPQ_056, but concentrations were documented for collection two and three and an increasing trend downstream with highest concentrations quantified of CPQ_056 from the most downstream site of PONU (Figure 6-4). More than double the rainfall occurred preceding event 1 (Figure 6-3) that may have flushed and diluted water column concentrations prior to the first collection event and possibly reduced settling into sediments. This may explain the relatively lower HFMG sediment concentrations measured from event 1 (Figure 6-4).

The least rainfall occurred within the 19 and 23 days before collection events two and three (Figure 6-3). The lack of HF183 detections from collection events two and three may suggest these periods were sufficient to see decay in freshwater sediments of this HFMG (Figure 6-4). The concentration documented from collection event four may reflect leakage from the private network that ponded in sediments higher up in the catchment and was washed into The Ponds Creek 2023 study reach with increased loading as the catchment area increased downstream (Figure 6-4).

The outcomes from the study of sediment at the PONU2 site in 2020-21, where a confirmed Sydney Water network issue was identified (and repaired) from field investigations, has defined HFMG sediment concentrations (from the duplex assay HF183 and CPQ_056 markers) to provide a threshold for future assessments (above 100,000 gene copies per gram). Establishing decay time frames for these two HFMGs would further assist in interpretation of the age of sewage contamination of sediment to make best use of this potential tool in assessing dry-weather leakage.

Water column HFMG concentrations at the three sites in The Ponds Creek (2023)

Figure 6-5 plots the HFMGs assessed in the water column, illustrating the elevated concentrations in samples collected after each of the two rainfall events from the three sites. The highest concentration in samples was observed in PONU samples collected downstream of the ERS. This water column testing also suggested the sewer carrier repair of the 2020-21 dry-weather leakage issue from the Sydney Water sewerage system was successful.

The water column HFMG concentrations observed after rainfall in 2023, together with either not detected or much lower concentrations documented for events two and three from a dry period, suggest there may be some private network faults that flow to the upper two sites PONU3 and PONU2 (Figure 6-5).

Water samples from events one and four, collected three days post rainfall, had concentrations about or above the 50th percentiles from Risk Based Thresholds (RBTs) of the QMRA (Section 3.3.2) that represented an average illness level of 0.032. Although these RBTs have been calculated from the study of estuarine sites in the Sydney region, they are perhaps more meaningful than comparisons to overseas literature. Hence, these Sydney specific RBT thresholds provide context for measured HFMG concentrations in the water column. This may be of use in prioritising sites for further evaluations such as field inspections. An example of this was provided in Section 3.2.2 (Ahmed et al., 2019a) with detection of persistent leakage of influent from a trunk main sewer carrier into the freshwater creek that flows into the estuary at Gymea Baths, based on HFMGs used in campaign sampling in dry weather.

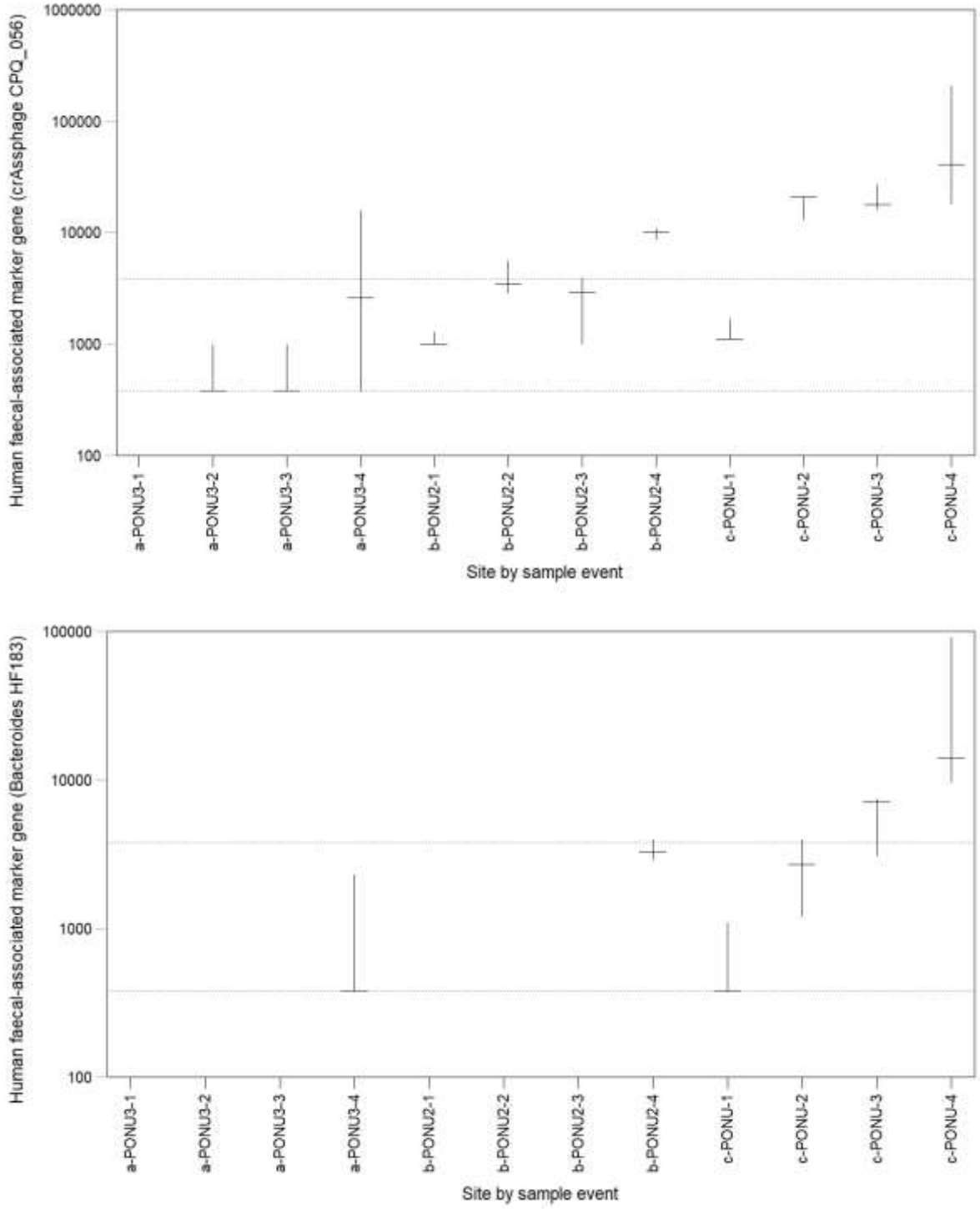


Figure 6-4: The Ponds Creek HFMG concentrations (gene copies per gram) from sediment samples across three sites and four collections

Sites: a = PONU3; b = PONU2; c =PONU. Events -1 = May 3; -2 = May 22; -3 = June 14, -4 = July 27. Upper dotted line represents LOQ (3800 gene copies per gram) and lower dotted line LOD (380 gene copies per gram). Samples with HFMG detection but not quantifiable were assigned LOD value.

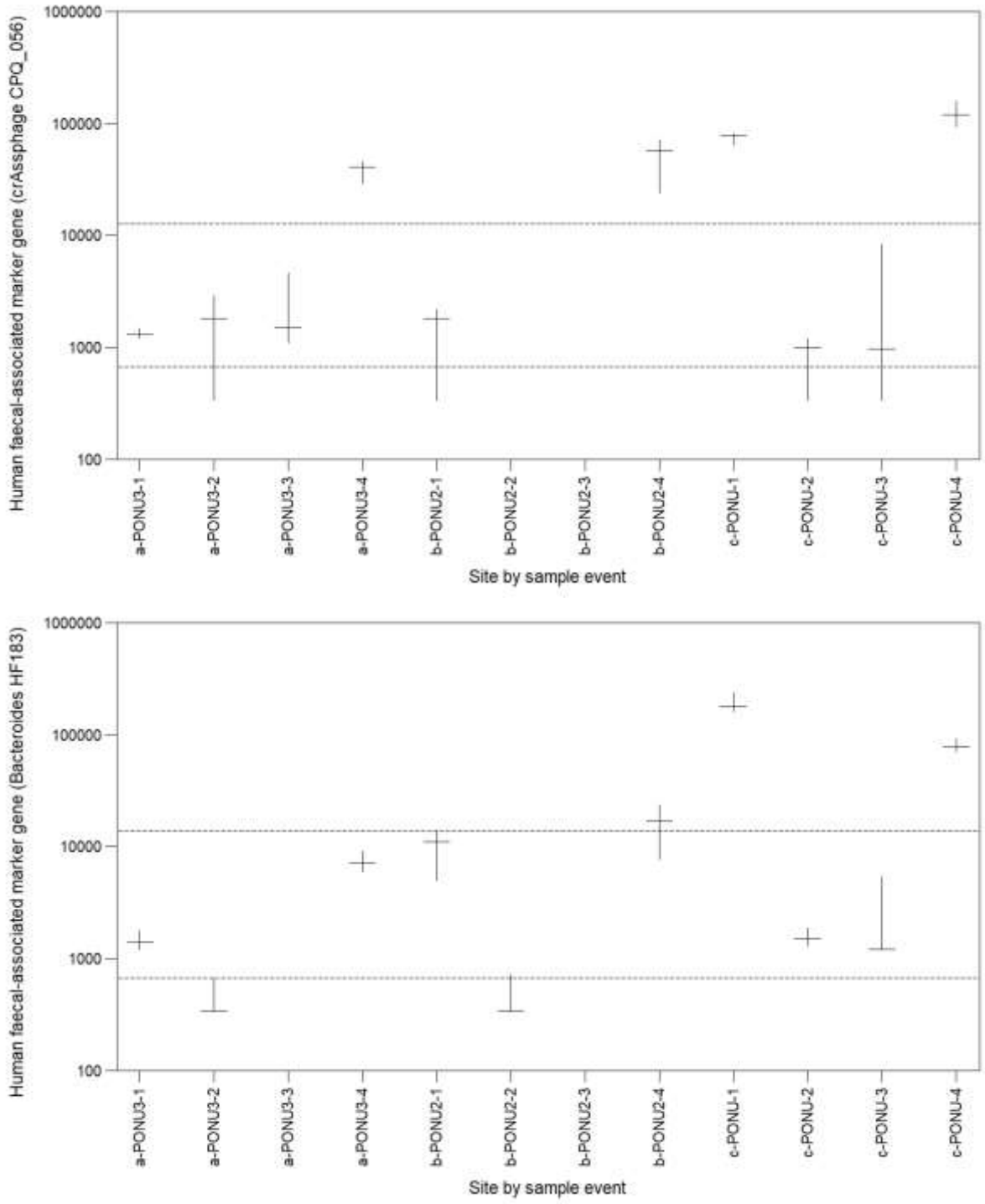


Figure 6-5: The Ponds Creek HFMG concentrations (gene copies per Litre) from water samples across three sites and four collections

Sites: a = PONU3; b = PONU2; c =PONU. Events -1 = May 3; -2 = May 22; -3 = June 14, -4 = July 27. Lower dotted line represents the LOQ (670 GC/L) and upper dotted line represents 50th percentiles from Risk Based Thresholds (3-day aged contamination) of the QMRA (Section 3.3.2). Samples with HFMG detection but not quantifiable were assigned half the LOQ

Summary and future studies

The Ponds Creek field investigations confirmed DWSO leakage from the Sydney Water sewerage system into this creek and developed our understanding of HFMG concentrations in sediment. Outcomes of this case study from The Ponds Creek has defined a threshold (above 100,000 gene copies per gram of sediment) for HFMG sediment concentrations (from the duplex assay HF183 and CPQ_056 markers) for future assessments of DWSO leakage from the Sydney Water sewerage system.

The follow-up study conducted in 2023 confirmed that repair work in 2021 was successful.

HF183 was observed to be less persistent than CPQ_056 during the 2023 study, which was also observed in sediment samples collected for WWOM in 2020-21. This persistence difference suggests that detection of HF183 may indicate more recent sewage contamination. Although, as discussed in Section 3.3.2, modifications to the existing adsorption extraction concentration method should be explored to determine if a more consistent capture of both viruses (CPQ_056) and bacteria (HF183) can be obtained simultaneously. Despite this consideration under The Ponds Creek case study, both HF183 and CPQ_056 markers returned quantified results well above the proposed 100,000 gene copies per gram threshold.

Further studies of HFMGs in sediment

Gathering more persistence data for the HF183 and CPQ_056 HFMGs would assist in understanding the decay of DNA from these bacterial and viral indicator microorganisms (respectively) in sediment. These data may develop an understanding of which factors such as porosity, grain size, clay content, oxygen conditions, reduced predation by protozoans, organic carbon, that may be influential in decay (Ahmed et al., 2020b). Learnings from those studies would further assist in application of this tool to assess sewage contamination.

References

Ahmed, W., Payyappat, S., Cassidy, M., Harrison, N., Besley, C., 2020a. Sewage-associated marker genes illustrate the impact of wet weather overflows and dry weather leakage in urban estuarine waters of Sydney, Australia. Sci. Total Environ. 705, 135390. https://doi.org/10.1016/j.scitotenv.2019.135390

Ahmed, W., Payyappat, S., Cassidy, M., Harrison, N., Marinoni, O., Besley, C., 2020b. Prevalence and abundance of traditional and host-associated fecal indicators in urban estuarine sediments: Potential implications for estuarine water quality monitoring. Water Res, 184, 116109. https://doi.org/10.1016/j.watres.2020.116109

ANZG (2018) Australian and New Zealand guidelines for fresh and marine water quality. Australian and New Zealand Governments and Australian state and territory governments, Canberra ACT, Australia. Available at: https://www.waterquality.gov.au/anz-guidelines

Birch, G.F., Drage, D.S., Thompson, K., Eaglesham, G., Mueller, J.F. 2015. Emerging contaminants (pharmaceuticals, personal care products, a food additive and pesticides) in waters of Sydney estuary, Australia. Mar. Pollut. Bull. 97, 56-66.

7 Monitoring tools assessed and employed in WWOM

Since the mid 1990s and the implementation of WaterPlan 21, Sydney Water has applied established tools for assessing our operating impacts on receiving waters. Monitoring tools developed in the last decade have not been widely adopted to date by Sydney Water. Environment Protection Licence requirement M5.1 obligates Sydney Water to conduct waterway health monitoring to assess the performance of the wastewater network and treatment facilties. This program is the Sydney Water Aquatic Monitoring (SWAM) Program (previously the Sewage Treatment System Impact Monitoring Program (STSIMP)). Implementation of new and improved tools and methods into routine monitoring supports the ongoing improvement of this program.

The historical focus of the WWOA program has been to mitigate the human health risk from WWOs. The current and future program has equal weighting assigned to the risk to ecological receptors and reflect the risk to human health. To do this effectively, monitoring and modelling techniques are required. The recent transition from a regulation by frequency target to a risk-based wet-weather overflow abatement program has provided a unique opportunity to investigate monitoring approaches new to Sydney Water. The WWOM program has undertaken a suite of comprehensive scientific studies and wastewater system investigations to test these 'new' monitoring technologies and methodologies and in doing so, develop a better understanding of how the wastewater system operates (in real time) and whether we can assess how WWOs effect the receiving environment. When designing the monitoring program, a key objective was to build capability uplift within Sydney Water Laboratory Services and provide cost-effective options applicable to other programs, such as the STSIMP.

Planning approval requirements prior to installing monitoring equipment

Monitoring equipment was deployed at numerous locations, with varying environmental, social and cultural sensitivities. Sydney Water conducted an environmental assessment under the requirements of the *Environmental Planning and Assessment Act 1979* (Multisite Review of Environmental Factors) for this work. Several sites were within National Parks, requiring additional assessment under the *National Parks and Wildlife Act 1974* of the potential for environmental and cultural impacts. Under the Access for Maintenance, Repair and Operation of Sydney Water Infrastructure in Parks and Reserves: Consent & Protocol (BMIS0128), we were able to undertake our installation works as a Type A proposal (defined as activities that are likely to have no impact, or a negligible impact, on the natural and cultural values of the park or on park visitors or neighbours), this streamlined the approval process and met the National Parks and Wildlife legisaltive requirements.

7.1 Genomic microbial source tracking capability development

Human Faecal Marker Genes

Dr David Roser, UNSW Water Research Centre, was initially engaged to conduct a literature review to inform the early concepts for sampling for parameters relevant to human health. He was introduced to WWOM by Sydney Water's internal human health specialist Dr Peter Cox. Dr Roser based his review on the 2008 Guidelines for Managing Risks in Recreational Water (NHMRC), which advocate a risk-based management approach to characterise contaminants of concern; map contamination pathways (dispersion, dilution and mixing of sewage) and key exposure locations and identify the exposed populations. Dr Roser's review advocated that the HFMG molecular assay of human *Bacteroides* measured with qPCR technology was a very high priority for development and validation. Dr Roser also advocated that the *Bacteroides* HFMG should be measured against *E.coli* and enterococci to measure the load of sewage in the presence of a faecal contamination dominated by non-human sources. These early concepts were presented to the WWOM expert peer panel, who endorsed his recommendations.

The subsequent engagement of Dr Warish Ahmed (CSIRO) in conjunction with the expert peer panel advice generated the human health studies (Figure 4-1) primarily based on water column testing. Our collaboration with Dr Ahmed started as a pilot study into the usefulness of HFMGs to detect WWOs, however it evolved into a highly successful project with 12 high-end journal publications documenting the outcomes. The research presented in Ahmed *et al*. (2024) reported on how risk based threshold (RBTs) concentrations for four specific sewage-associated markers were established, based on rigorous work characterising concentrations and decay rates of HFMGs and pathogens in the estuarine environments of the Sydney region conducted under the 11 human health pilot sub-studies. These RBT's, for HF183, Lachno3, CrAssphage and PMMoV, were derived through a QMRA model that simulates the risk of becoming ill from human Norovirus (HNoV), after ingesting sewage-contaminated estuarine water. The modelling used an equation published by Boehm *et al*. (2018) and Schoen *et al*. (2020) expressing concentrations of HNoV in water contaminated by sewage of a known age as a function of marker concentrations and decay rates, but with site-specific data for both. The median RBT values from the simulations represent the best estimate of values that correspond to a USEPA health risk benchmark of gastrointestinal illnesses rate of 32 per 1000 swimmers (USEPA, 2012). These RBT concentrations were calculated for fresh (day 0) and aged (day 1 to day 10) sewage contamination after a WWO spill at two specific sites. The assumptions used in the modelling are presented in the paper.

A future implementation option for the WWOA risk prioritisation methodology is detailed in Section 3.4 with further continuous improvement suggestions outlined in Sections 3.5 and 3.6.

7.2 Chemical tracers of sewage capability development

Sampling techniques to detect chemical markers indicative of wet-weather overflows and urban stormwater runoff

The sporadic nature and wet-weather event-driven characteristics of WWOs from the sewer, along with the risks of sampling during high flows, present logistical and financial constraints for the broad-scale collection of sufficient water samples to adequately characterise events and determine whether WWOs have occurred where gauging is not in place. Ort et al. (2010) and Vallejo et al.

(2013) suggested that spot samples from a few points in time are unlikely to capture fluctuations of rain-driven wet-weather spills and continuous sampling would be required.

Passive sampling methods have become more commonly employed for monitoring environmental concentrations of chemicals over longer periods in waterway assessments (Vrana et al., 2005; Harman et al., 2012) and as a screening tool to identify and prioritise locations for more detailed investigations (Mutzner et al., 2019; Taylor et al., 2020).

Historically, Sydney Water has utilised autosamplers and grab samples to collect instream water samples. Autosamplers are generally deployed over a set period; and are triggered after a meteorological station activates a rainfall trigger condition or after a sewer gauge activates a wastewater level trigger condition. Grab samples can be taken by staff accessing designated sampling locations via foot or boat. Both deployment of autosampler equipment and retrieval/collection of samples are time, cost and resource intensive. In comparison, passive samplers may be deployed over a few wet-weather events, subject to biofouling rates of local waters (biofouling is known to inhibit effective uptake of chemicals in passive samplers), before requiring maintenance (cleaning/minor repair) or replacement.

What are passive samplers?

There are a variety of commercially available passive sampler devices for detecting organic chemicals in water. The two most popular types of passive samplers are Chemcatcher™ (CC) and Polar Organic Chemical Integrative Samplers (POCIS). Both CC and POCIS use sorbent material covered or sandwiched by membranes. Water flows over an adsorbent (which is the active part of the passive sampler) to capture targeted substances from the water column without active transport or pumping. Passive samplers can easily be deployed within a waterway; must be submerged at all times; do not use electricity, have no moving parts and are simple to use.

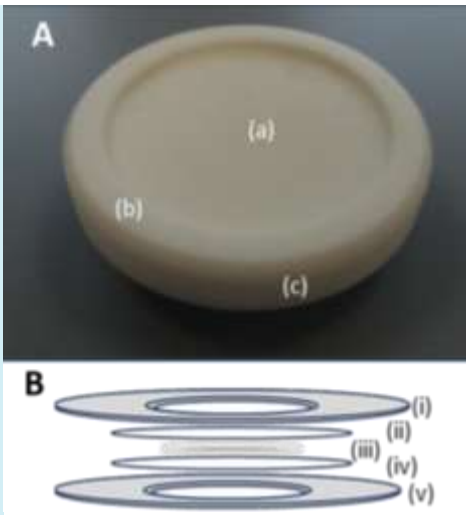


Figure 7-1: Passive samplers: (A) Chemcatcher® and (B) POCIS

The Chemcatcher® photograph (A) shows the components (a) Empore disc and Omnipore diffusion limiting membrane, (b) compression ring and (c) supporting base. The components of POCIS in figure (B) are (i) an upper compression ring, (ii) upper membrane disk, (iii) sorbent layer, (iv) lower membrane disk and (v) lower compression ring.

Before adopting passive samplers, we assessed whether passive samplers accumulate, at a minimum, the same chemicals as those detected in the water column from autosampler grabs over set time intervals. The study assessing passive sampling methods are described in Section 4.3.

Effectiveness of passive samplers as a tool for WWOA

Passive samplers cost approximately 10x less expensive than autosamplers to purchase and install. In the WWOM study, passive samplers also detected more sewage and urban stormwater runoff organic chemicals than water samples collected by autosamplers. However, utilisation of these devices during the WWOM program has highlighted the following limitations:

- a lack of quantified concentration data obtained unless a site-specific stream gauging component is added.

- deployment durations can vary between 2 to 6 weeks for some chemicals (as outlined in Section 4.3). Further site-specific calibration is required to account for differing sampling rates (Rs) across persistently detected chemicals. The ability to cost-effectively deploy passive samplers across a wider spatial area would be negated by the need for calibration of each individual passive sampler chemical.
- POCIS disks are fragile and can be damaged by the pounding of gravel/sand suspended in powerful water movement or destroyed caused by strong currents and/or biofouling (Figure 7-2). Damaged disks were observed during retreival, mainly torn or separated from the POCIS disk holder rendering the POCIS unable to be analysed. Biofouling is also known to inhibit effective uptake of chemicals in passive samplers and although biofouling wasn't initially observed in the testing phase some passive samplers employed at estuarine sites were found to incur significant biofouling. These POCIS disks were unable to be analysed. This loss of information typically coincided with more intense wet-weather events.
- saline environments caused rapid corrosion of passive sampler cages and replacement cages were required to be manufactured during the study. Electroplating can extend the lifespan of a passive sampler cage, minimising the need for replacement, but passive sampler cages would have to be viewed as expendable consumables as they will rust and need replacement.
- an additional step of cleaning of passive sampler cages and disk holders was introduced into study design and costings, as was the use of corrosion proof materials. This was not an anticipated cost, or resource impost.

Figure 7-2: Examples of passive sampler failures that can reduce or prevent chemical uptake, after a four-week deployment

(A) disks of CC and POCIS retrieved from Devlins Creek covered with sand/sediment; (B) a torn POCIS disk damaged from pressure of water flow; (C) biofouling covering surface of a CC; (D) and (E) biofouling of POCIS disks retrieved from Chowder Bay; and (F) POCIS disk separated from the compression ring

Observations of autosamplers effectiveness

Autosamplers are a proven way of sampling surface waters, reducing error with high reproducibility. Autosamplers are also a more reliable tool for calculating the concentrations of chemicals than passive samplers, specifically the organic chemicals in sewage and urban stormwater runoff analysed in WWOM. As with all in-situ monitoring equipment, autosamplers have limitations to deployment across a large area. The WWOM program observed the following during the various studies outlined in Sections 4.3, 4.4 and 4.5:

- the complete autosampler equipment required included the ISCO Avalanche Sampler; a kiosk to house the sampler; a power source (WWOM utilised up to three Lithium batteries); polyethylene collection bottles and suction tubing.
- installation of equipment can cause inadvertent environmental impacts, for example attachment of a kiosk to a rock outcrop within a National Park at the Devlins Creek site. Additional environmental impact assessment and care is necessary when siting autosamplers in sensitive environments.
- in a similar vein, the autosampler kiosks require installation on suitable surfaces and may require a cement plinth to be constructed.
- suitable placement is vital to ensure samples are collected as needed and equipment is not impacted by flooding (Figure 7-3), as seen at Bidjigal Reserve during an extreme weather event in 2020.
- autosampler kiosks are conspicuous and a target of vandalism. This was observed at several sites, for example, Gymea Bay where an opening was forced into the kiosk which was then used as a waste bin; and at Devlins Creek, where the kiosk was broken apart and the autosampler damaged. Damaged kiosks require replacement and equipment repairs, were an additional cost and resource commitment.
- ISCO Avalanche samplers require regular and ongoing maintenance, including:
    - purchase and replacement of batteries
    - reprogramming of an autosampler to obtain events and after failed events sample collection
- sample collection is dependent on wet-weather events, and can often occur on weekends, which increases the cost to deliver. As part of sample collection after an event, the retrieval of samples needs to occur within established sample holding-times for laboratory results to be valid.
- most sampling occurs during wet-weather events; hence resources are limited and sampling has to be balanced between the competing projects.
- some wet weather events can be dangerous when waterways flooded, creating Work, Health and Safety (WHS) concerns. A number of autosamplers were damaged due to flooding. Flooding also precluded access to autosamplers on some occasions, with sampling holding times exceeded before samples could be safely recovered.

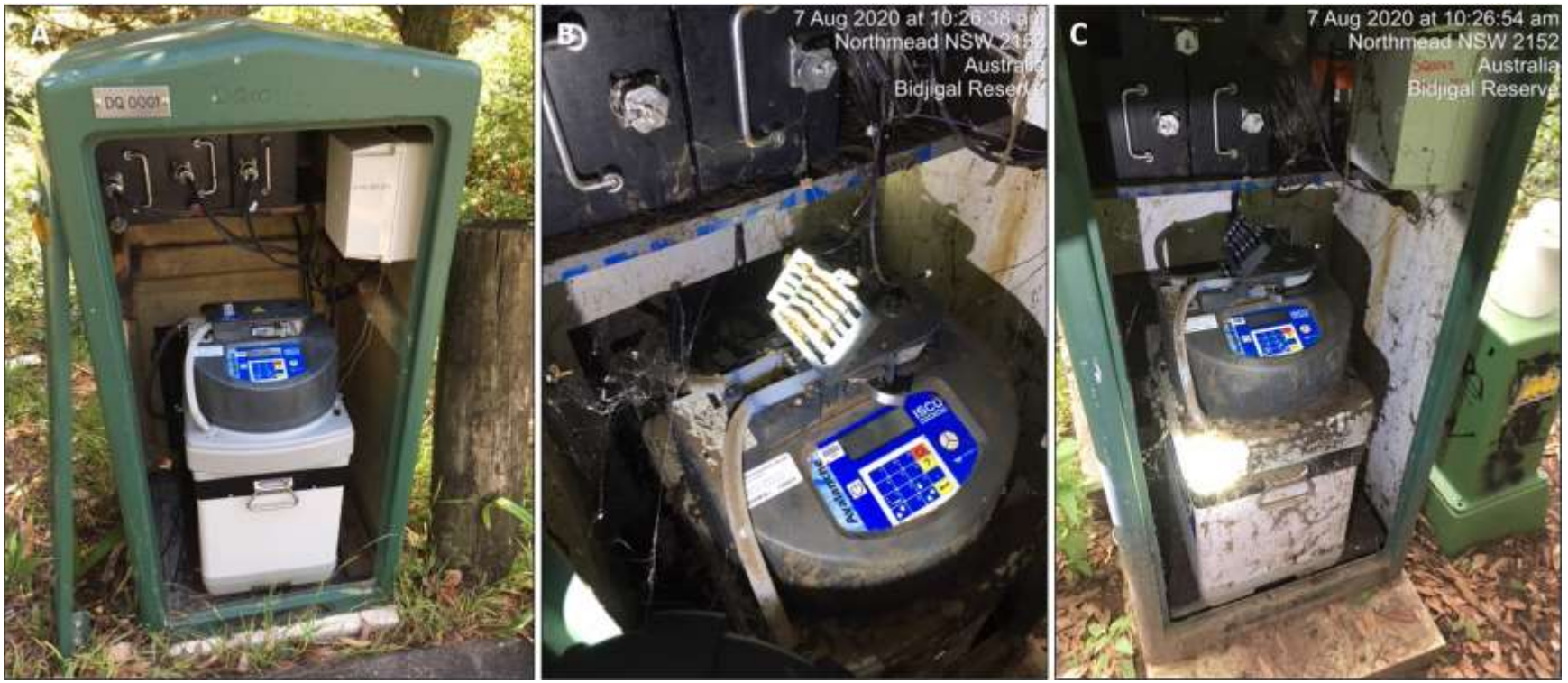


Figure 7-3: Autosamplers deployed to collect chemical markers

(A) an example of the ISCO Avalanche auto-sampler; (B) and (C) autosampler at Bidjigal Reserve, Northmead, after being flooded during a large wet-weather event.

Sampling techniques for chemical markers recommendations

Passive samplers are more cost effective than autosamplers to deploy across a wider area. They can detect chemical markers of sewage contamination and urban stormwater runoff provided the laboratory has the capability to identify target chemicals. However, unless passive samplers are calibrated for each specific marker chemical at a site, their use is limited to a presence/absence detection. The additional cost of stream flow gauging is required to allow calculation of quantitative information.

Based on WWOM project experience of low-level recovery of samples, autosamplers are not recommended. Purchase, set-up time and cost, ongoing maintenance and resourcing commitment to collecting samples is not cost-effective and impractical to implement in a broad-scale assessment. The dangerous conditions during wet weather also prevented sample collection on a number of occasions.

The above listed issues outweigh benefits and as such it is not recommended to apply passive samplers to detect WWOs in a broadscale assessment of potential impacts.

References

Allinson, M., Cassidy, M., Kadokami, K., Besley C.H. 2022. In situ calibration of passive sampling methods for urban micropollutants using targeted multiresidue GC and LC screening systems. Chemosphere, 311, 136997. https://doi.org/10.1016/j.chemosphere.2022.136997

Harman, C., Allan, I.J., Vermeirssen, E.L.M., 2012. Calibration and use of the polar organic chemical integrative sampler-a critical review. Environ. Toxicol. Chem. 31, 2724–2738. https://doi.org/10.1002/etc.2011

Mutzner, L., Vermeirssen, E.L.M., Mangold, S., Maure, M., Scheidegger, A., Singer, H., Booij, K., Ort, C., 2019. Passive samplers to quantify micropollutants in sewer overflows:accumulation behaviour and field validation for short pollution events. Water Res. 160, 350–360

Ort, C., Lawrence, M.G., Reungoat, J., Mueller, J.F., 2010. Sampling for PPCPs in wastewater systems: comparison of different sampling modes and optimization strategies. Environ. Sci. Technol. 44 (16), 6289–6296

Sharp, S., Allinson, M., Pettigrove, V. (2013) Application of Passive Sampling for Sydney Water's Wet Weather Overflow Abatement Program. Centre for Aquatic Pollution Identification and Management, Technical Report No. 91, University of Melbourne, Victoria, Australia

Taylor, A.C., Fones, G.R., Mills, G.A., 2020. Trends in the use of passive sampling for monitoring polar pesticides in water. Trends Environ. Anal. Chem. 27 https://doi.org/10.1016/j.teac.2020.e00096.

Vallejo, A., Prieto, A., Moeder, M., Usobiaga, A., Zuloaga, O., Etxebarria, N., Paschke, A., 2013. Calibration and field test of the Polar Organic Chemical Integrative Samplers for the determination of 15 endocrinedisrupting compounds in wastewater and river water with special focus on performance reference compounds (PRC). Water Res. 47, 2851–2862

Vrana, B., Mills, G.A., Allan, I.J., Dominiak, E., Svensson, K., Knutsson, J., Morrison, G., Greenwood, R., 2005. Passive sampling techniques for monitoring pollutants in water. Trends Anal. Chem. 24, 845–868. https://doi.org/10.1016/j. trac.2005.06.006

7.3 Investigations with DNA under WWOM

Sydney Water applied the following ecological sampling methods for assessing adverse ecological effect of WWOs:

1. widespread (56 urban stream sites) macroinvertebrate sampling with morphometric taxonomy (Figure 7-4A)
2. diatom settling plates to obtain colonising diatoms for subsequent eDNA extraction in urban freshwater streams (Figure 7-4B)
3. sediment corer sampling in estuarine waters along with sediment grab sampling in freshwater streams (Figure 7-4C) for subsequently extraction of eDNA (Figure 7-4D).

Each approach is outlined below, with an example of the method depicted in Figure 7-4.

Figure 7-4: Various methods tested for collecting DNA data

(A) Traditional macroinvertebrate collection; (B) deployment of diatom plates; (C) UWITEC single corer collecting estuarine sediment samples; and (D) collecting sediment samples for eDNA analysis

Macroinvertebrate sample collection

Macroinvertebrates were sampled upstream and downstream of ERSs, with additional upstream sites sampled where there was a confluence of two streams located above the ERS. Samples were collected from pool edges using rapid assessment methods (for example, Chessman 1995, Turak et al, 2004). The pool-edge habitat is an area of unbroken water surface within 2 m of the bank.

Approximately 10% of the edge habitat of the overall site or 10-metres of a 100-metre reach was sampled using a plankton net. Macroinvertebrates were picked with forceps and pipettes into a collection jar containing 85% ethanol over a period of 60 minutes. A counter was used to tally the number of animals collected and a timer/stopwatch used to monitor picking times.

Identification and enumeration were conducted by Sydney Water Laboratory Services using the National Association of Testing Authorities (NATA) accredited in-house test method for macroinvertebrate identification and enumeration. Identification and counting was carried out up to genus taxonomic level where possible, using published keys (Hawking, 2000), web links (for example Centre for Freshwater Ecosystems Identification and Ecology of Australian Freshwater Invertebrates), or using descriptions and reference specimens maintained by the Sydney Water laboratory.

Outcomes from the analysis of the morphometric macroinvertebrates assemblages are outlined in Sections 5.1 to 5.3. Bulk and single specimen macroinvertebrate samples were collected for metabarcoding. The comparison of biotic index scores from assemblages of morphometrically identified macroinvertebrates and DNA is detailed in Section 5.5.1.

Diatom settlement plates

Diatoms are silicon-based microorganisms, found in all types of waterways, and are used to monitor ecological condition. In ‘A diatom species index for bioassessment of Australian rivers’, Chessman et al (2007) assessed ~248 species and assigned sensitivity values (that is pollution sensitivity) as a basis for an eDNA diatom index. Both the European Union and the United States are currently working on developing a molecular-based diatom index from eDNA. Under the WWOM, an initial eDNA pilot study (2017) of water samples identified a good representation of diatom taxa. Subsequent discussion with the peer-review panel recommended further exploration of diatoms (and potentially a diatom index) using settlement plates (April 2018 peer-review workshop) to obtain a time-integrated measure of a site. Settlement plates would potentially capture changes in taxa present in response to periods with and without overflow activity as they would be deployed for four-weekly periods. It was proposed that this would avoid the potential for historical WWO overflow bias that could be observed in sediment derived diatom assemblages.

Development of a climax community of diatoms can take about four weeks (Taylor et al., 2007), dependent on environmental variables and time of year. There were five diatom Perspex™ settlement plates (Figure 7-4C) deployed at each freshwater location on an eight-week collection/redeployment cycle. After each eight-week placement, the plates were removed and processed to collect a DNA sample and clean plates were installed. Plates were placed at random spots within the sampling location on each deployment; avoiding the same positioning, grouping and patterns, this was to remove any pseudo-replication. Settlement plates were located within the site reach (nominal 100 m), but with similar flow conditions, for each deployment period. Plates were positioned horizontally with the roughened side facing upright, and at an optimal depth of

0.4 - 0.5 m. Plates were roughened on one side during manufacturing to provide a suitable surface for diatoms assemblages to establish. Taylor et al. (2007) suggests performance of diatom indices is not affected at depths up to 0.5 m. Urban streams provide a dynamic environment with varied substrates and water clarity. To mitigate plate loss and movement during deployment, plates were attached to varied support structures, including bricks and star pickets (Figure 7-4B illustrates an example of a brick configuration) and were located in pools that would provide stable water depth and level.

Key learnings from the implementation of diatom plates

The diatom plate method was abandoned in March 2019 with agreement from the peer panel. This was decided after continued deployment periods experienced significant loss of plates (~50%) across almost all locations and varied complications that compromised sample integrity. These issues are detailed below:

- The initial deployment periods coincided with drought, and many stream pools became dry and left plates exposed
- When the drought ceased, and Sydney experienced moderate to large rainfall, the significant velocity and volume of stormwater resulted in plates and the support structures being displaced and washed downstream and onto stream banks (Figure 7-5A)
- The ‘flashy’ and dynamic nature of the urban streams outside of the more extreme events also resulted in the periodic exposure of plates and the movement of sediment beds smothered plates (Figure 7-5)
- The roughened face of the plates also provided a suitable surface for the recruitment of filamentous algae, and in some locations this algal growth thickly covered the full face of the plate

These issues combined resulted in limited data collection each deployment period and compromised sample integrity. Considering these issues, and with peer panel support, it was decided to manually collect sediment samples at the freshwater locations.

Figure 7-5: Diatom plates sampling challenges

(A) example of monitoring tools after a severe rain event; (B) pedestrian bridge downstream of diatom settling plate deployment area; and(C) the pedestrian bridge after a severe rain event

Sediment sample collection as a basis for extraction of eDNA

The collection and identification of macroinvertebrates is costly (Marshall et al., 2006), and applying molecular techniques in the WWOM program was seen as an opportunity to investigate methods currently not employed by Sydney Water. The WWOM expert peer panel introduced Associate Professor Anthony Chariton to the project to advise and train Sydney Water Laboratory Services on eDNA sampling methods and was subsequently engaged to undertake the analysis (bioinformatics) and modelling of the DNA-sequenced data.

Sediment samples were collected via two methods:

- At estuarine sites, core sediment samples were collected via UWITEC single corer (Figure 7-4C). Cores were collected from 2 m depth below the Mean Low Water Spring (MLWS) tide level. Each core was taken at random locations within the depth constraint for each study location (to avoid pseudo-replication). Four cores were collected at each study location. To prepare the cores and minimise cross-contamination, all cores were pre-soaked in 10 to 15% bleach for 24 h, rinsed five times will Milli-Q water and wrapped in food ‘cling’ wrap for transportation. Once a core was collected, 0 to 2 cm of sediment (approximately 30 grams of sediment) from the top of the core was transferred into a 50 mL Falcon tube. To avoid contamination of samples, samplers replaced gloves; used new

spatulas; and did not reuse Falcon tubes after each core sample. Samples were transported to the laboratory in a portable freezer and transferred to a -80°C freezer until undergoing the DNA extraction process.

- In freshwater streams, grab samples using 50mL sterilised and nuclease-free test tubes (Figure 7-4D). Samples were collected upstream and downstream of ERSs. At each freshwater sampling location four sediment samples were collected at a minimum 1 m apart and at a depth of no greater than 0.6 m in a sediment-rich area comprised of fine sand to silt texture. Areas with organic detritus and iron-oxidising bacterial flocs were avoided. Samples were collected by sweeping an uncapped test tube across the surface of the creek bed for approximately 40 cm, collecting a maximum 30 mL sample of fine sediment from the bed substrate. Samples were transported to the laboratory in a portable freezer and transferred to a -80°C freezer until undergoing the extraction process.

Blank sediment samples were used to establish background conditions. Blanks were generated by destroying (bleach wash and rinse cycles) the DNA from sediment collected at the start of the program from one location. Confirmation that no DNA was present in the samples was obtained by performing PCR amplification of 16S primers. Blanks were exposed at sample site for the same collection duration of time to fill a sample tube. One blank was also used in the laboratory to assess the freezer storage by placing one Falcon tube sediment blank in within the inward storage a -80°C freezer.

To extract the DNA, samples were thawed in batches of 10 (9 samples and a control), homogenised in a bead mill homogeniser (Omni Ruptor 24) and extracted using the DNeasy PowerMax® Soil Kit from Qiagen using the manufacturer's instructions. Extracted DNA was aliquoted into DNA/RNA free cryovials (Thermo Fisher) and stored in a separate -80°C freezer. A tube of nuclease free water (extraction blank) was processed alongside the nucleic acid extraction to assess the quality of the extraction process.

Companion grain size data were collected as grain size is a known influence on benthic fauna. For the estuarine cores, a 250 mL grain size sample container was filled from each of the three eDNA replicate sample sites within a location, once in 2019 and once in 2020 (based on peer-panel advice that estuarine sediments were unlikely to dynamically change). In contract, freshwater grain size sediment samples were collected at eight-week intervals as urban streams are known to have continually moving bedloads. Freshwater sediments were sampled in a 200 mL container via the same technique used for DNA sediment collection. Samples were analysed for 3 grain sizes (<0.063 mm, 0.063 mm to <2.0 mm and >2 mm, representing mud, sand and gravel, respectively).

Sydney Water engaged the UNSW Ramaciotti Centre for Genomics (Ramaciotti) to conduct sequencing of the DNA-extracted samples. Once samples were received by Ramaciotti:

- PCR condition tests were performed to assess the viability of the extracted DNA assays
- PCRs were conducted on assays
- Replicates for each product was assessed on an agarose gel for the presence of absence of a band, which indicated if the reaction passed or failed
- Pooled PCR products were then sequenced on MiSeq v2 Nano sequencing equipment to check the quality
- Once the quality of the pooled PCR products has been determined, each bulk pool was sequenced on either a NovaSeq 6000 SP or Illumina NextSeq800

- Following sequencing, the run was demultiplexed and quality assessed

Section 5.5.3 outlines the relevant primers used in the sequencing runs for both freshwater and estuarine samples. Sections 5.6 and 5.7 outline the project objectives and outcomes from the molecular eDNA study and Section 5.8 outlines the key findings.

Key learnings from implementation of ecology studies with DNA

- Costs to be considered in adoption of genomic methods include:
    - data storage and database development requirements from the increased amounts of information provided by molecular derived data
    - purchase and training in use of related software required to process laboratory derived molecular data reads into relative abundance molecular OTUs
    - next generation sequencers (that is, the latest models) can provide sequenced data more cost-effectively. However, continually purchasing the latest sequencing equipment is likely to be cost-prohibitive if there is not the generation of enough project work to justify the investment
    - There are ongoing costs of regular maintenance of the sequencing equipment. It was a prudent decision not to purchase at the start of program when other equipment was purchased, given the rapid advances in technology has achieved much better quality.
    - WWOM purchased three -80°C Thermo Scientific™ TSX Series Ultra-Low Freezers to store inbound sediment samples and extracted DNA samples. These freezers have a backup system which maintains the temperature with liquid nitrogen during power outages. Note: a power outage occurred during the project, hence this additional option was critical to protect the integrity of the samples
- Determining the number, type and compatibility of primers used during the sequencing process took longer than expected. A pilot study to assess these factors was implemented to confirm the suite of primers applied to freshwater and estuarine DNA. It was found that multiple primers provided greater coverage for DNA assessment.
- WWOM generated a large number of samples for sequencing, the estimated workload for the project was in excess of the annual workload for the DNA sequencer provider, Ramaciotti. Due to resourcing impacts and supply limitations during the covid-19 pandemic, delays to the delivery were experienced.
- Sediment sampling may require fewer members of staff for sample collection.
- Community-DNA based on bulk sampling methods would still require similar staffing as morphometric sampling, although the laboratory-cost component would be reduced and replaced with sequencing and bioinformatics to obtain the taxonomy dataset.
- Bioinformatics is a necessary step to cleanse raw sequenced data, engagement of a bioinformatician is not currently cost-effective until the whole DNA analysis process is implemented in-house.
- Sediment samples encapsulates a longer time record of taxa present compared with water samples, which represent transient taxa. When water sampling was conducted well after an overflow event, this collected transient water may not capture the pulse nature of recent WWOs.

DNA and traditional morphometric tools recommendations

Traditional morphometric identification of hand-picked macroinvertebrates remains a suitable method to collect data to determine adverse ecological effects based on spatially close (upstream and downstream) paired-site assessments (Section 5.3).

Diatom settlement plates are not recommended. There are limitations for deployment in Sydney freshwater streams, as during storm events flooding and high energy flows can cause the equipment to be deposited outside of the waterway. Along with the inability to assess for diatoms, this is a public and employee safety issue.

The UWITEC sediment corer was most effective at collecting recent depositional sediments (that is, the top layer) without mixing the surface layer. In contrast, grab samplers can mix surface and lower layers. If the aim of a project is to obtain recent deposition the UWITEC sediment corer is highly recommended.

Collection, sequencing and generation of eDNA taxonomy datasets from sediment samples successfully returned high-quality data (refer to Sections 5.6 and 5.7).

Broad region modelling with the morphometric or eDNA-based taxonomy datasets against assembled metadata variables were unable to encapsulate sufficient predictive capacity to enable a viable modelling approach (Sections 5.4, 5.6 and 5.7) for WWOs as an input to the risk prioritisation methodology.

Future application of DNA derived taxonomy may most effectively be based on community-DNA analysis from bulk sample collection of macroinvertebrates, as a basis for calculating sample biotic index SIGNAL-SG scores. This future application of eDNA will be explored under the SWAM program. This exploration would be best evaluated with pilot studies that compare DNA and traditional samples to determine if at least similar outputs are obtained. Proposed pilot studies are outlined in Section 5.9.The purchasing and maintenance of a sequencing platform and employment of suitable operators will be assessed for cost-effectiveness after determining whether community-DNA or eDNA is the basis of implementation.

Future programs utilising DNA under a metabarcoding approach should employ a pilot study to determine which primer-pairs (amplicons) are compatible and suitable for the project aim. Future project aims potentially fall into two categories: to target a particular taxonomic group, to provide the basis for biotic index calculation; or, to understand the diversity of a site/study area.

If Sydney Water were to utilise an external sequencing provider, a review of providers would be prudent to ensure commitments can be delivered as there are now more than one suitable provider in Australia/NZ marketplace.

References

Chessman B.C. (1995). Rapid assessment of rivers using macroinvertebrates: a procedure based on mesohabitat-specific sampling, family-level identification and a biotic index. Australian Journal of Ecology. 20: 122-129.

Hawking J.H. (2000). A Preliminary Guide to the Keys and Zoological Information to Identify Invertebrates from Australian Inland Waters. Guide No. 2 Second Edition (Cooperative Research Centre for Freshwater Ecology: Albury VIC).

Marshall, J.C., Steward, A.L. & Harch, B.D. Taxonomic Resolution and Quantification of Freshwater Macroinvertebrate Samples from an Australian Dryland River: The Benefits and Costs of Using Species Abundance Data. Hydrobiologia 572, 171–194 (2006). https://doi.org/10.1007/s10750-005-9007-0

Taylor, J.C., Harding, W.R. and Archibald, C.G.M., (2007) A Methods Manual for the Collection, Preparation and Analysis of Diatom Samples, Version 1.0. Report to the Water Research Commission, Gezina, Pretoria, South Africa TT282/07

Turak E., Waddell N. & Johnstone G. (2004). New South Wales (NSW) Australian Assessment System (AUSRIVAS) Sample Processing Manual. Online available at http://ausrivas.canberra.edu.au/Bioassessment/Macroinvertebrates/

7.4 Equipment installed to monitor WWO events and assist with data collection

Sewer Gauges

Since 2000, Sydney Water has been using models of the sewerage network for each of its Sewage Treatment Systems (STSs) to assess the hydraulic and environmental performance of each system. The wastewater system trunk model represents the operation of WWOs on the larger trunk mains and designates nodes to capture WWO points on the reticulation system. The wastewater system trunk model is updated annually to reflect changes in the wastewater system and reported to the EPA. This assessment has historically been on a suite of roaming gauges (approximately 30 across 3,000+ ERS). Our STS licenced models are recalibrated every five years.

Depending on the asset construction, inflow and infiltration inputs, local rainfall, and surrounding topography, each ERS can operate differently. There are numerous types of complex overflow structures, for example, some ERSs operate in a siphonic manner, others require differing weir heights to compensate for frequency of spills, and outlet pipes may have an upward sloping connecting pipe. Most overflow structures have overflow arrangements that include flaps that may be affected by flooding in the creeks or waterways into which ERSs spill Figure 7-6B).

With such diversity, it is important to understand the operation of the ERS, as well as the waterways surrounding these assets and how flooding will affect the spills from WWOs. As such, the WWOM program invested in increasing the number of gauges (Figure 7-7) to better represent WWOs and improve our ability to define their influence on receiving waters.

With ERSs built to different designs, this effectively requires each ERS to be treated differently when installing equipment. In some cases, the location of an ERS was difficult to locate, requiring

the use of dyes to locate the discharge point (Figure 7-6A). As access into the ERS also presented a safety risk to staff, Sydney Water purchased a 3D scanner (Figure 7-6B illustrates a 3D scan of an ERS) for the WWOM gauging installation work to address these safety concerns. The imaging of the ERS structures from the 3D scanners proved to be successful, and additional scanners were purchased for other programs.

Figure 7-6: (A) environmentally sustainable dye used to locate ERSs; (B) 3D image of an ERS; and (C) configuration of a sewer gauge and RTU

Across the four main coastal sewerage treatment systems, ERSs representing the overflows of highest risk identified by the initial 2015 risk prioritisation (based on the STS licenced overflow points) were identified for sewer gauging. WWOM installed 173 level sensors (Figure 7-6C and Figure 7-7), either a sewer level or velocity gauge, with low powered remote telemetry units (RTUs) (Figure 7-6C) within each ERS.

Sewer gauges were set to record any overflow levels above the crest of the overflow weirs (30 mm above the weir to indicate the start of an overflow, and at 0 mm for the cessation of the spill). The gauges were installed within the ERS or within maintenance holes designed to overflow (Hartlids). This measured level data were transmitted to our Supervisory Control and Data Acquisition (SCADA) system in approximate 'real time'. The RTUs were set up to notify sampling crews via SMS when ERSs started spilling into the receiving waters.

The aim of the widespread gauging was to obtain more accurately measured overflow volumes and frequencies, than provided by the modelled overflow performance. These data were then compared to the analyses of chemicals, eDNA and macroinvertebrates. The raw data from the sewer gauges informs the level and measurement values, which are then converted into theoretical rating tables to estimate the overflow volume.

Theoretical rating tables were generated by

- Accessing historical and current data on each ERS and applying the most suitable quality data
- Developing relationships between depth and spill
- Generating the final rating curve and assessing the measurement uncertainty

In some cases, rating tables could not be developed, due to the complexity of the ERS design, the modelled volumes were used for these ERSs. Measurements of uncertainty in volume for the rating tables ranged between <10% to 100%. Uncertainty in calculations arose from the potential for MH spills when the MH lid 'pops' under pressure from the sewer; where backwater (either flows from a creek or tidal influences) overtops the weir crest at the same time as high sewage levels in the system; and whether the data available for the ERS structure are able to be cross-validated to ensure high quality inputs. Once a rating curve had been developed, this information was coded into IICATS BI to calculate the volume of spill.

Text in this section is drawn from the Post Implementation Report: Sydney Water (2019). Report on Post Implementation Review (PIR) of Wet-Weather Overflow Monitoring Programs, Stage 1 Pilot Studies, Project ID: 20033051.

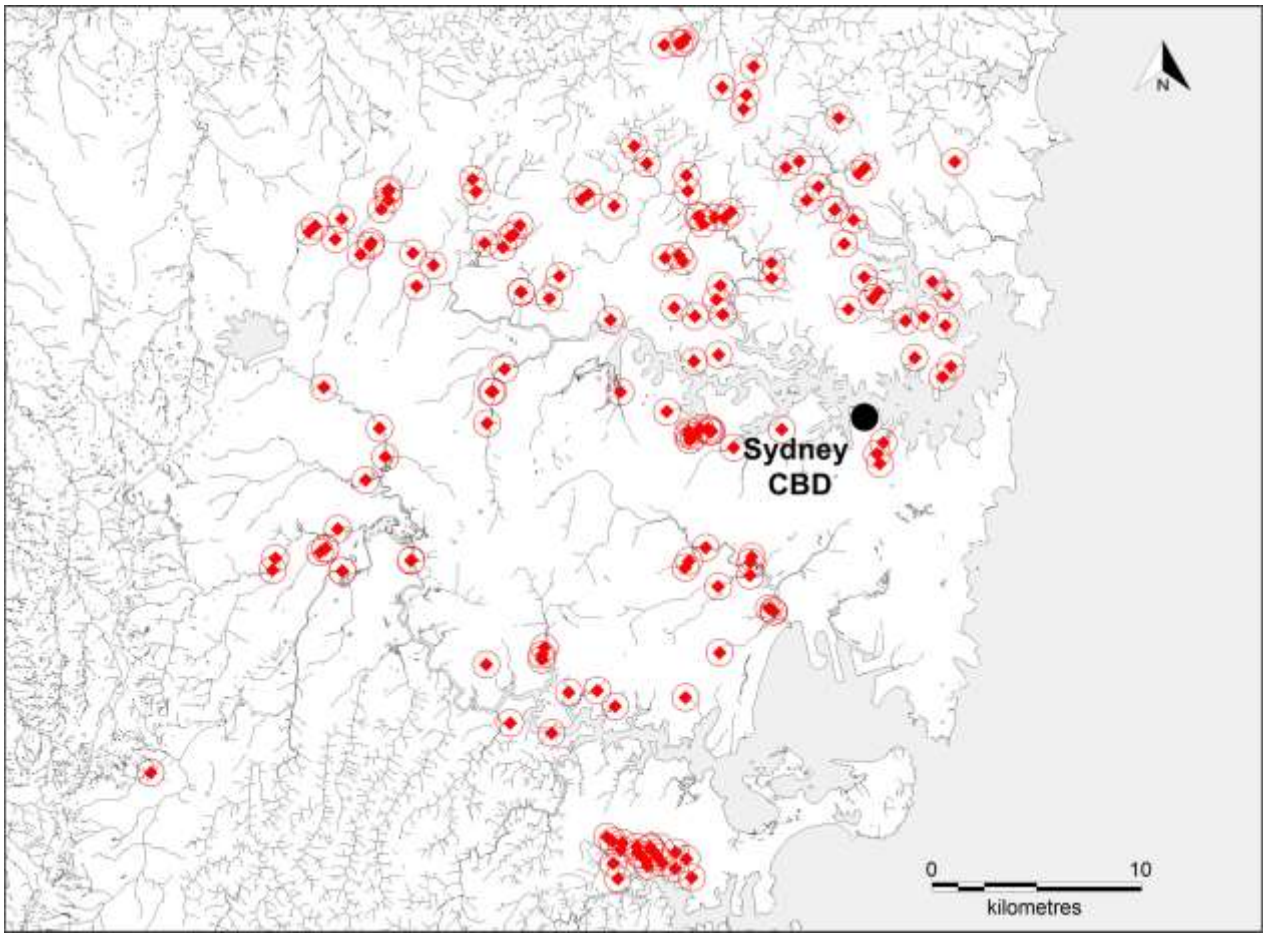


Figure 7-7: Locations where overflow gauges were installed to monitor WWO activity at ERSs

Rain Gauges

Rainfall in the Sydney metropolitan area falls unevenly across the region during small to medium wet-weather events. This complicates the ability to analyse site specific effects. The Bureau of Meteorology (BoM) and Sydney Water combined have approximately 230 rain gauges within the Sydney Water area of operations. Whilst data from established rain gauges covered most of the sampling sites for this program, six additional rain gauges were installed to cover those areas with insufficient coverage. The rain gauges used for this project are Tipping Bucket Rain Gauges (TBRG) which use a 0.5 mm bucket (Figure 7-8A).

Instrument buoys

In the estaurine waterways, we deployed our monitoring equipment on existing structures, such as bridges, pontoons, or piers. In seven locations, no suitable deployment options were available, and temporary buoys were installed. Figure 7-8 shows the deployment of a mesocosm described in Section 3.2.11) in Port Jackson/Middle Harbour (5), Georges River (1) and Port Hacking (1).

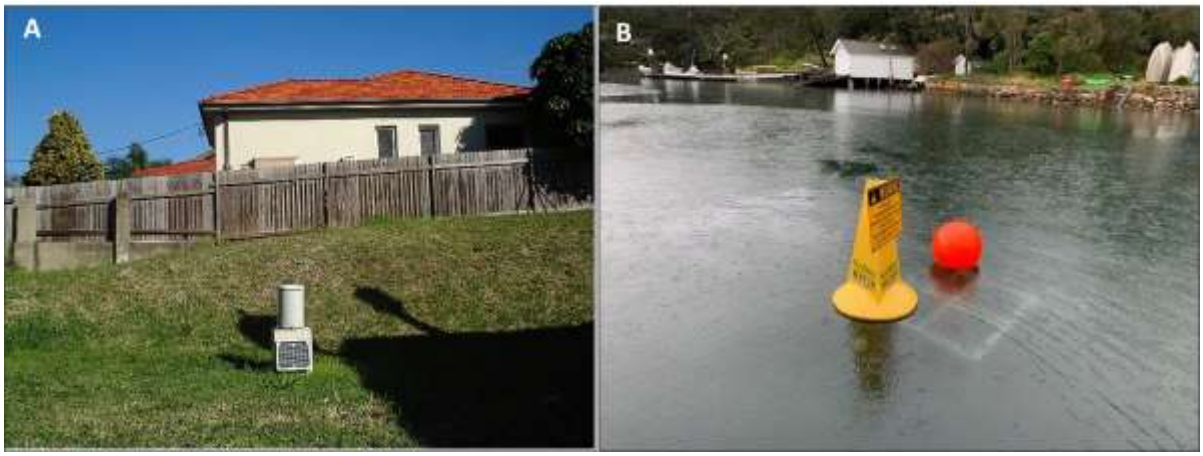


Figure 7-8: Example of a rain gauge (A) and a typical buoy deployment with mesocosm (B)

Monitoring equipment installation learnings and outcomes

Installation of a variety of monitoring equipment over a large spatial area comes with its challenges. Important learnings were gained through the process, which have been applied to the WWOM program, and used to inform other projects, such as the Port Hacking Integrated Catchment Model pilot project. The key learnings include:

- Overflow structures may be located in difficult to access locations, such as private properties, main roads, and heavy bushland, while mobile phone reception (needed for telemetry) can be patchy. It is important to conduct detailed scoping of the various locations where equipment is to be deployed. A comprehensive understanding of costs to deploy, maintain, repair and replace damaged or faulty equipment is important for project success. Unforeseen events experienced during the project included flooding, corrosive damage from sewer gases, saltwater damage and vandalism (human and animal).
- Confined space entry requires additional staff and work health and safety requirements. 3D scanners allowed for safe assessment of the construction and condition of an ERS.
- Understanding the uses and limitations of Sydney Water monitoring equipment, software and data storage platforms is crucial:
    - Data from the SCADA system are stored in IICATS BI, as raw data. Verification of raw data is key to excluding false positive overflow activity, as gauges can malfunction.
    - When conducting assessment of the influence of WWOs on receiving waters, wastewater hydraulic modelling of overflows should be validated with gauging to accurately capture overflow activity.
- The sewer gauging employed for WWOM program, has been an effective input into other projects, wastewater modelling validation and strategic decisions on future gauging needs.

The above learnings should be kept in mind for application to the overarching WWOA that the WWOM was developed to inform.

8 Conclusions & recommendations

Until the current 2020 – 2024 IPART price path, Sydney Water has historically managed WWO abatement predominantly via storage and containment solutions. However, large infrastructure solutions are expensive and resource- and time-consuming to achieve. They can result in increased costs to our customers and localised inconvenience during construction. In the United Kingdom, customer interest and concern over the management of overflows and their potential impacts to ecosystem and public health, has initiated a government investment directive to water utilities for the reduction of overflows. The United Kingdom's Storm Overflows Discharge Reduction Plan (Defra, 2023) has outlined stringent improvement targets for ecosystem and public health to achieve a target of no storm overflows, except in unusually heavy rainfall events, by 2050. The reduction targets have been calculated to cost approximately £60 billion (across all English water utilities, and 25 million customers), with an averaged estimated £45 per annum increase to water bills until 2050 (Defra, 2023).

Where the UK have implemented a frequency target to achieve overflow discharge reductions, from 2020 Sydney Water has been obliged to meet an interim target to reduce volume and/or frequency, which will change from 2024 to achieving a 6% overflow volume reduction (or 1.6 GL) in the four coastal catchments of the North Head, Malabar, Bondi, and Cronulla EPLs. The remaining sewage treatment systems are currently still required to maintain or achieve frequency targets.

To meet the revised targets, Sydney Water has transitioned from storage to source control management of infiltration and inflow into the wastewater system. Calculations developed for the 2025 – 2030 IPART submission identified that the cost of storage would be an investment of 17x more than source control activities and would take over 350 years to complete, based on the current source control spend rate. Whereas it is estimated that all source control works would be completed by 2060. Source control to abate WWOs is both cheaper and more time efficient than storage. However, effective identification of where to invest in source control is paramount for successful abatement of WWOs and to enable Sydney Water to achieve its compliance obligations.

Under the Pollution Reduction Program (PRP) 307 of the North Head, Malabar, Bondi, and Cronulla EPLs, Sydney Water is required to implement a risk prioritisation methodology starting in 2020. This methodology establishes a baseline prioritisation profile for WWO abatement works to the four largest coastal sewerage systems. In addition, the EPA has also implemented a continuous improvement requirement, Pollution Study (PS) 307, '*to resolve technical issues and uncertainties (that is, limitations and assumptions) with the current prioritisation methodology through identification and use of new and improved tools and information to achieve a refined methodology*'. A considerable challenge of effectively prioritising individual ERS for their risk of impact to sensitive receptors in receiving waterways or to human health of waterway users, is to distinguish between the influence of stormwater and wastewater as waterways can receive both WWO spills and urban stormwater runoff (containing other pollutants).

Determining the impacts of WWOs from our sewerage network is difficult. There are over 3000 ERSs across the Sydney Water area of operations discharging into freshwater and estuarine receiving waters, in highly urbanised or industrialised areas with varying levels of riparian bushland. Historically, ERSs have not been constructed uniformly and tend to operate

inconsistently depending on the volume and duration of a rain event and the performance of the local wastewater system. This can be a result of localised rainfall patterns, the asset itself not functioning as designed or the location of an ERS in the system (those located in downstream points on the network may discharge later and longer). It is also difficult to detect or measure the effect of WWOs on environmental and human health and any benefits from abatement solutions. Confirming the presence of wastewater in waterways is not a direct measure of the impact. Suitable cost-effective, targeted, robust, and reproducible waterway monitoring is imperative to inform where to invest. Investment must be targeted to the infrastructure where there is greater certainty around the impact from WWOs on the receiving environment and the users of these waterways.

In 2016, Sydney Water commenced the WWOM program to address PS 307. WWOM was comprised of a suite of scientific investigations, designed to investigate monitoring options to support future revisions of the methodology and to assess the benefits of these investments to the environment and the community. WWOM addressed the two risk streams of adverse ecological effects and human health assessed by the current risk prioritisation methodology by investigating methods and technologies able to detect WWOs and using these methods, to assess if WWO impacts can be separated from the underlying urban stormwater impacts. WWOM comprised investigation of four potential influences of WWOs (Figure 1-4):

- chemical assessments and toxicity testing to determine contaminants of concern;
- trash net collection technology and aesthetic assessments to gain insights into the discharge of sewage-derived gross pollutants from ERSs;
- traditional morphometric and emerging eDNA techniques to investigate if adverse ecological effects from intermittent-sporadic or ongoing disturbance in receiving waters can be ascertained; and
- molecular (DNA/RNA) tools to reflect the human health risk to recreational users from sewage contamination.

The following conclusion sections contain key outcomes and recommendations from these WWOM studies:

- Section 8.1 provides an overview of human and animal faecal-associated marker gene studies and provides case studies illustrating scenarios from observations of contaminants of concern and companion toxicity investigations.
- Section 8.2 details the conclusions from the four research areas into the potential influences of WWOs and outlines the potential application of successfully tested monitoring tools.
- Section 8.3 proposes how direct measurements using the identified tools can be applied to a revised WWOA risk prioritisation methodology.
- Section 8.4 outlines the application of successfully tested monitoring tools to other Sydney Water programs.

8.1 Overview of key findings

An outline of when HFMGs may reflect risk to human health from WWO spills is outlined in the flow chart (Figure 8-1) drawn from key observations from the eleven human health sub-studies (Section 3.2). These sub-studies informed a QMRA model that established site-specific Risk Based Thresholds (RBTs) for four HFMGs (HF183, Lachno3, crAssphage and PMMoV) (Section 3.3), that are aligned to a benchmark level of GI illness risk, for fresh (day 0) and aged (day 1 to day 10) sewage contamination. Future direct measurement of human faecal-associated marker genes (HFMGs) in the water column during/after a WWO event would provide data to rank departures from RBTs. These departures would then provide an input into the WWOA risk prioritisation methodology to allow for a categorisation of each ERS based on highest illness risk to lowest illness risk (Sections 3.4 and 8.2.2).

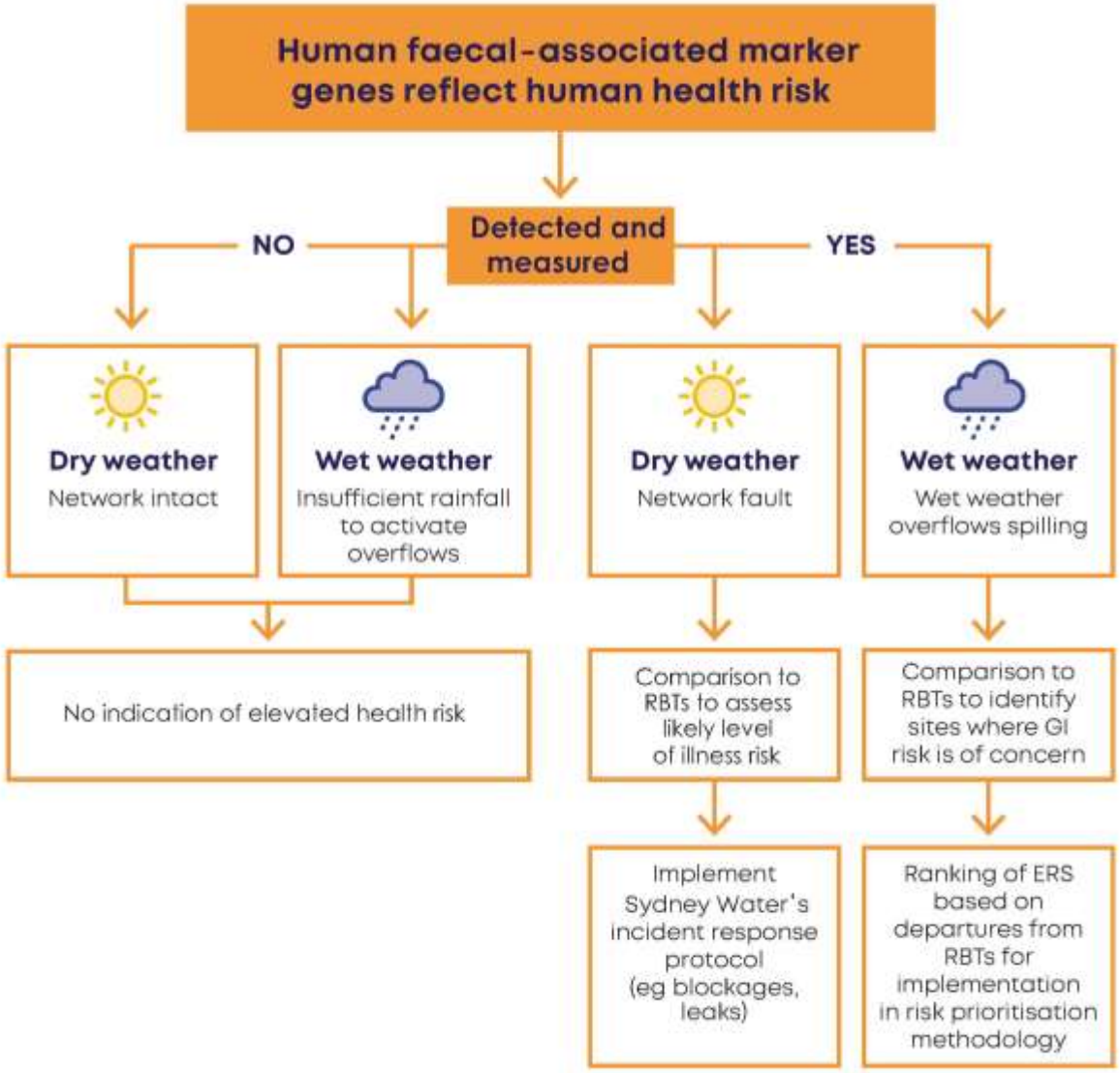


Figure 8-1: Flow chart of key observations from the human health pilot studies outlining when HFMGs may reflect risk to human health from WWOs

The new insights provided from tracking faecal contamination in receiving waters with molecular marker genes has been able to distinguish human- from animal-faecal contamination. Minimal human health illness risk was suggested from observations of faecal contamination from common urban animals (birds and dogs) (Figure 8-2) based on animal faecal-associated marker gene studies (Section 3.2). WWOM study findings clearly illustrate that under wet-weather conditions human faecal contamination is a source of risk to human health. While under dry-weather conditions when the sewerage network is intact, faecal contamination from bird species would represent concentrations of the traditional bacterial indicator of faecal contamination, enterococci.

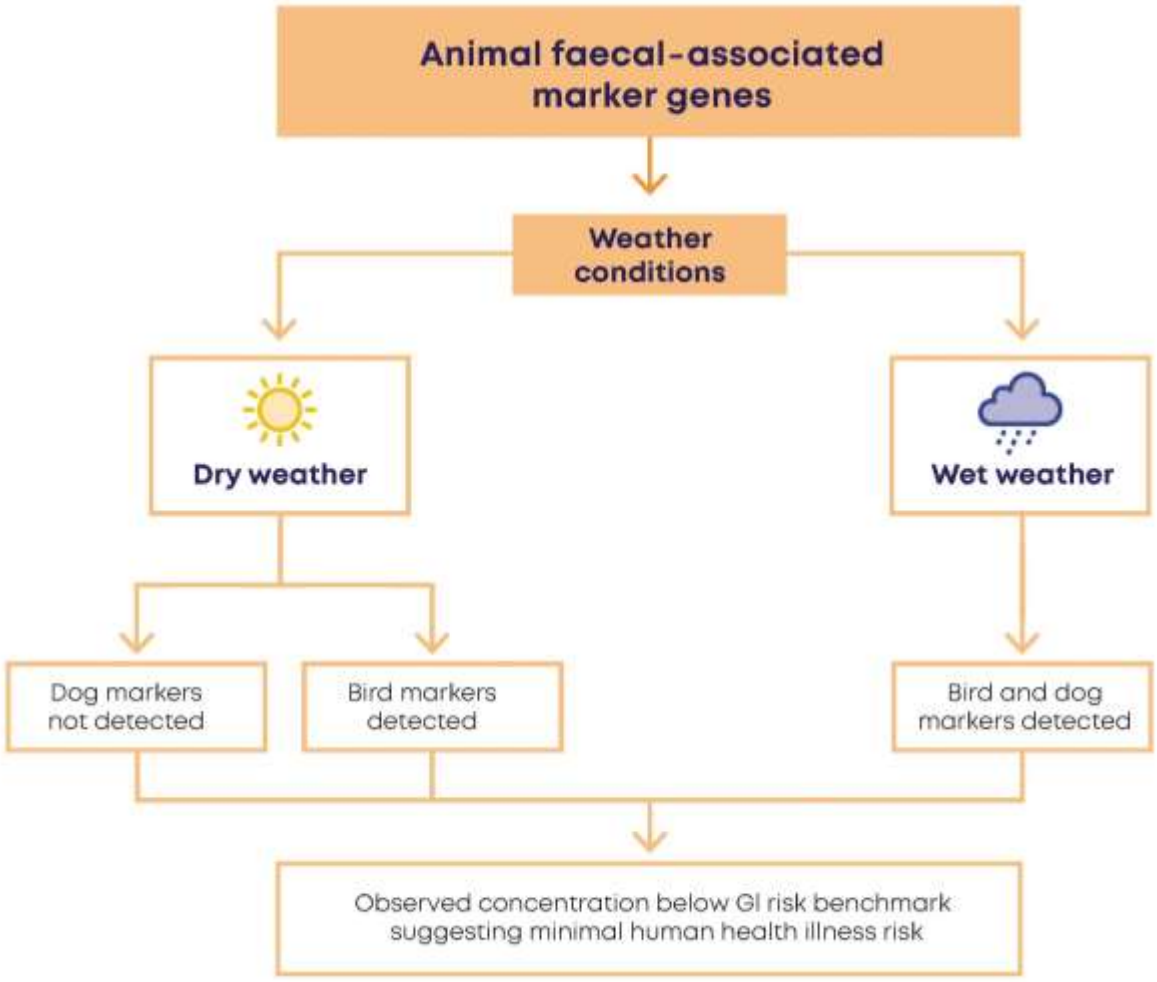


Figure 8-2: Flow chart of observations from animal faecal-associated marker gene studies

*As outlined in Ahmed et al. (2019b) there is no marker gene threshold available for BacCan-UCD (dog) and GFD (avian) marker genes. Ahmed et al (2019b) note that Brown et al (2017) determined that the median concentration of illness equals the GI risk benchmark when gull feces associated *Catellicoccus* marker gene concentration is 6.60 log10 GC/100 mL. This risk benchmark was used to assess quantifications of the avian GFD marker gene. The next most observed animal marker (across the human health sub-studies) was for dogs but with infrequently quantified concentrations in wet-weather samples that were less than the highest observed for the avian maker.

Outcomes of studies evaluating contaminants of concern (Section 4) and companion toxicity investigations (Section 4.5) have established ammonia as a contaminant of potential concern in receiving waters where dilution of WWO spills is insufficient. This insufficient dilution was attributed to adverse ecological effect observed under three receiving water situations from the companion study of morphometric macroinvertebrates (Sections 4.7 and 5.3). These key learnings allow an illustration of scenarios with and without WWOs in combination with baseline stormwater inflows to urban streams (Table 8-1).

References

Ahmed, W., Payyappat, S., Cassidy, M. Besley, C., 2019b. Enhanced insights from human and animal host-associated molecular marker genes in a freshwater lake receiving wet weather overflows, Sci. Rep. 9, 12503. https://doi.org/10.1038/s41598-019-48682-4

Brown, K. I., Graham, K. E., Boehm, A. B. 2017. Risk-based threshold of gull-associated fecal marker concentrations for recreational water. Environ. Sci. Technol. 4, 44–48.

Table 8-1: Case studies illustrating scenarios from observations of the contaminants of concern and companion toxicity investigations

Site example	Scenario	Dilution of WWOs	Chemistry	Toxicity	Ecology	Explanation
Rudder Creek	Urban stormwater flows ✓	**NA**	+	+	+	• The metals, copper and zinc exceed guideline values in stormwater source (without WWOs) • Toxicity due to metals that form about 80% of loading to urban streams from stormwater inflows • Adverse ecological effects seen when compared to near pristine bushland stream where metal sensitive taxa are present (for example mayflies)
	WWO ✗					
Buffalo and Darling Mills	Urban stormwater flows ✓	**> 2x**	+	–	–	• Ammonia less than guideline values • Metals exceed guideline values • Chronic toxicity observed but not consistent with duration of overflow events (<1 day). Effects reduced based on pulsed exposure durations of test taxon • No adverse ecological effects seen
	WWO ✓					
Vineyard Creek	Urban stormwater flows ✓	**< 2x**	+	+	+	• Ammonia and metals above guideline values • Acute toxicity not observed while chronic toxicity seen • TIE confirms ammonia, copper and zinc as sources of toxicity • Adverse ecological effects seen
	WWO ✓					

TIE – Toxicity Identification and Evaluation; ✓ indicate presence and ✗ indicates absence of stormwater or WWOs

8.2 Conclusions from the four research areas into WWOs

The WWOM program comprised four areas of research into the potential influences of WWOs on the receiving waters (Figure i and 8.3). The corresponding overarching research questions were:

- Section 2: Does spill rate and/or volume contribute to sewage-derived gross pollutants being spilt to receiving waters?
- Section 3: Is human or animal faecal contamination the dominant source?
- Section 4: Are there any contaminants of potential concern in WWOs?
- Section 5: Are adverse ecological effects apparent?

The corresponding report Sections 2, 3, 4 and 5 discuss outcomes and provide recommendations from these four research areas, respectively.

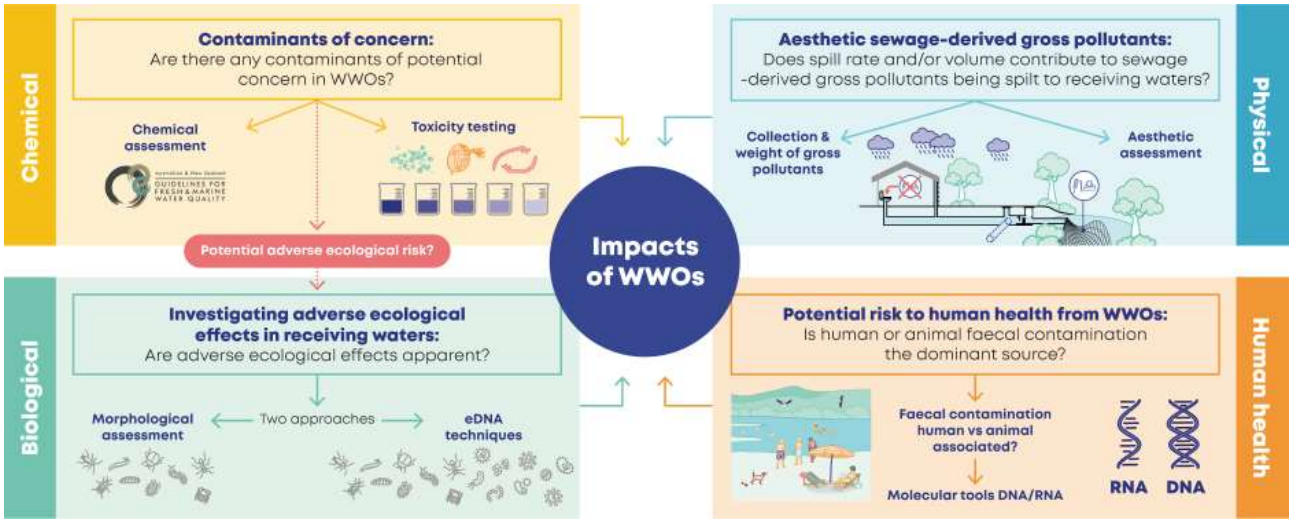


Figure 8-3: Summary of WWOM program showing components investigated and overarching research questions

8.2.1 Gross pollutants reactive and proactive management options (outcomes from Section 2)

Support for periodic community education campaigns (such as, Toilet Blockers Anonymous to support the Only flush the 3Ps: Pee, Poo and Toilet Paper! initiative) for the proactive management of gross pollutants appears necessary, as cited literature in Section 2.1 predicts an upward trend in usage of wet-wipe products.

Results from the study of incorrect disposal of gross pollutants (including wet wipes) into the wastewater system suggested that if future screening with modified trash nets (Section 2.2.4) was implemented, it would reduce the presence of sewage-derived gross pollutants and would minimise the risk of public contact with sewage-derived gross pollutants.

A potential reactive management approach is outlined in Section 2.2.4, that recommends selection of ERSs with relatively higher overflow volumes, such as from siphonic overflows that have very

high spills rates (> 1000 L/s, Figure ii), would yield the highest capture of sewage-derived gross pollutants. While selection of gravity-fed ERSs that have relatively high spill rates (> 150 L/s) and associated relatively higher volumes (Figure ii) would yield the next highest capture.

A permanent in-pipe screening approach, as discovered in reviewing literature while preparing a scientific publication for this study, may be an option for the highest volume siphonic overflows, such as those to be located within Sydney Airport lands discharging to the Mill Stream (a fringing waterway), to ameliorate the spread of sewage-derived gross pollutants through the receiving environment to reduce current manual labour clean-up effort after WWOs.

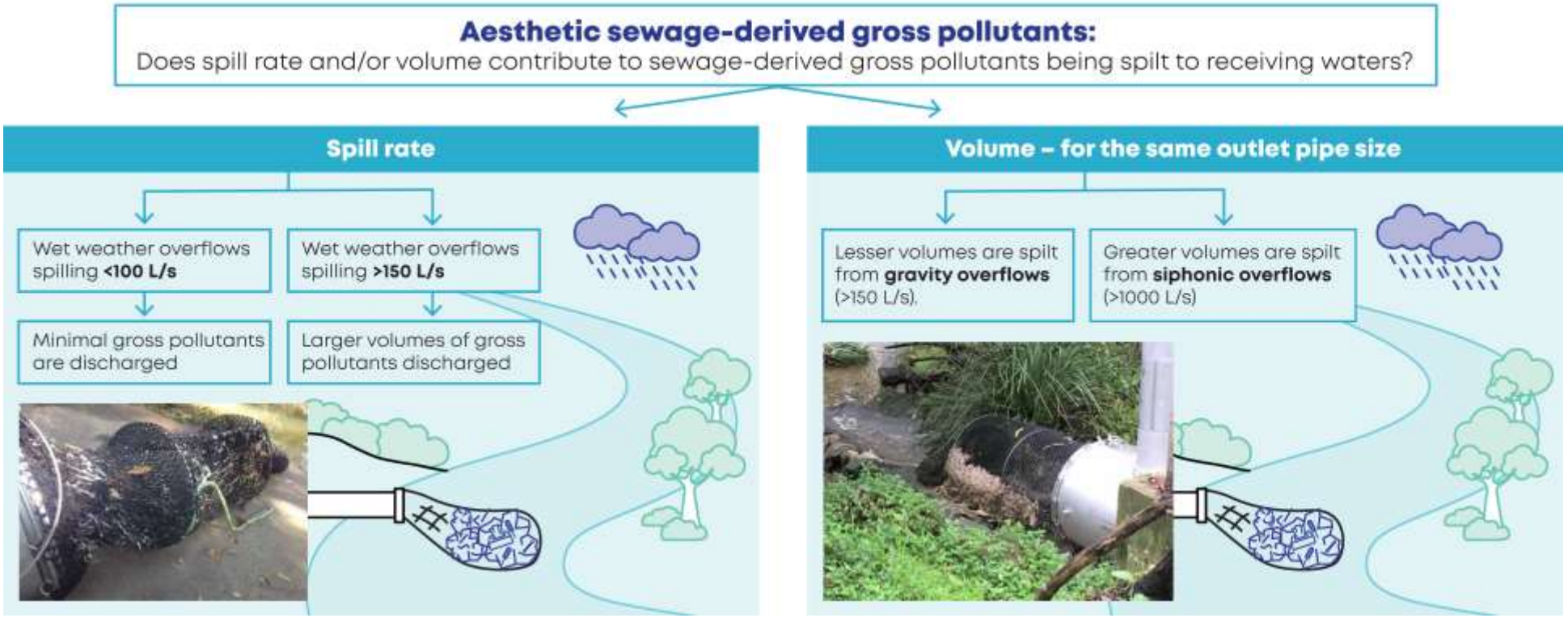


Figure 8-4: Conceptual model raised from the study of sewage-derived gross pollutants

8.2.2 Potential risks to human health from WWOs (outcomes from Section 3)

In the Sydney region, microbial source tracking with HFMGs (HF183, Lachno3, crAssphage, and PMMoV) has successfully distinguished human from non-human (animal) faecal contamination in urban estuarine waters at all 13 studied sites after WWOs from the sanitary (separate to the stormwater system) sewerage system (Section 3.2.4 and Section 3.2.8) and identified leakage of sewage under dry weather (Section 3.2.4). These outcomes indicate monitoring for HFMGs under both dry- and wet-weather conditions provides an additional tool to supplement information from Beachwatch assessments.

As discussed in Section 3.2.4, the charts of enterococci concentrations by rainfall category contained within State of NSW and Department of Planning and Environment (2021) reporting may represent a useful source of information. Charts with elevated enterococci concentrations under the 'dry' and lowest rainfall category (0.1 to 4.9 mm) (as displayed in Figure 3-9A) may represent swimming locations with similar dry-weather issues, as identified with HFMGs for Gymea Bay, or may represent abundant animal populations. Further assessment of those locations with HFMGs and animal faecal-associated marker genes would allow an understanding of the identified faecal sources contained within these enterococci data patterns. Determining the source of faecal contamination would enable cost-effective remediation at those locations with detected human faecal contamination. Hence, it would be prudent to incorporate Beachwatch sanitary inspection surveys as a first pass based on locations selected from reporting, such as State of NSW and Department of Planning and Environment (2021) to identify this type of leakage issue.

Under moderate to heavy rainfall, stormwater entry into the wastewater system is 5x to 10x more than dry-weather flows and can exceed the designed capacity of 3x average dry weather flow as outlined under Section 1.1. These conditions allow rainwater ingress and inflow to cause diluted sewer influent that exceeds dry weather capacity causing WWOs. Under these conditions, ERSs are a necessary component of the sewerage system to protect public health by stopping sewage backing up into homes and businesses (Bickford et al. 1999).

The current focus of the 2024 to 2030 WWOA capital solution planning for the four main coastal sewerage systems (North Head, Bondi, Malabar and Cronulla) is to minimise infiltration and inflow into the sewerage system. Each sewerage system is divided into Sewer Catchment Asset Management Plans (SCAMPs), Bondi – 9; Cronulla – 11; Malabar – 74; and North Head – 61.

Remediation works focusing on source control of infiltration and inflow into the sewerage system (theoretically aimed at restoring dry-weather flow capacity of pipes) are most likely to abate WWOs under light rainfall conditions from this restored pipe capacity. Together with corrected dry-weather leakage issues, these solutions should reduce illness risk on some swimming days currently impaired by human faecal contamination (Figure 8-5).

To avail of learnings from raising the QMRA (Section 3.3, Ahmed et al., 2024), as well as applying the CSIRO recommendation of using two or more HFMGs (Section 3.1), the duplex assay (crAssphage and HF183) together with PMMoV are recommended for future application. If continuous improvement work (as outlined in Section 3.5) is successful in defining another HFMG, with better performance and acceptable RBTs, then that HFMG should be implemented in place of PMMoV.

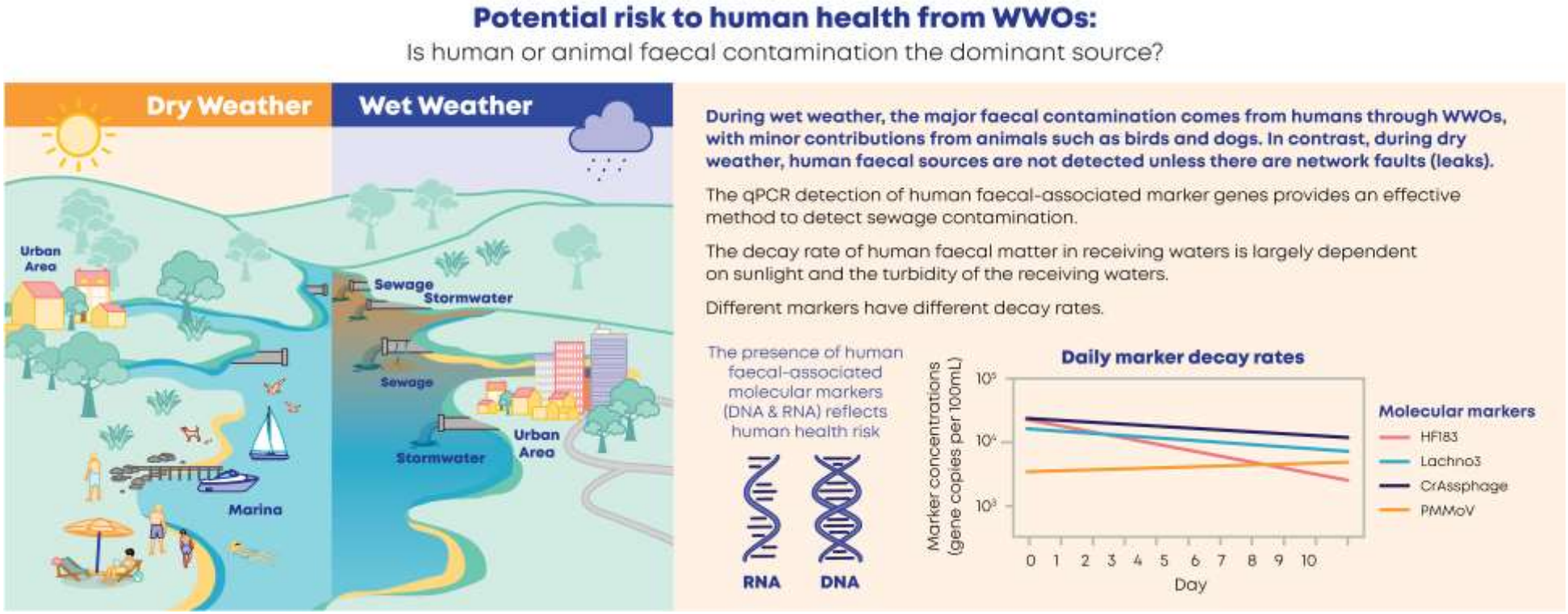


Figure 8-5: Conceptual model raised from the human health pilot studies

Outline of future application of HFMGs in the WWOA risk prioritisation methodology

To support the WWOA program's overarching aim to prioritise ERSs for source control works in a cost-effective manner, the following recommendations are made for campaign monitoring:

1. **Under dry-weather conditions**, conduct surface water column monitoring with HFMGs and animal-associated faecal marker genes (for at least two avian faecal-associated marker genes) as a first pass based on locations selected from State of NSW and Department of Planning and Environment (2021) reporting (as discussed in Section 3.2.4) that have similar data patterns of elevated enterococci, as observed to those of Gymea Bay. The assessment for avian faecal contamination is recommended as detection of avian faecal contamination was documented under both dry- and wet-weather conditions in receiving water studies of the 13 estuarine sites (Section 3.2.4 and Section 3.2.8).
2. **Undertake surface water column monitoring with HFMGs under a few differing magnitude storm events** for each SCAMP, ideally with:
    - 10 to 20 mm of rainfall in the preceding 24 hours
    - greater than 20 mm of rainfall in the preceding 24 hours
    - greater than 35 mm of rainfall in the preceding 72 hours for context of protracted storm events at a location

    Under each event, turbidty data must also be collected to allow comparison of crAsspahage results to RBTs calculated from DP or HCB, described in Section 3.3.3.

    The understanding gained under the WWOM of WWO spill behaviour suggests these rainfall amounts represent conditions on which to base ranking locations for source control that is most likely to provide information toward restoration of the designed dry-weather capacity of the sewerage system
3. Once HFMG concentrations have been returned from laboratory testing, those concentrations are compared with the Sydney specific RBTs established through the site-specific QMRA modelling conducted under the WWOM project
4. **Rank swimming locations by numerically ordering greater concentration departures from RBTs** to least outcomes from concentration departures from RBTs for each respective HFMG assessed. Another consideration in ranking locations accounts for Dr Ahmed's advocation, for example, if three assessed HFMGs return RBTs at a concentration of concern, then that location would be prioritised over a location where only two, one or no HFMG returned RBTs at a level of concern
5. in some cases, a finer dissection of locations is required and **further assessment under sanitary inspection surveys with HFMGs** may assist in establishing problem ERSs under low magnitude rainfall events

Resourcing to conduct storm event sampling

The inherent flexibility of RBTs allows HFMG concentration results collected from any day of the 11-day window to be compared (Figure 8-5). This 11-day sampling window to measure RBTs also minimises the need to expose human resources to under bad weather (unsafe) conditions on the day or day after a WWO. It also avoids the need to collect samples within 24 – 48 hrs (holding time), depending on analyte, for valid results from laboratory processing, which is a problem with autosampler collection. The sampling window afforded by the RBTs helps reduce pressure on limited human resources and be available for reactive event monitoring, such as for keeping the drinking water system safe.

It should be noted that this approach is only appropriate when rainfall has not occurred within those 11 days. If rainfall did occur within the 11 days post a storm event, then comparison to RBTs would be limited across the days up to the next rainfall event to provide a conservative assessment.

Continuous improvement obligation

Additional investment in applied research to assess the suitability of other novel emerging marker assays is recommended. This investment would conform to the pollution study (PS307) objective of continuous improvement to the prioritisation methodology. It is proposed that the following continuous improvement activities be considered:

- tomato brown rugose virus (ToBFRV Mo), bacteriophage Bifidobacterium (Bifi) and culturable bacteriophage (GB-124) HFMGs are assessed for prevalence and abundance, host specificity and host sensitivity
- development of a triplex assay should be investigated to further improve cost-effectiveness of implementation of this genomic tool. The three markers selected for the triplex assay should be compatible for input into the assay and have good specificity and sensitivity
- Further work should be done to modify and improve the adsorption extraction concentration method for viruses and bacteria to minimise the chance of not detecting HF183
- Conduct a double-blind study with CSIRO to investigate the potential benefits of digital PCR technology

This is further discussed in Section 3.5 above.

8.2.3 Contaminant of potential concern to assess risk of adverse ecological effect

The study at Vineyard Creek of an atypical node of ERSs, where ammonia was confirmed as a contaminant of potential concern (Sections 4.4 and 4.5.1), this established ammonia as an assessment indicator for potential adverse ecological risk. Ammonia concentration results evaluated against the ANZG (2018) guideline value (for 95% species protection) best represented toxicity testing outcomes. These WWOM investigations found that the risk of adverse ecological effect is reduced under increased dilution from inflow and infiltration of rainwater ingress into the sewer system (Figure 8-6, Section 4.5.1).

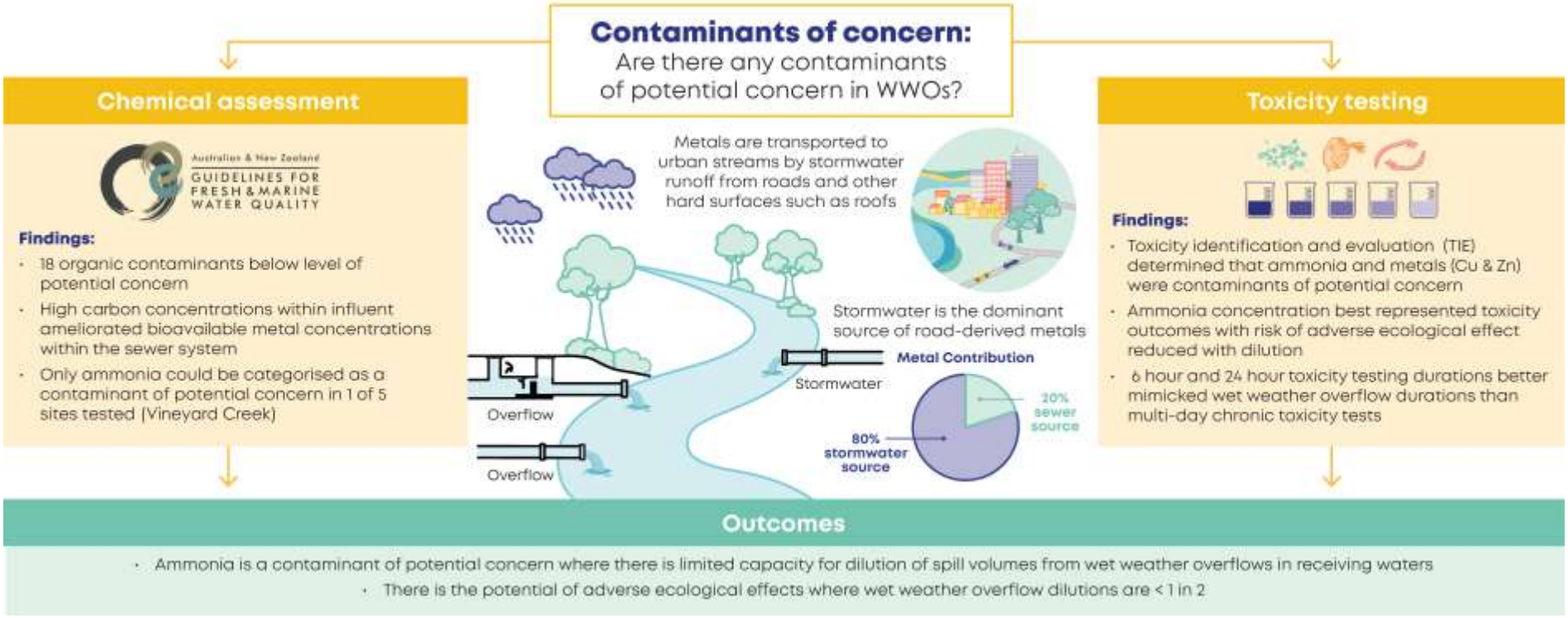


Figure 8-6: Conceptual model raised from investigations of the contaminants of concern

Follow up toxicity testing that better mimicked typical WWO spill durations (Section 4.5.2), based upon 24 h and 6 h pulse exposures of wet-weather influent, established required dilutions greater than 1 in 2 (Section 4.5.3). This indicates that the receiving waters needs to be of an adequate capacity to achieve a greater than 2x dilution of a WWO spill to remove the risk of an adverse ecological effect (Figure 8-7) based upon *Ceriodaphnia dubia* toxicity test outcomes.

Insufficient dilution of WWO spills was attributed to adverse ecological effect observed under three receiving water situations (Sections 4.7 and 5.3). These were represented by:

1. too many ERSs spilling to same point of an urban stream (Vineyard Creek)
2. oversized ERSs to a very small urban stream (Kittys Creek)
3. too many spatially separated ERSs spilling to a small urban stream reach (Girraween Creek)

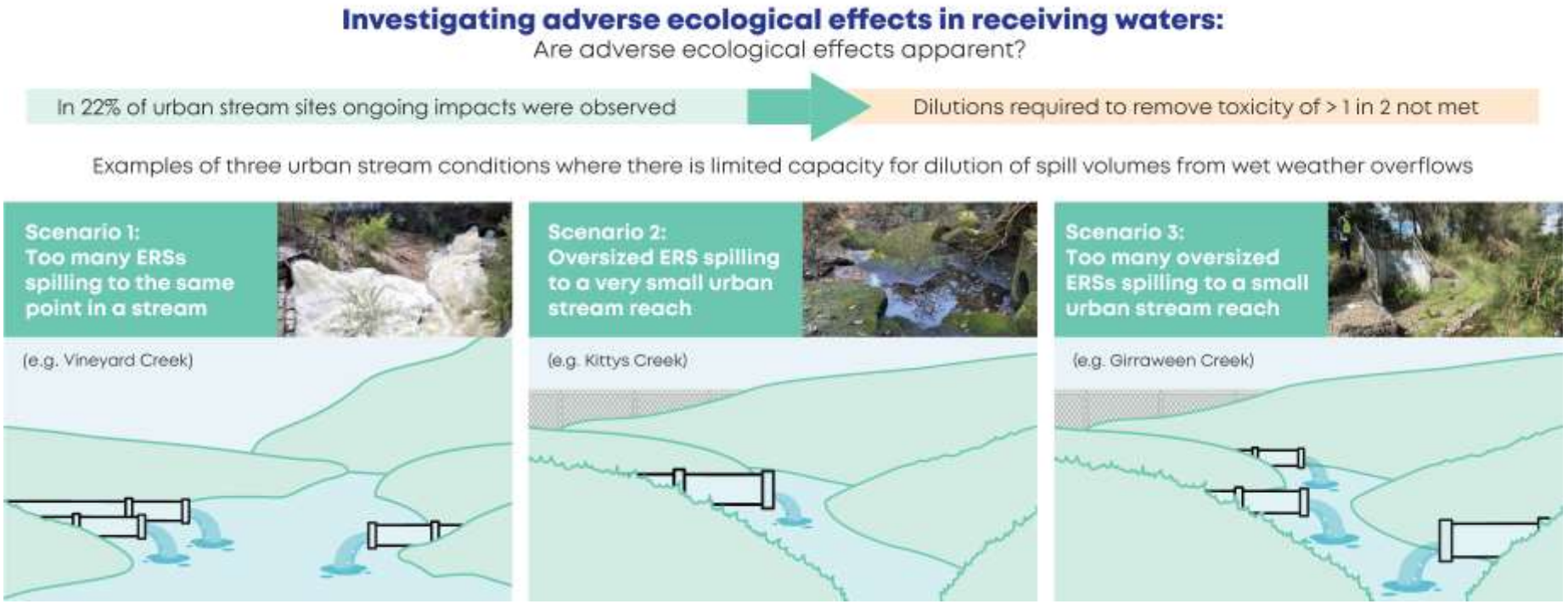


Figure 8-7: Conceptual model raised from investigating adverse ecological effects in receiving waters from WWOs

Outline of proposed future application of ammonia in the WWOA risk prioritisation methodology of the WWOA

In determining where the risk of adverse ecological effects is likely to be observed, identifying urban streams with a diminished capacity for dilution of WWO spilled volumes would be prudent. **As a first pass** in assessing this risk:

1. it is recommended that for each stream, the number of ERSs and their spatial proximity along the stream is evaluated in conjunction with the stream size, particularly width and depth under dry-weather base flow for an understanding of potential receiving water dilution capacity. Catchment walks maybe required to undertake this evaluation
2. the size of each individual ERS outlet pipe should also be evaluated against stream size
3. an understanding of WWO exposure expressed as the modelled volume (in the absence of gauged data) from all ERSs spilling to a stream or stream reach is required

Consideration of these three measures in concert will identify streams that may have a risk of adverse ecological effect. Assessing these measures would help determine situations where multiple ERS spill into a very small stream and for situations where an oversized ERS spills into a very small stream. Situations where atypical agglomerations of ERSs discharge in close proximity, such as in a stream like Vineyard Creek, should be assessed in the future program.

The morphological ecology line of evidence identified 22% of urban streams assessed had apparent ongoing disturbances attributed to WWO spills (Section 5.3). This in turn suggests up to 660 of 3000 ERSs may require further assessment as recommended below. The earlier risk assessment conducted by Bickford et al. (1999) indicated 200 of the 3000 ERSs operate in smaller rain events. This perhaps informs the level of effort that may be required in the next phase of assessment to define a revised input for the WWOA risk prioritisation methodology.

Second phase assessment. Once candidate ERSs have been identified to pose a risk of adverse ecological effect from diminished capacity for dilution out of the first-pass assessment, instrument arrays would be deployed for periods of time to assess ammonia in receiving waters before, during and after WWOs.

1. Deployment of instrument arrays that can continuously measure and provide total ammonia. For example, from one potential supplier, their ammonium sensor will automatically derive ammonia and total ammonia when used in conjunction with the pH/ORP and conductivity sensors.
2. Ideally instrument arrays are deployed to record data across a few differing magnitude wet-weather events to effectively ground truth the desktop and field walk assessment from phase 1 outlined above
3. Gathered total ammonia data would then be evaluated against ANZG (2018) guideline value for 95% species protection. The 95% guideline value is considered prudent given the stormwater delivered metal contamination has excluded metal-intolerant species from urban streams
4. Ammonia concentration departures above ANZG (2018) guideline value for each respective ERS would then be an input into the WWOA risk prioritisation methodology to represent risk of adverse ecological effect from WWOs from an ERS.

5. ERSs not taken forward into the second phase could be assigned ammonia concentrations equal to half the guideline value to allow inclusion in overall ranking across the 3000 ERSs

Deployment of instrument arrays avoids the cost associated with:

- setup of auto-samplers, collection of samples, laboratory processing of samples
- limited human resourcing to respond after wet-weather events, and
- also avoids the cost of buying and installing and maintaining sewer gauges into ERS

Downloading of data from retrieved instrument arrays and interpreting against Hydstra rainfall data would represent post processing to identify ammonia concentration maxima in wet-weather to rank ERSs as an input into the risk prioritisation methodology and in the case of dry-weather ammonia detections to further explore undetected leakage issues.

It is envisaged after an ERS or stream reach with a number of ERSs are assessed that instrument array would be redeployed to another location. This could be progressively performed to continuously improve the WWOA risk prioritisation methodology.

8.2.4 Measurement of ecological indicators on an ERS basis is cost-prohibitive

As stated in Section 5.8, predictive modelling of ecological assemblages is not an option to inform the risk prioritisation methodology going forward based upon metadata assembled under this study. As also stated in Section 5.8, it would be cost prohibitive to undertake a pair-site approach as demonstrated with the morphological indicator to assess about 660 ERSs (from the broader 3000 ERSs) based on extrapolation of the above morphological findings. Another consideration against a paired-site approach was illustrated under the study of the Darling Mills Creek system where upstream and downstream sites could not be positioned spatially close (within a few hundred metres) around the target ERS, as the ERS was situated at a stream junction with 10 km of additional stream length on one branch (Section 5.3.7). The same issue would be faced if a paired-site approach was attempted based upon DNA derived assemblage data. Hence inclusion of measurements of ecological assemblages would be an unacceptable cost for Sydney Water, with customers at the heart, and as such is not recommended in future assessments. Rather, direct measurement of ammonia with subsequent comparison against the corresponding default guideline vale of ANZG (2018) is advocated as a cost-effective approach to understanding potential adverse ecological effects from WWO spills from ERSs as proposed in Sections 4.8 and 8.2.3.

8.3 Continuous improvement of the WWOA prioritisation methodology

The WWOA risk prioritisation methodology is a comparative risk assessment for the waterway values of public health and ecosystem health. The 2024 - 2030 comparative risk assessment aligns with the pressure–state–response (PSR) model which is consistent with the approach adopted by DPE when setting waterway health objectives (NSW Department of Planning, 2022), and is based on the 'likelihood' and 'consequence' of an event occurring, which has been defined as:

- Likelihood: the chance of poor water quality that does not support public health or ecosystem health objectives
- Consequence: the extent or scale of potential water quality impacts on public health or ecosystem health objectives

The process for the 2024 – 2030 WWOA risk prioritisation methodology is detailed in the Sydney Water 2024 – 2030 WWOA Prioritisation Methodology Process Report (Sydney Water, 2022). To maximise the area over which the comparative risk could be assessed, this current methodology relies on data inputs from base datasets that provided the most complete coverage across the study area for the parameter of interest. Datasets were selected as a default where no higher quality, detail or resolution data were available.

The current WWOA risk prioritisation methodology inputs are also sourced from water quality models, which are effectively creating 'a model of a model'. WWOM has identified that direct measures can replace these modelled outputs. Potential integration of outcomes for future improvements to the WWOA risk prioritisation tool are discussed below.

Recommended improvements to the WWOA risk prioritisation methodology

Public health

The 2024 -2030 WWOA risk prioritisation methodology applies the 95th percentile enterococci concentration from Beachwatch (DPIE, 2020). Beachwatch data, as representative of real-world data, were adopted as the primary data source and water quality modelled data were incorporated where Beachwatch data were not available.

As discussed in Section 3, enterococci measurements cannot differentiate between human and other sources (other animals) of faecal contamination, or with those naturally growing in the environment. WWOM has developed a QMRA to simulate the health risks of GI illness for four HFMGs (HF183, Lachno3, crAssphage, and PMMoV). The health risk from the QMRA is presented as risk-based thresholds (RBTs) for each of the four HFMGs assessed in the human health pilot study. Outcomes from the WWOM proposes that the 95th percentile enterococci concentration be replaced with an assessment of HFMG concentration results of the four HFMGs against the risk-based thresholds (Section 3.3) for estuarine waters in the revised risk prioritisation methodology.

Ecosystem health

Currently, modelled 80th percentile chlorophyll-a (Chla) concentration data from Sydney Water's RMA water quality models raised from data collected at limited estuarine sites as the input used as

a proxy to rank the likelihood of poor water quality. The EPA identified in their review of the 2024 – 2030 WWOA risk prioritisation methodology, that the use of modelled Chla was not a preferred stressor input.

The United Kingdom has implemented direct measurements for ecosystem health. Their Storm Overflows Discharge Reduction Plan (Defra, 2023) for the identification and measurement of ecological health impacts, applies the Urban Pollution Management Fundamental Intermittent Standards (FIS) or 99th percentile standards for ammonia and dissolved oxygen downstream of discharge points (inland waters) as the measure of ecological impact. WWOM has established ammonia as a chemical of potential concern, which mirrors the UK's direction. Dissolved oxygen measurements from samples collected under wet-weather conditions for toxicity testing of influent and companion receiving water samples were above ANZECC/ARMCANZ (2000) recommended thresholds of > 6 mg/L or > 80-90% saturation (Section 4.5.1, Table 4-2) suggesting this physical parameter was not a concern.

Ammonia concentration results can be evaluated against the ANZG (2018) guideline value (for protection of 95% of species) to represent risk of adverse ecological effect from WWOs. WWOM has established that ammonia as a stressor input (Sections 4.4 and 4.5.1) would be a suitable improvement to the prioritisation methodology and would replace modelled Chla as the stressor indicator for potential adverse ecological risk. It is proposed that ammonia data be collected using the methods outlined in Section 8.2.3 from both freshwater streams and estuarine sites near ERS in low dilution settings. This will remove the need to upgrade existing water quality models.

Benefits of direct measures

There are several benefits of using direct measures as the input in the WWOA risk prioritisation methodology:

- Direct measurement removes a level of uncertainty from the WWOA risk prioritisation methodology as highlighted by the EPA. Direct measurement improves sensitivity resulting in a more robust prioritisation methodology.
- Collection of ammonia and HFMG samples, at the ERS or at key locations along a waterway (both freshwater streams and from estuarine locations), will increase the granularity that is currently not available via modelled outputs that were raised from a limited suite of estuarine sites.
- Integrated catchment models (ICM) are logistically complicated to develop, they take considerable time, resources and financial commitment as well as ongoing maintenance and validation costs. Direct measurement of ammonia and HFGMs would provide real-world data and remove the need for a modelled water quality data input into the risk priortisation methodology
- Direct measurement removes the issue of human resource scarcity in physical collection of samples from auto-samplers within sample processing holding times (to enable vaild laboratory analysis results) during wet-weather events, specifically:
    - collection of MST data can take place up to 10 days after the wet-weather event to then be compared to the RBTs)

- in-situ ammonia continuous data collection instrumentation would remove the need to mobilise staff to collect samples after wet-weather events and analyse samples in the laboratory. Instead field visits can be undertaken under drier weather conditions to avoid safety risks associated with wet-weather sample collection (for example, flooding, high winds that prevent boating, and dangerous river level, wave conditions and road driving conditions)

- To directly asess ammonia with in situ instrumentation at high-risk ERS sites the investment cost is 1.5x less than the proposed ICM models. This allows for cost savings for our customers and/or reuse of funds into the implementation of solutions.
- The NSW Government has proposed that the Beachwatch Program in the Sydney Coastal area will transition from a fully funded service model to an opt-in Beachwatch partnership model with local councils. Future monitoring of current Beachwatch sites will be dependent on the relevant councils to fund and implement. As such, the broadscale Beachwatch forecasting models and reporting for swim sites currently used as an input indicator for the WWOA risk prioritisation methodology in uncertain. Transitioning to a direct measure using RBTs derived from HFMG assessments would future proof against reliance on the Beachwatch Program and any potential for any information gaps resulting from changes.

References

DPIE, 2020. Protocol for Assessment and Management of Microbial Risks in Recreational Waters. Department of Planning, Industry and Environment, NSW. Accessed 7 June 2024

NHMRC, 2008. National Health and Medical Research Council. Guidelines for Managing Risks in Recreational Water. Australian Government Publication Services ISBN 1864962666.

Sydney Water, 2022. 2024 – 2030 WWOA Prioritisation Methodology Process Report, Document number: SW 121 04/22 (D0002152)

8.4 Application under other Sydney Water programs

8.4.1 Studies to inform continuous improvement under WWOA

Refinements to the existing MST toolbox for monitoring surface waters with HFMGs are outlined in Section 3.5, while recommendations for enhanced in-house capability with animal faecal-associated marker genes are outlined in Section 3.6. Application of direct measurement of HFMGs is outlined under a stepped approach in Sections 3.4 and 8.4 for future monitoring under the WWOA to inform the risk prioritisation tool.

The identified direct measurement of ammonia enables comparison of concentration results against the ANZG (2018) 95% protection of species default guideline value to assess risk of adverse ecological effects. Application of monitoring with ammonia is outlined in a stepped approach under Sections 4.8 and 8.4 for future monitoring under the WWOA to inform the risk prioritisation tool.

8.4.2 Applicability of findings in other Sydney Water programs

Sydney Water Aquatic Monitoring program (SWAM)

The abovementioned MST single molecular target approach using HFMG and animal faecal-associated marker genes is well advanced and is already being implemented into monitoring by Sydney Water.

In contrast, the outcomes of the WWOM DNA pilot studies clearly indicate that further evaluation of Biomonitoring 2.0 ‘rebuild’ and ‘renovation’ approaches with multiple molecular targets are required. A series of recommendations to undertake further evaluation of the application of DNA based taxonomy with paired-site assessments under the SWAM program is outlined in Section 5.9. to establish if at least equivalent or better assessment of adverse ecological effects are afforded by community-DNA or environmental DNA.

These recommendations are outlined under the two Biomonitoring 2.0 approaches (Section 5.9). A ‘renovate’ approach looks at implementing the existing biotic index tool SIGNAL-SG based upon taxonomy from DNA samples using the macroinvertebrate size-class animals that are the basis of the traditional morphometric taxonomy method. While a ‘rebuild’ approach explores the applicability of other taxonomic groups afforded by DNA, such as bacteria, invertebrate meiofauna or diatoms, on which to base new biotic indices.

To enhance outcomes of the above pilot studies, a gap analysis should be conducted between DNA-barcoded taxa hosted in the AIA and the common taxa of the SWAM Section 5.5.2. A further round of DNA barcoding should be undertaken to infill those defined gaps with the view to implementation of the ‘renovate’ Biomonitoring 2.0 into that long-term regulatory project. Some recommendations on further comparison studies with morphometric data are outlined at the end of Section 5.5.3, while considerations outlined under topic heading ‘Application in biomonitoring’ in Section 5.5.1 need to be applied.

Dry weather sewer overflow investigations

Environmental performance of dry-weather sewer overflows (DWSOs) is managed under the Sydney Water Enterprise Initiatives (T1.2). DWSOs include sewer influent spills due to a choke in

the sewerage network or from breaks within this network. Monitoring of ammonia concentrations results against the ANZG (2018) 95% protection of species default guideline value to assess risk of adverse ecological effects is already in place within DWSO investigations. Findings under the contaminants of concern studies (Sections 4.4 and 4.5.1) support continued use of this chemical marker in DWSO investigations (Section 6.1).

Already implemented in the DWSO investigations is the use of HFMGs in post cleanup assessments where elevated enterococci concentrations are returned. Further development of in-house laboratory capability with animal faecal-associated marker genes (Section 3.6) has commenced to support DWSO investigations to assess if animal faecal matter may be responsible for elevated enterococci concentrations, such as in the case of waterfowl observed at cleanup sites. Use of both human- and animal-associated faecal marker genes would add certainty to determine if a human faecal source remains, and further repairs are warranted. If animal faecal contamination was indicated this would then allow human resourcing to be prioritised elsewhere (Section 6.1).

Another potential tool to establish sewage contamination (DWSO) is provided by the suite of eight organic chemical (acetaminophen, ibuprofen, metformin, sucralose, theobromine and three benzotriazoles) markers (tracers) (Section 6.2). The presence of these chemicals that have short half-lives of up to two days in the water column would indicate recent / active sewage contamination.

Under the WWOM a dry-weather leakage threshold (above 100,000 gene copies per gram) for HFMG sediment concentrations (from the duplex assay HF183 and CPQ_056 markers) was apparent. Use of HFMG sediment concentrations for future assessments of DWSO leakage from the Sydney Water sewerage system may be another possible tool to employ in investigations (Section 6.3). A further pilot study has been advocated to gather more persistence data for the HF183 and CPQ_056 HFMGs to assist in understanding the decay of DNA from these microorganisms in sediment. That would then allow a better understanding of the age of detected sewage contamination (Section 6.3).

9 Acknowledgements

Core Project team

Michele Cassidy (Project Manager, Wastewater & Environment): Project administration; funding acquisition; conceptualisation; visualisation; writing (primary) - review & editing; writing - original draft

Colin Besley (Technical Lead and liaison between expert peer panel and engaged collaborators, Monitoring Design & Reporting): Conceptualisation; methodology; visualisation; validation; formal analysis; data curation; writing (primary)- original draft; writing - review & editing; supervision

Nathan Harrison (Project and Technical Support and project administration, Wastewater & Environment): Writing - review & editing; writing - original draft; resources; investigation

Field and Laboratory Investigations, Resources and Validation

Field Sampling and Testing

The Field Sampling and Testing team provided invaluable support to the WWOM project, in particular the agility to respond to weather dependent event sampling and in the building, procurement and testing of novel sampling equipment and methods. The team would like to specifically acknowledge the contribution of Elanor Young and Tom McLoughlin, and the below staff:

Alex Michie; Susan Schwartz; Lee Hayes; Danielle Clarkson; Dan Nolan; Dinesh Kunasegaran; Elizabeth Fabian; Bevan Yiu; Amelia Hanns; Bobby Missaghian; Courtney Keep; Matthew Vernon; Leon Ow; Alexandra Grzechnik; Luke Pedini; Daniel Bradley; Andy Johnston; Daniel Whitley; Adam Clark; Liam Baker; Toby Patti; Angelito Jader; Renee Wilsher; David Connolly; Yehuda Shalem; Aaron Duncan; Liam Doyle; Maeve Malikides; Evan Giovanni; Bailey Davis; Binaya Sharma, Erdal Ozen

Hydrometrics

The hydrometrics team deployed and monitored over 180 sewer gauges, rain gauges and autosamplers, providing invaluable data to the WWOM project. This work provided great insights into not just the real performance of the overflows but additional benefit to the wastewater models and even identified dry weather surcharges. The project team would like to specifically acknowledge the contribution of Adam Wadey and Leigh Harris and the staff listed below

Cameron Loader; Chris Tornatore; Darren Rees; Jeremy Hearfield; Ken Sutton; Leon Whittingstall; Phillip Matthews; Chris Harris; Dennis Fruci; Thomas Hoppe; William Denny; James Slattery; Shaun Henderson; Shaun Haywood

West Ryde Laboratories

The West Ryde Laboratories provided invaluable support to the WWOM project. In particular the contribution of the Aquatic Ecology, Trace and Microbiology teams in processing a huge number of

samples. This also required the development and testing of new methodologies and processes. The team would like to specifically acknowledge the contribution of Sudhi Payyappat, Mark Hopkins, Jamie Devine and Justice Esoun-Nyarkoh and the staff listed below:

Catherine Hooper; Harrison Ritchie; Monica Mulder; Justine Djajadikarta; Annalisa Humphrys; Hamish Pritchard; Luke Clay; Greg M^{c}Donald; Zoe Goold; James MacNamara; Simeon Pocha; Xanthe Petridis; Aidan Salvatore; Samuel Kroon; Sabina Kafle; Jin Kato; Iandiana Strachan; Henry Nemme; Erica Leal; Deniz Yildiz; Dorota Chamos; Carol Sutherland; Anna Ly; Mitchell Parker; David Choueifati; Haley Bambridge; Megan Dreyer; Greg Stonehouse

Technical review of journal papers and synthesis report

Merran Griffiths, Kaye Power, Peter Cox

Project support

Natasha Abulafia – initial conceptualisation; funding acquisition

Susan Westcott; Tory Walker – REF preparation and environmental advice/support

David Farrugia – project resources and support

Tiffany Chen – GIS support

Management team support for resourcing and project sponsorship

Yvonne Sinanovic; Catherine Cunningham; Iain Fairbairn; Julia Bartlett; Paul Higham; Kris Mobberley; Paul McCartan; Craig Earl

We acknowledge that many more people have contributed to the WWOM program over the years that we have not been able to include above. Everyone involved has added to the success of the project, and we thank them for their contribution.

10 Appendices

10.1 Appendix 1 - Freshwater morphological macroinvertebrates

As outlined in Section 5.1, Sydney Water Laboratory Services NATA-accredited facility has long-term capability (1990 onward) in assessing freshwater (benthic) macroinvertebrates based upon morphological (external physical) characteristics forming one line of evidence. Peer panel advice, when indicators were explored, suggested this existing high capability with morphological macroinvertebrate assessments should be included as a line of evidence for assessing urban freshwater streams under the WWOM. This line of evidence was assessed on a taxonomic assemblage (community structure) basis with multivariate statistical approaches or with a biotic index as an evaluation of stream health at each study site.

10.1.1 Graphcial overview

The following four graphical summaries were raised for each study stream:

- control chart of SIGNAL-SG biotic index scores by site and collection periods
- shade plot of taxonomic assembalage data by site and collection periods with SIGNAL-SG grades annotated for each taxa (where available from Chessman et al., 2007)
- metric Multidimensional Scaling ordination plot based on taxonomic assembalage data
- tree diagram from classification analysis based on taxonomic assembalage data.

These graphical summaries are presented in Figure 10-1 to Figure 10-23 for each of the study streams.

The primary assessment of scores calculated from the SIGNAL-SG biotic index was done visually using plots along the lines of a process control chart for ecological monitoring presented by Burgman et al. (2012) to display information in a simple, practical and scientifically credible way. This style of control chart illustrates temporal trends and allows interpretation of data against background natural disturbance and variation of the respective streams. In these control chart plots, the range of each site period has the mean plotted together with error bars of ± one standard deviation of the mean, as recommended by ANZECC (2000) for basing ecological decisions. These ± one standard deviation of the mean formed ranges of stream health for period displayed. These charts were plotted on collection period basis. These comparisons had three possible outcomes:

1) Mean downstream stream health was within the range recorded for the upstream site of the same collection period suggesting no impairment from inflow sources between sites
2) Mean downstream stream health was lower than that of the upstream site suggesting impairment from inflows between sites
3) Mean upstream stream health was lower than that of the downstream site suggesting impairment from inflows above the upstream

Visual summaries from multivariate assessment of taxonomic assemblages were then inspected for clear separation of downstream site samples from upstream sites to look for supporting evidence to the control chart plots while keeping in mind potential natural influences such as sub-surface geology (Section 5.2.2), natural mesohabitat differences between sites (Section 5.2.3) and stormwater influences (Section 5.2.1) on taxonomic assemblages.

10.1.2 Statistical overview from ANOSIM testing

To assess for differences in taxonomic assemblages between upstream and downstream sites of each creek, ANOSIM R values of pairwise comparisons were inspected, as large R values (close to unity) are indicative of complete separation of the groups, whereas small R values (close to zero) imply little or no segregation (Clarke et al., 2014).

Natural heterogeneity from substrate differences within the same mesohabitat type between sites are also a known influence on community structure data within the Sydney region. Besley and Chessman (2008) documented site-specific macroinvertebrate assemblages of near-pristine streams situated in National Parks or bushland reserves of the Sydney region. They found assemblages for the pool-edge water mesohabitat significantly differed between the Georges River and O'Hares Creek for both freshwater streams (R = 0.44, p = 0.001) that were 4 km apart on the same stream system sampled each six months between 1997 and 2003 (Besley and Chessman, 2008). On O'Hares Creek another four sites at three different distances from source of stream were sampled on 4 to 6 occasions between 1995 and 1997 with significant ANOSIM R-values returned between 0.38 and 0.58 (13 and 25 km apart) (Besley and Chessman, 2008). A non-significant ANOSIM R-value was returned for the pair of sites 0.3 km apart.

A similar non-significant outcome was observed for a pair of sites (0.2 km apart) on Erskine Creek sampled on 4 occasions between 1995 and 1997, although for the riffle habitat (R = 0.46) a significant outcome was observed between 0.2 km apart between a boulder section versus a cobble section (Besley and Chessman, 2008). These ANOSIM test R-value outcomes from pairwise tests of near-pristine stream sites across varying spatial differences provide context to assess R-values against from pairwise comparisons of urban stream macroinvertebrate assemblages assessed under the WWOM pilot study. Taking note of this context in interpretation of ANOSIN results avoids falsely declaring adverse ecological effects when in fact community assemblage differences are from differences in natural heterogeneity in substrate between sites.

Further statistical assessments

After inspection of the graphical overview together with ANOSIM results and consideration of potentially low diltuion situations (as described in Section 4.7) a site within each of following listed streams were considered to have adverse ecologcial effects from WWO spills:

- Vineyard (Section 5.3.1)
- Girraween (Section 5.3.4)
- Kittys (Section 5.3.5)
- Fenchs (Section 5.3.6)
- Blacktown (Section 5.3.8)

In the case of Avondale Creek, a catchment walk identified a network fault that appeared to be responsible for the apparent adverse ecological effect as opposed to WWOs (Sections 4.7 and 5.3.3). Further statistical analysis was undertaken on morphometric data from the sites of these six streams at sections listed above and for Buffalo Creek (Section 5.3.2) and for the Darling Mills creek system (Section 5.3.7) as the downstream sites of these two streams were assessed under toxicity testing (Section 4.5.3).

Table 10-1: ANOSIM pairwise test summary by stream for morphometric data across 17 collection periods (approximately 8 weekly) October 2018 to April 2021

Stream	Position	Site code	2021-22
Cowan	upstream	COWSU	R = 0.17
Cowan	downstream	COWSD	
Hornsby Cockle	upstream	COCU	COCU versus COCD R =0.13
Hornsby Cockle	upstream	HORNU	HORNU versus COCU R = 0.37
Hornsby Cockle	downstream	COCD	HORNU versus COCD R = 0.39
Carroll	upstream	CARU	R = 0.27
Carroll	downstream	CARD	
Frenchs	upstream	FRENU2	R = 0.52
Frenchs	downstream	FREND	
Devlins	upstream	DEVU1	DEVU1 versus DEVU2 R = 0.24
Devlins	upstream	DEVU2	DEVD versus DEVU1 R = 0.34
Devlins	downstream	DEVD	DEVD versus DEVU2 R = 0.56
Avondale	upstream	AVONU	R = 0.49
Avondale	downstream	AVOND	
Gloucester / Lane Cove	upstream	CONGU	LANEU versus LANED R = 0.01
Gloucester / Lane Cove	upstream	LANEU	CONGU versus LANEU R = 0.85
Gloucester / Lane Cove	downstream	LANED	CONGU versus LANED R = 0.81
Rudder	upstream	RUDU	R = 0.40
Rudder	downstream	RUDD	
Blackbutt	upstream	BBUTU	R = 0.72
Blackbutt	downstream	BBUTD	
Kittys	upstream	KITU	R = 0.52
Kittys	downstream	KITD	
Buffalo	upstream	BUFFU1	BUFF1 versus BUFF2 R = 0.01
Buffalo	upstream	BUFFU2	BUFFD versus BUFF1 R = 0.23
Buffalo	downstream	BUFFD	BUFFD versus BUFF2 R = 0.14
Excelsior Darling Mills	upstream	EXCELU1	EXCELU1 versus EXCELU2 R = 0.31
Excelsior Darling Mills	upstream	EXCELU2	DARLD versus EXCELU1 R = 0.36
Excelsior Darling Mills	downstream	DARLD	DARLD versus EXCELU2 R = 0.46
Hunts	upstream	TRIBHUNU1	TRIBHUNU1 versus TRIBHUNU2 R = 0.54
Hunts	upstream	TRIBHUNU2	TRIBHUND versus TRIBHUNU1 R = 0.38
Hunts	downstream	TRIBHUND	TRIBHUND versus TRIBHUNU2 R = 0.10
The Ponds	upstream	PONU	PONU versus PONU2 R = 0.37
The Ponds	upstream	PONU2	POND versus PONU R =0.49
The Ponds	downstream	POND	POND versus PONU2 R =0.53
Vineyard	upstream	VINU1	VIND versus VINU1 R = 0.57
Vineyard	upstream	VINU2	VIND versus VINU2 R = 0.70
Vineyard	downstream	VIND	VINU1 versus VINU2 = 0.31
Blacktown	upstream	BLACU1	BLACU2 versus BLACD R = 0.13
Blacktown	upstream	BLACU2	BLACU1 versus BLACD R =0.80
Blacktown	downstream	BLACD	BLACU1 versus BLACU2 R = 0.81
Trib Blacktown	upstream	TRIBBLACU	R = 0.24
Trib Blacktown	downstream	TRIBBLACD	

Stream	Position	Site code	2021-22
Lalor	upstream	LALU	R = 0.15
Lalor	downstream	LALD	
Girraween	upstream	GIRU	R = 0.56
Girraween	downstream	GIRD	
Toongabbie	upstream	TOOU1	TOOU1 versus TOOD R = 0.12 TOOU2 versus TOOD R = 0.57 TOOU2 versus TOOU1 R = 0.62
Toongabbie	upstream	TOOU2	
Toongabbie	downstream	TOOD	
Bunbury Curran	upstream	BUNU	R = 0.03
Bunbury Curran	downstream	BUND	
Brickmakers	upstream	BRICKU	R = 0.50
Brickmakers	downstream	BRICKD	
Prospect	upstream	PROSU	R = 0.24
Prospect	downstream	PROSD	

References

Besley, C.H., Chessman B.C. 2008. Rapid biological assessment charts the recovery of stream macroinvertebrate assemblages after sewage discharges cease. Ecol. Indic. 8, 625-638.

Burgman, M., Lowell, K., Woodgate, P., Jones, S., Richards, G., and Addison, P., 2012. *An endpoint hierarchy and process control charts for ecological monitoring* in (eds) Lindenmayer, D., and Gibbons, P. Biodiversity Monitoring in Australia. CSIRO Publishing, Collingwood, Australia.

Chessman, B.C., Williams, S.A., Besley, C.H. 2007. Bioassessment of streams with macroinvertebrates: effect of sampled habitat and taxonomic resolution. J. North Am. Benthol. Soc. 26, 546-565.

Clarke, K.R., Gorley, R.N., Somerfield, P.J., Warwick, R.M., 2014. Change in Marine Communities: An Approach to Statistical Analysis and Interpretation, 3rd ed. PRIMER-E, Plymouth, U.K.

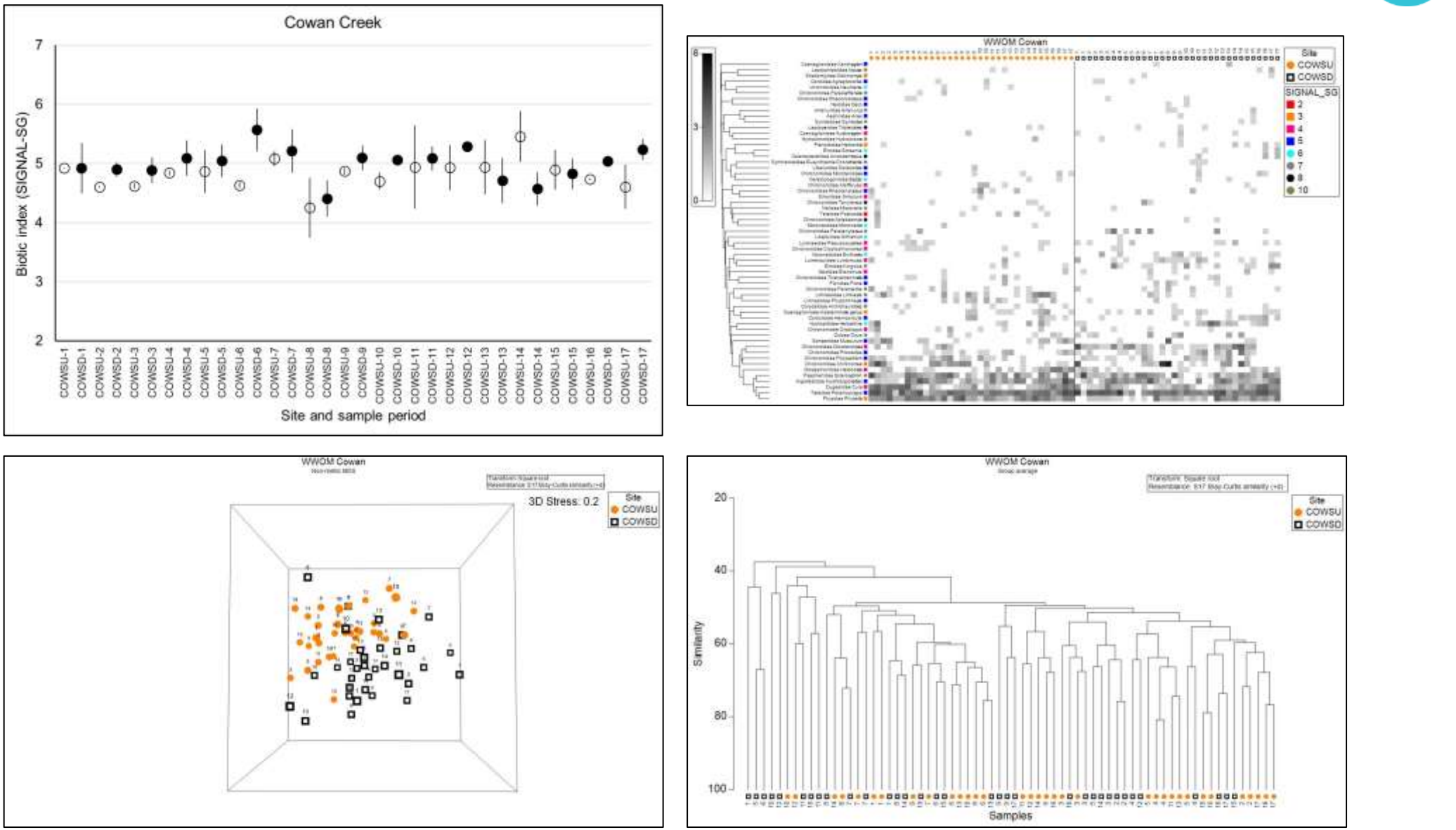
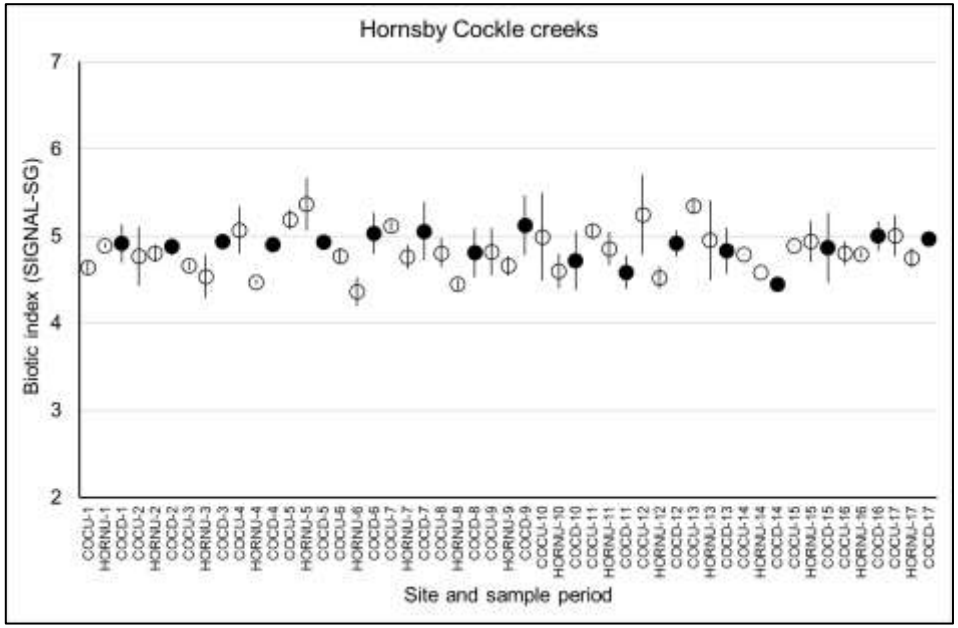
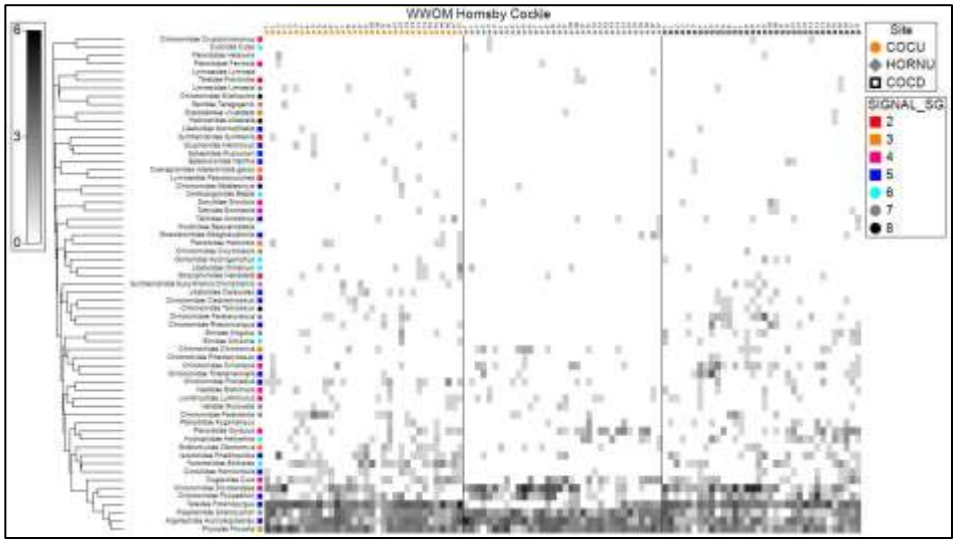
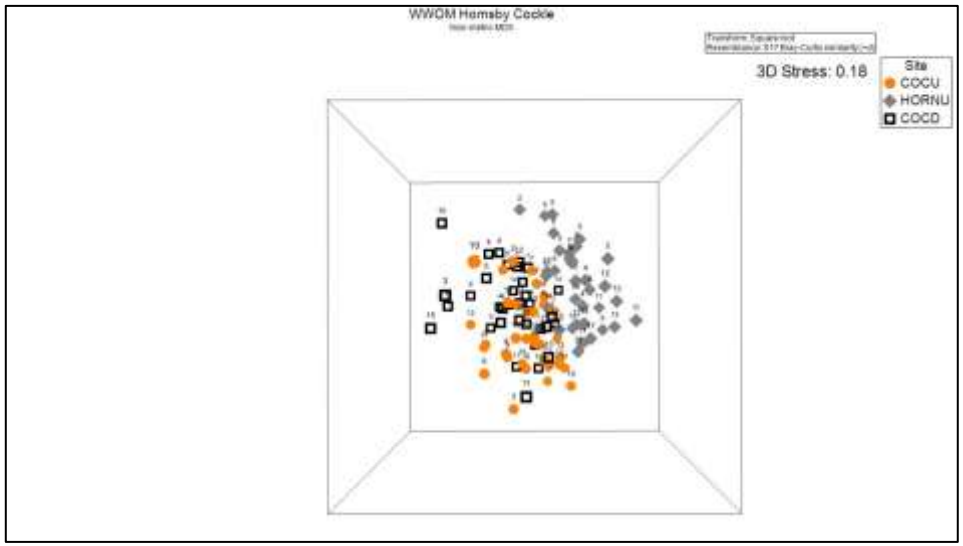


Figure 10-1: Overview of morphometric macroinvertebrate data for Cowan Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot







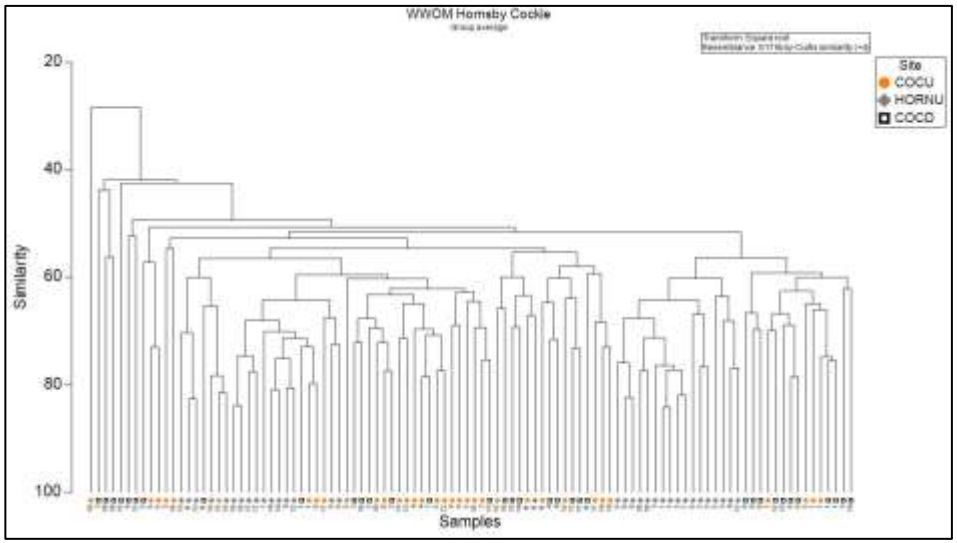
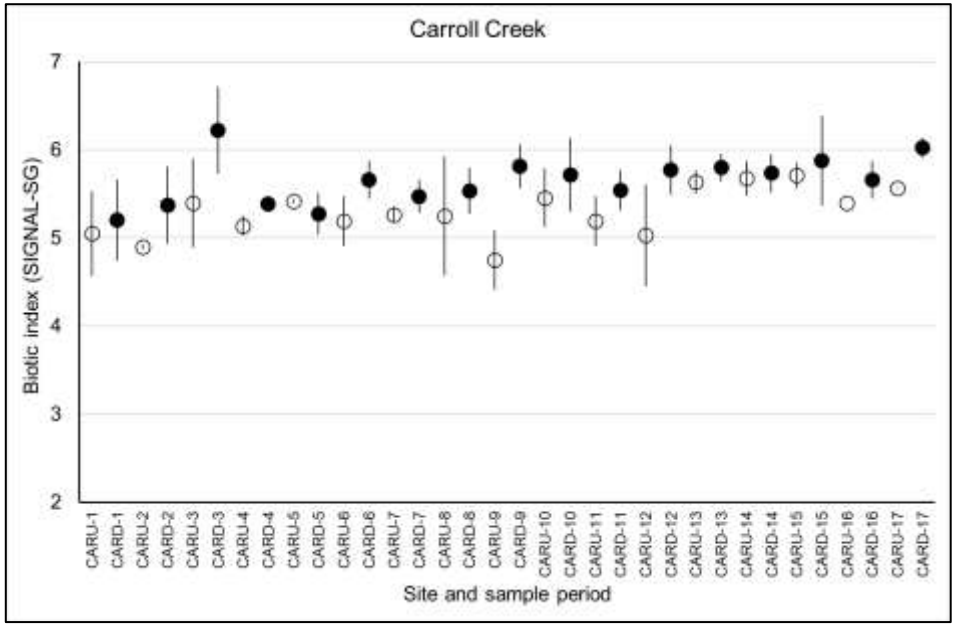
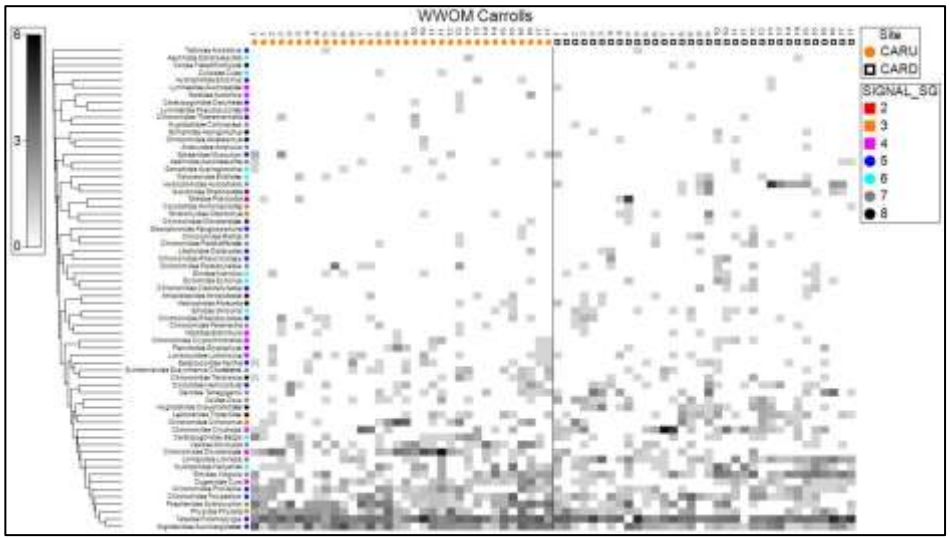
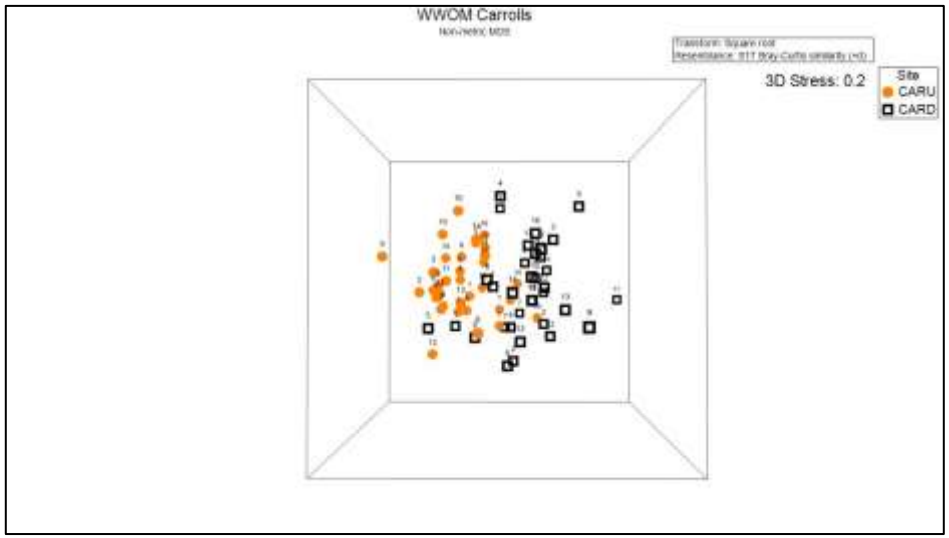


Figure 10-2: Overview of morphometric macroinvertebrate data for Hornsby and Cockle creek system

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot







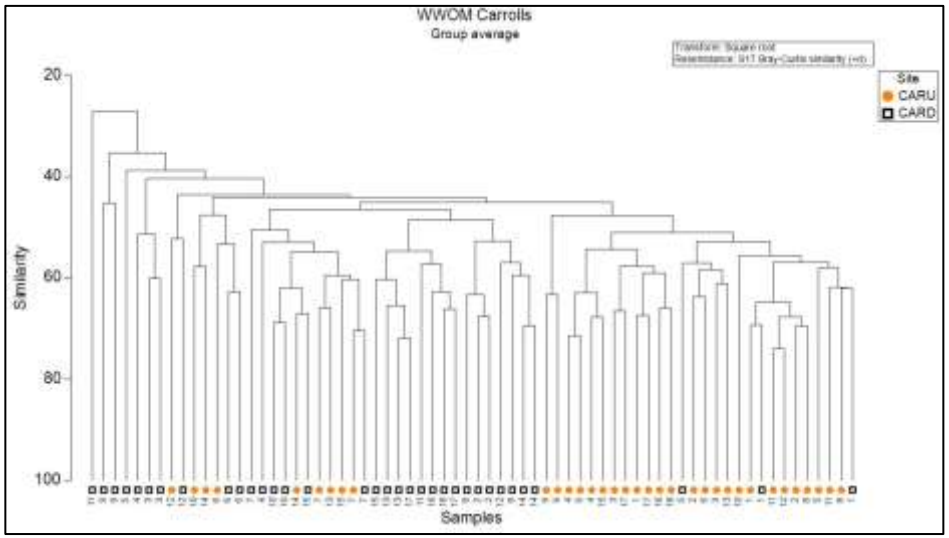
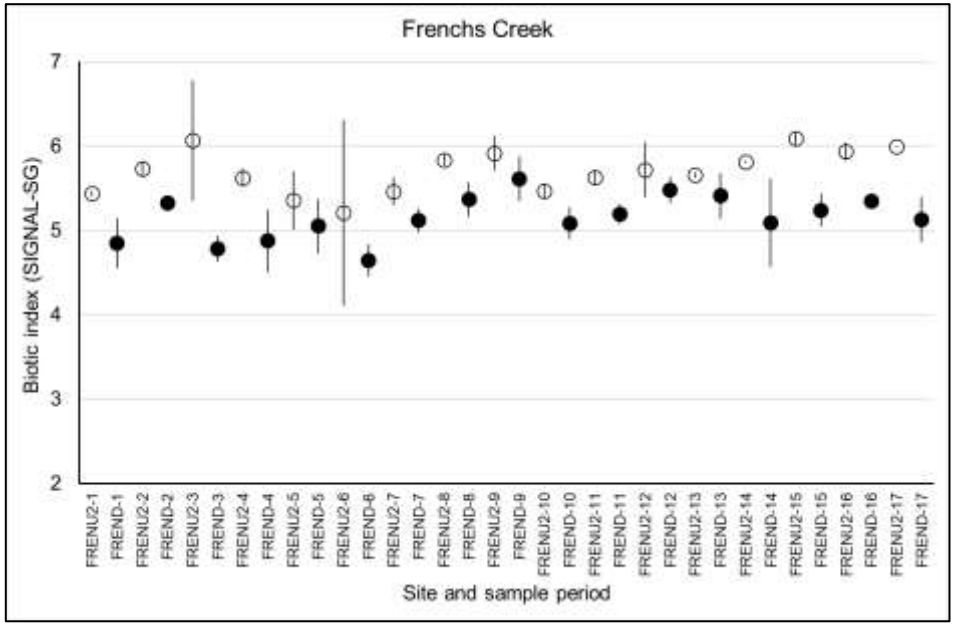
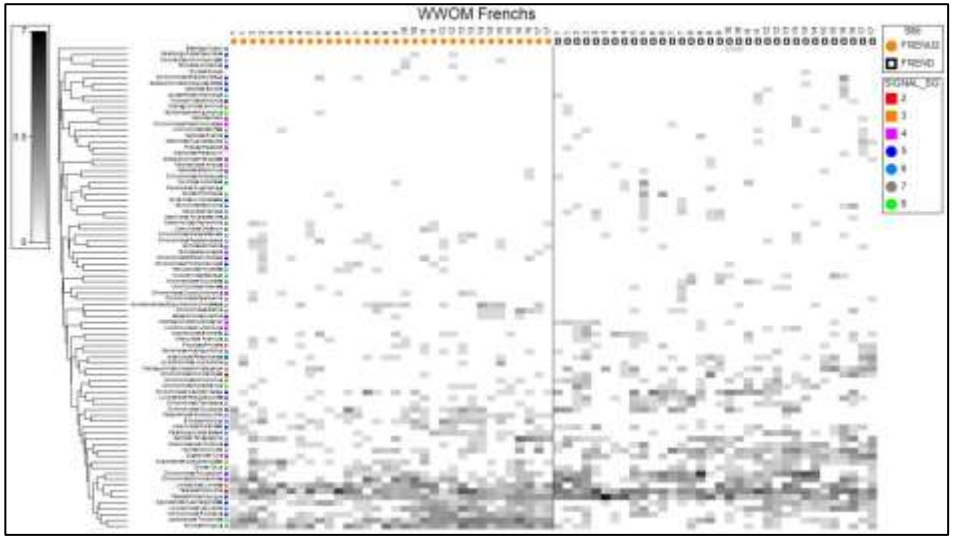
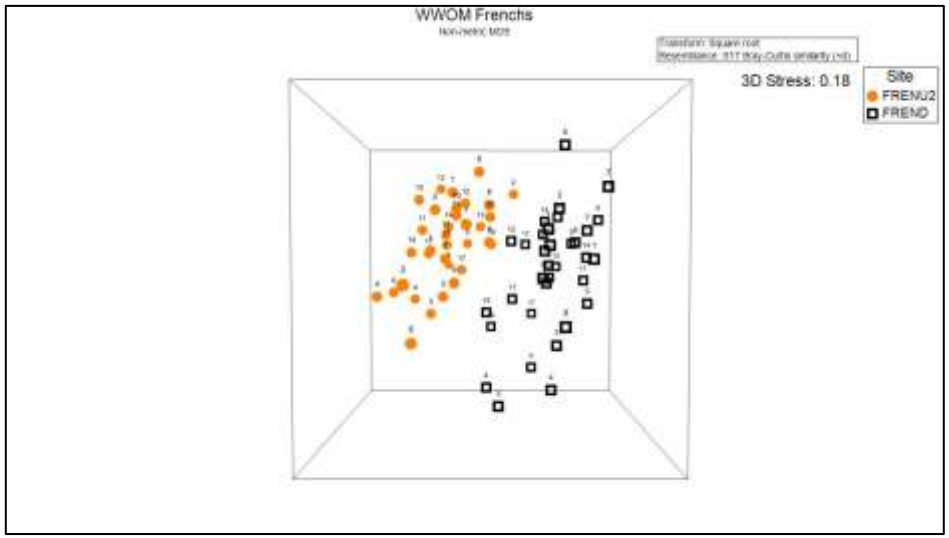


Figure 10-3: Overview of morphometric macroinvertebrate data for Carroll Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot







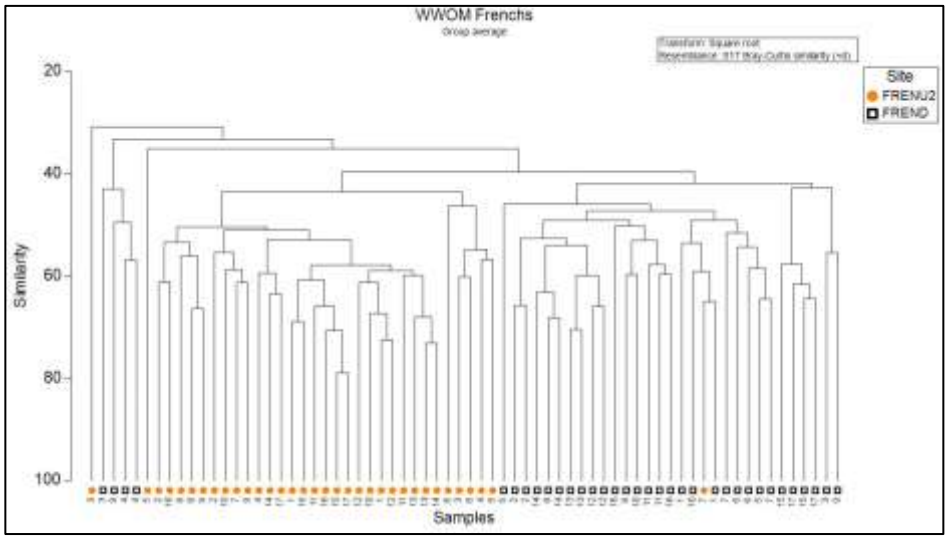
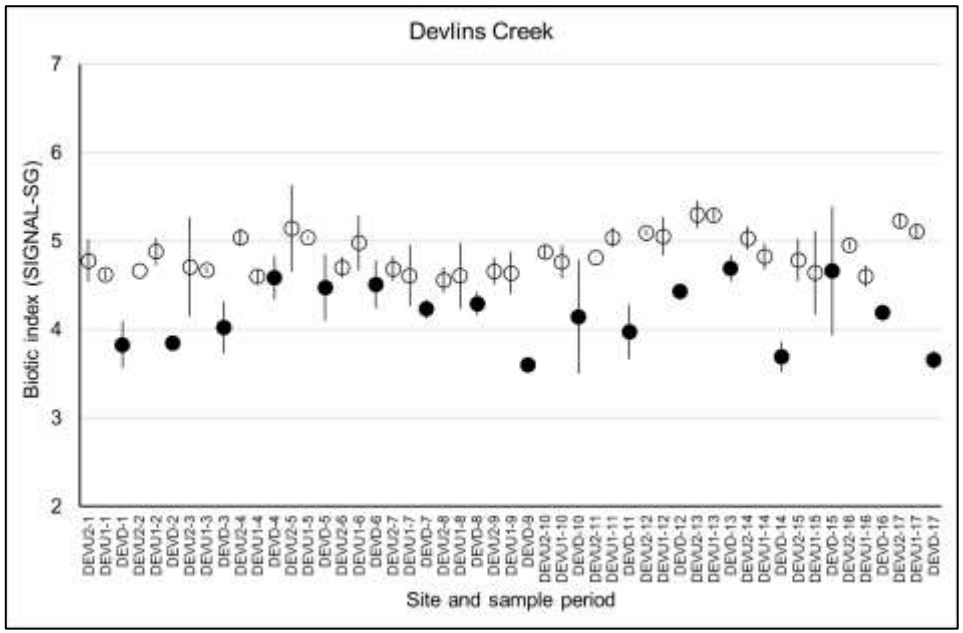
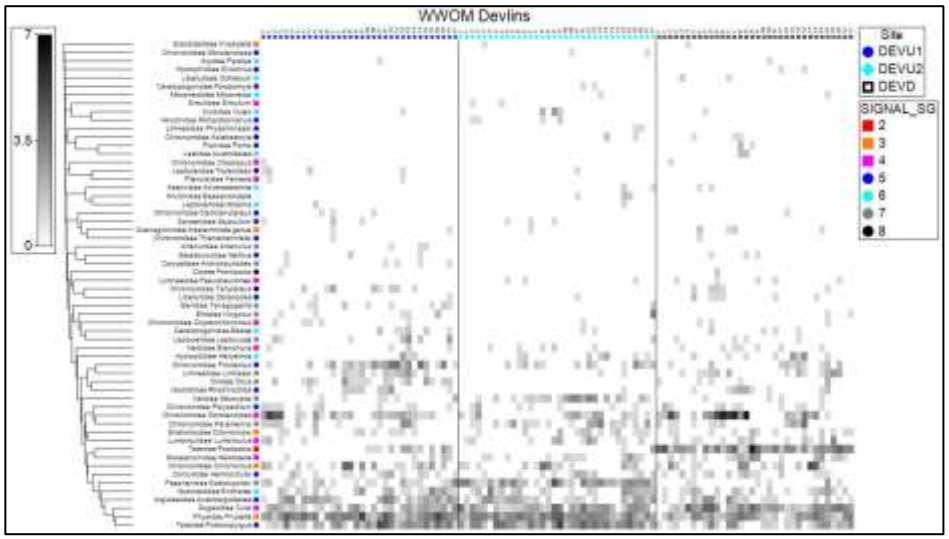
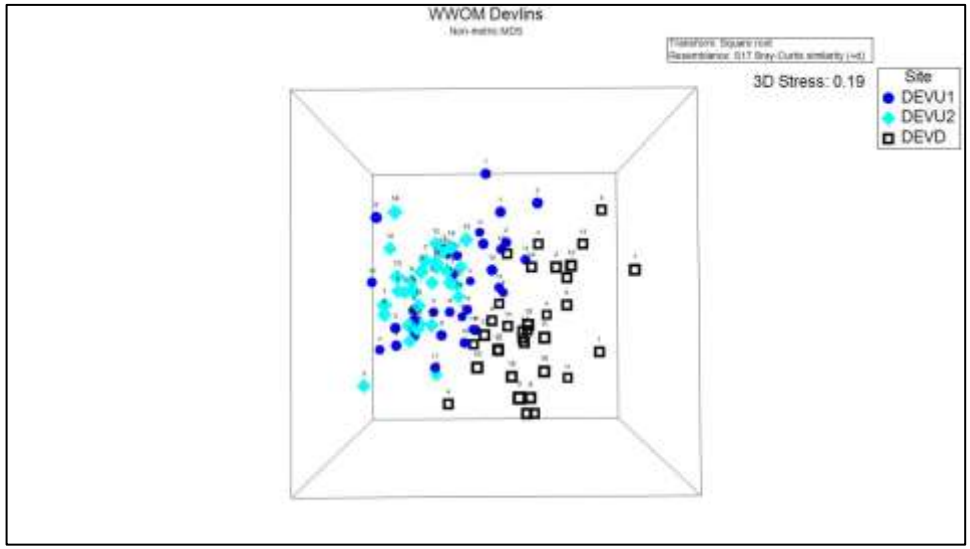


Figure 10-4: Overview of morphometric macroinvertebrate data for Frenchs Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot







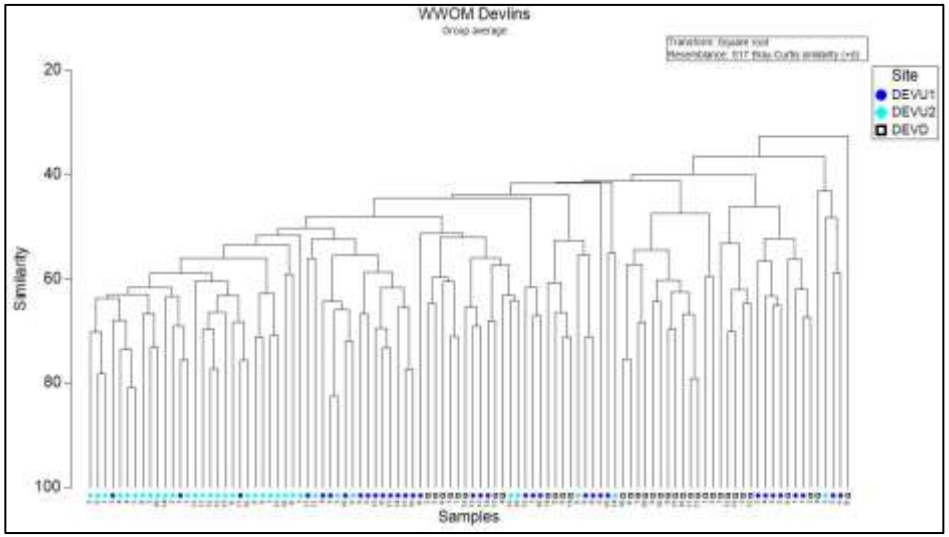


Figure 10-5: Overview of morphometric macroinvertebrate data for Devlins Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

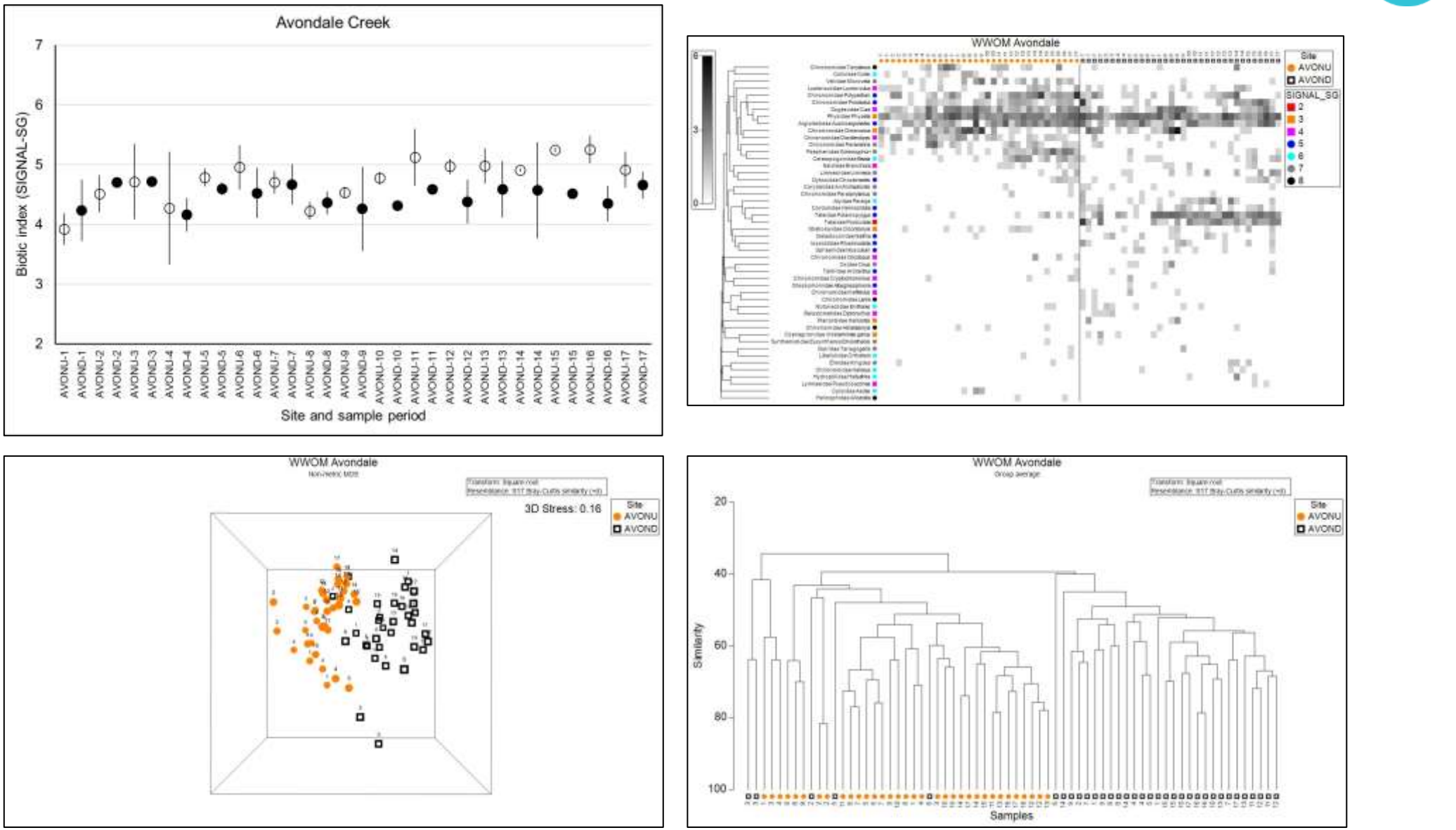


Figure 10-6: Overview of morphometric macroinvertebrate data for Avondale Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

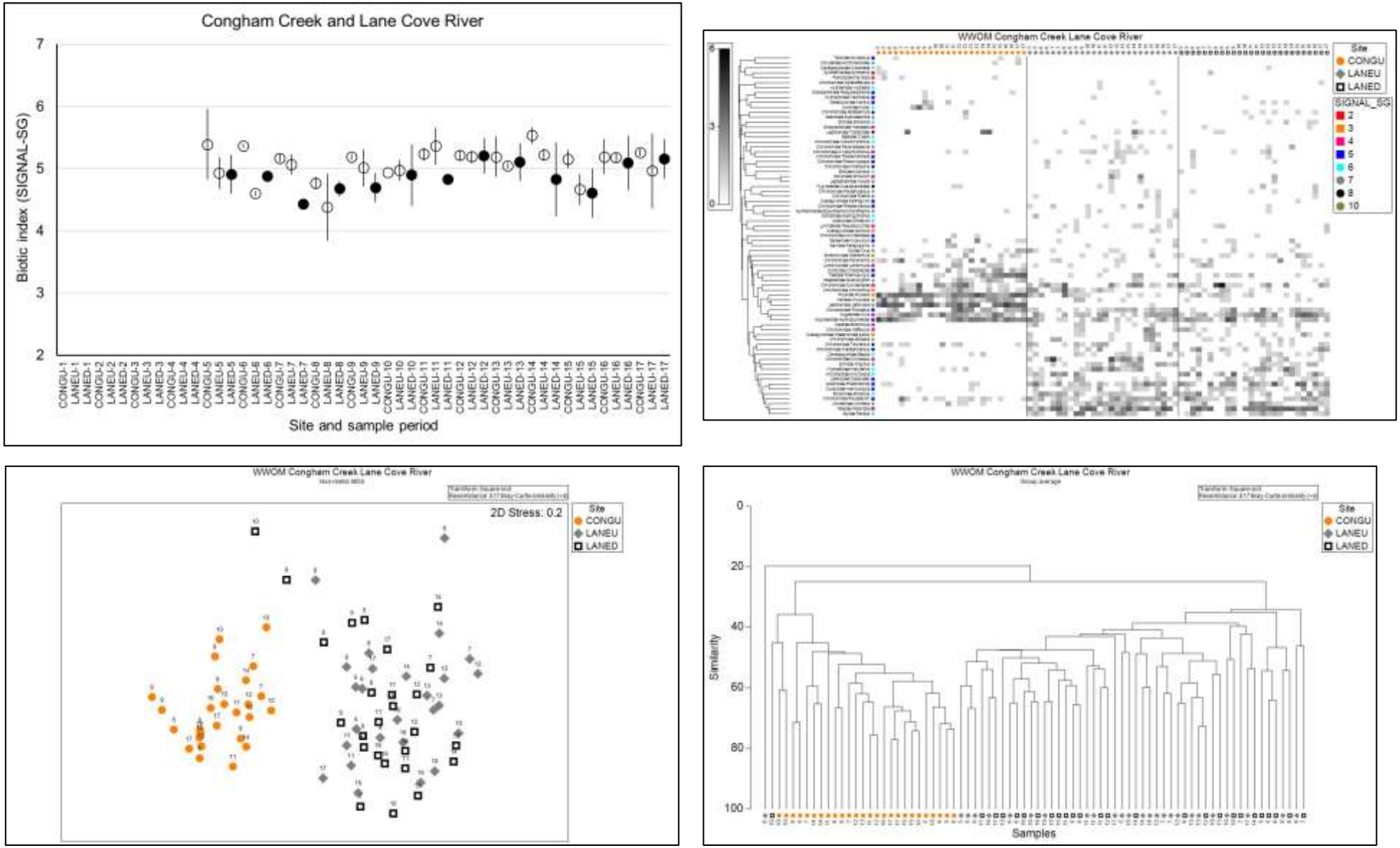
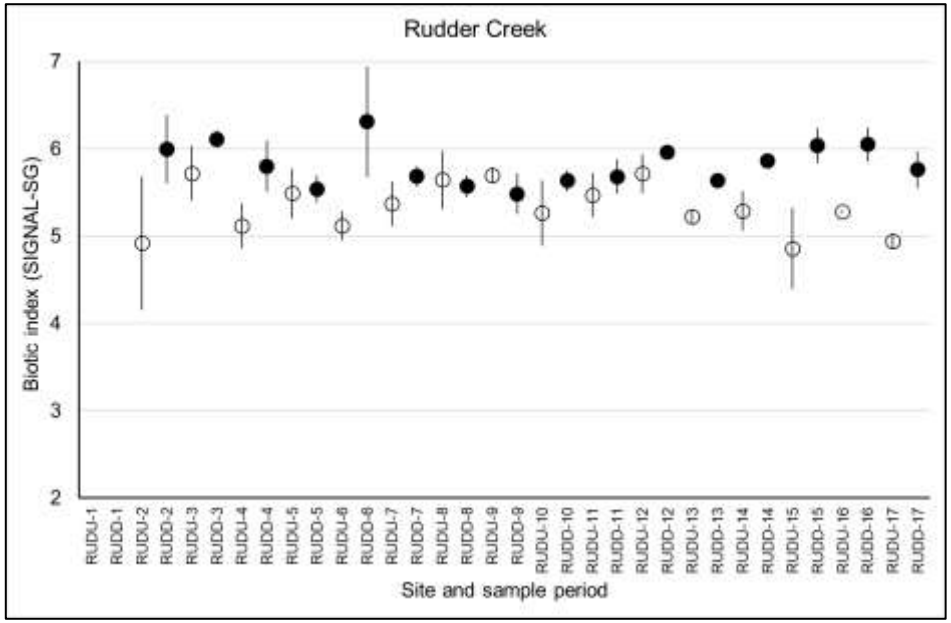
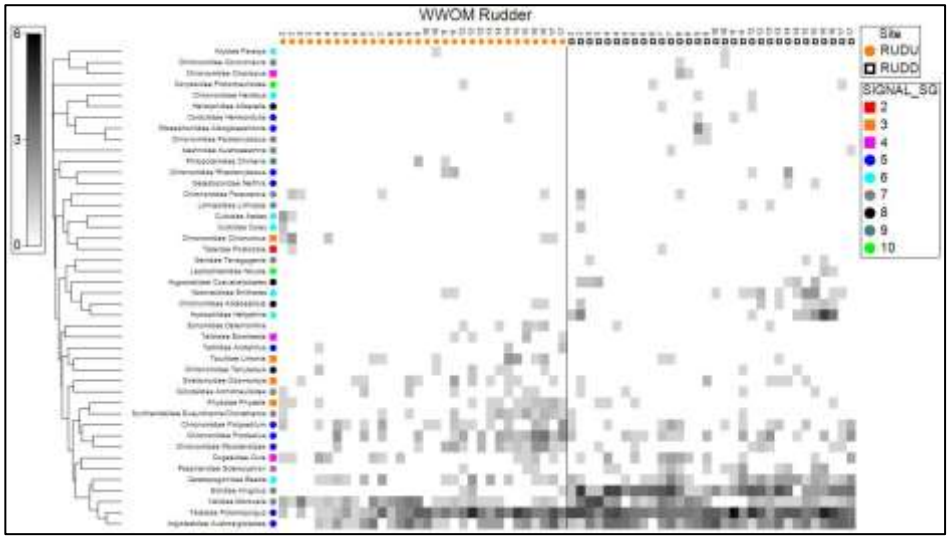
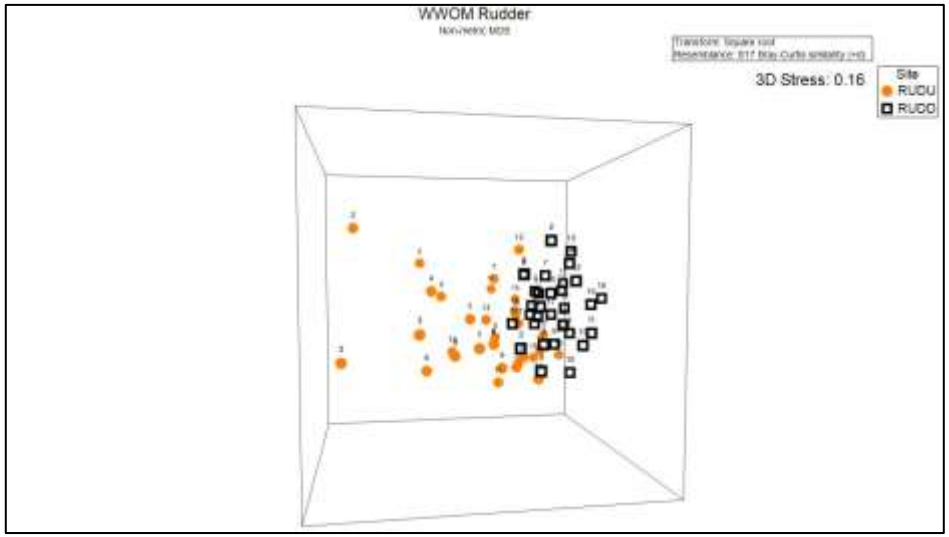


Figure 10-7: Overview of morphometric macroinvertebrate data for Congham Creek and Lane Cove River

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot







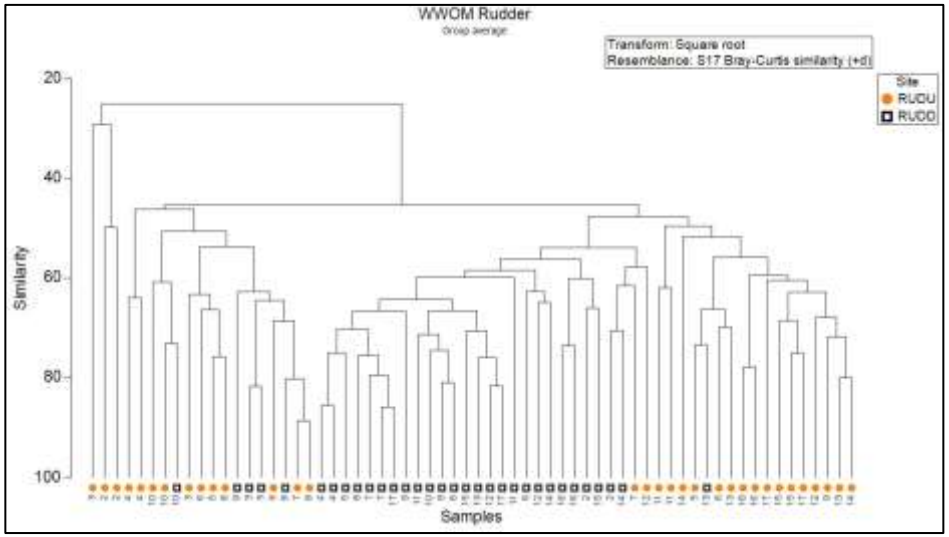


Figure 10-8: Overview of morphometric macroinvertebrate data for Rudder Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

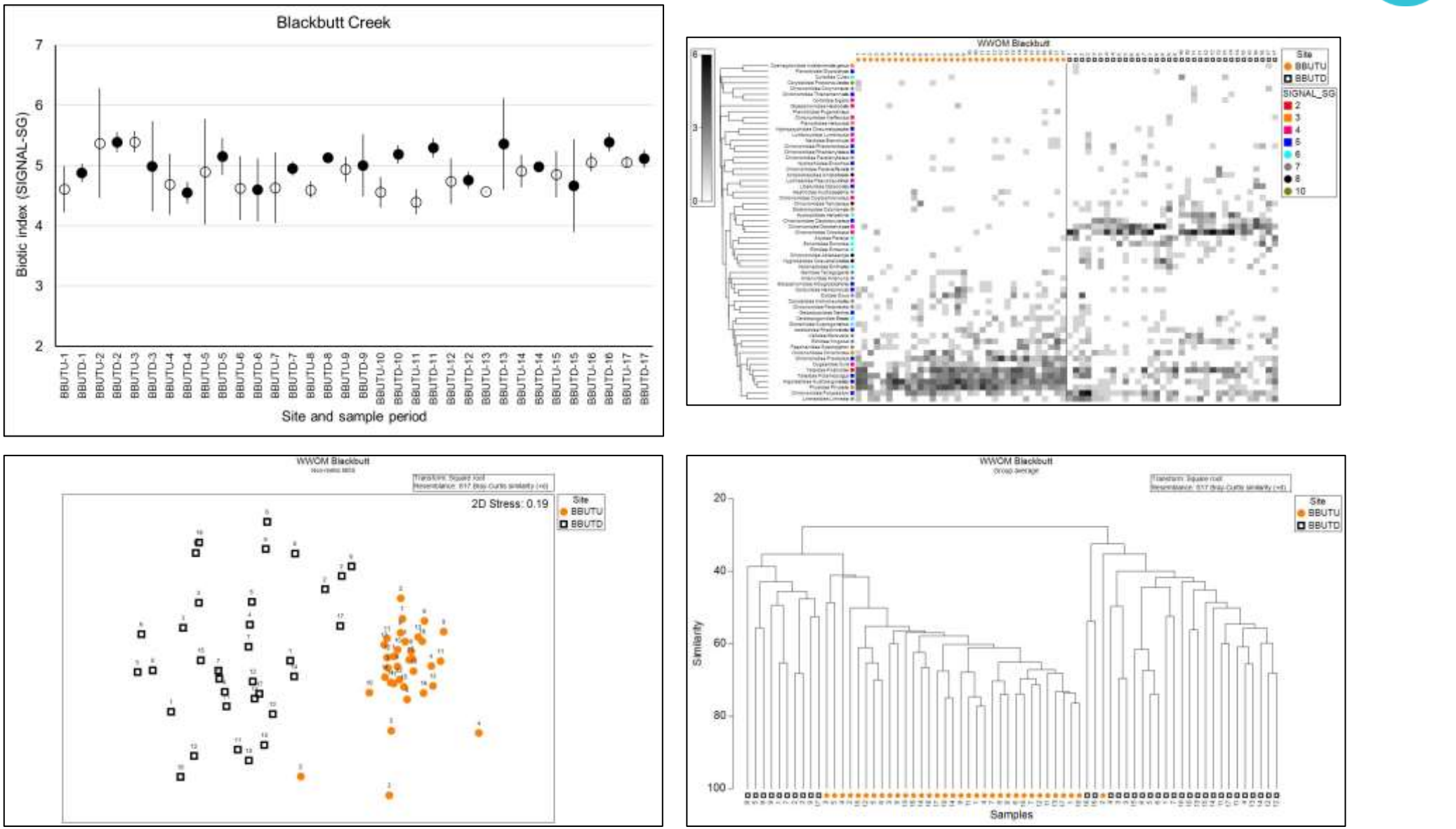


Figure 10-9: Overview of morphometric macroinvertebrate data for Blackbutt Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

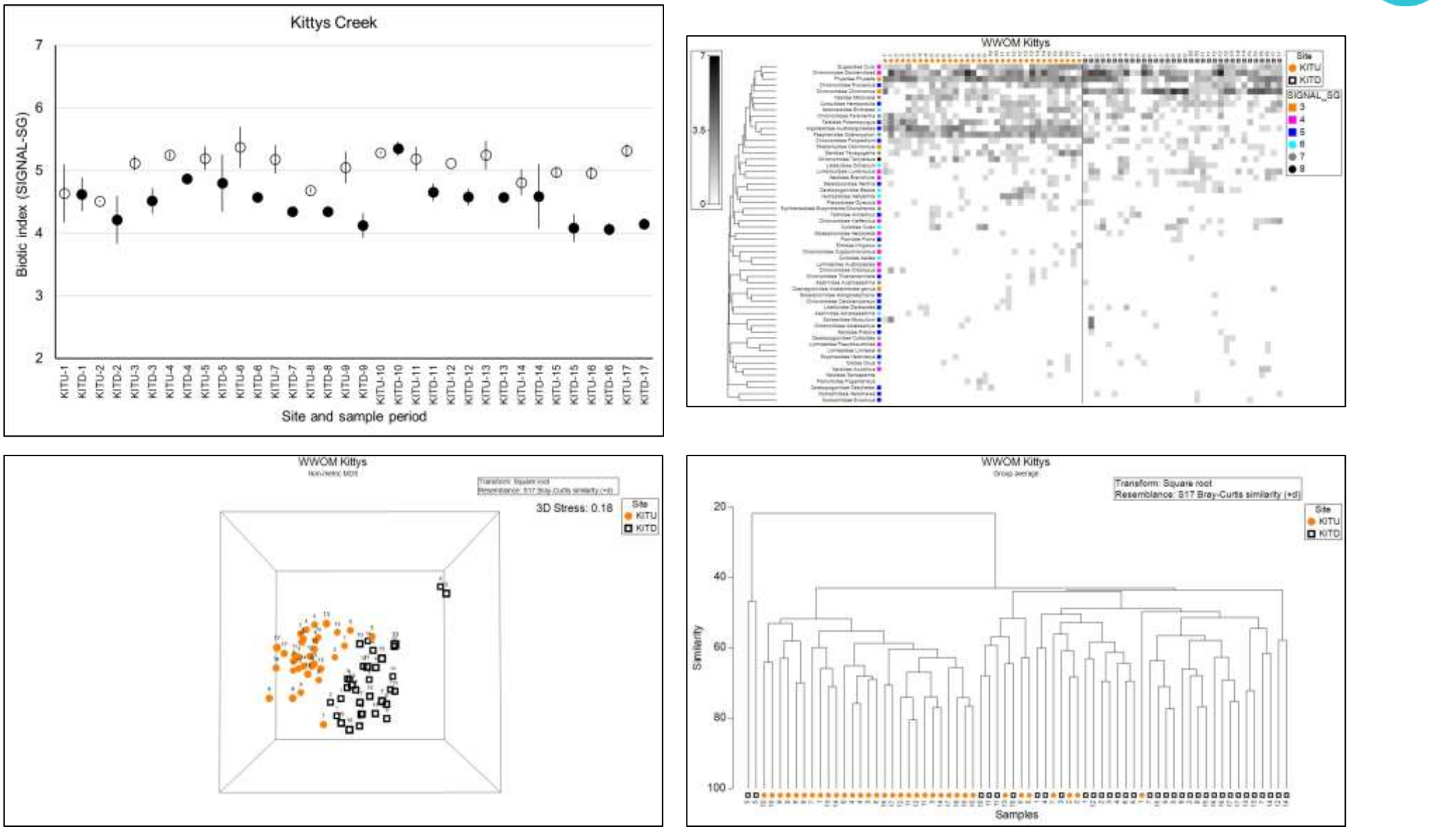


Figure 10-10: Overview of morphometric macroinvertebrate data for Kittys Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

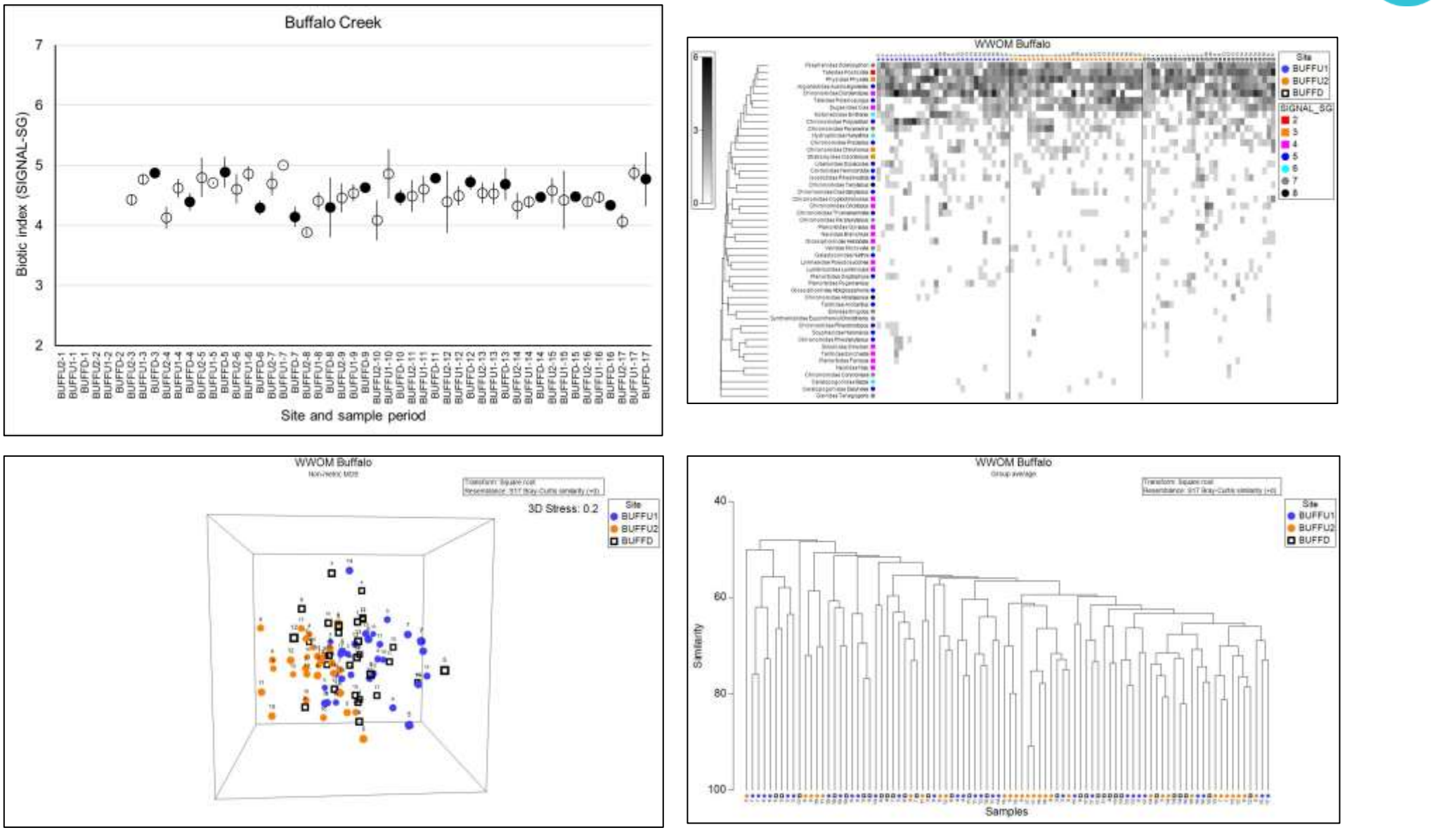


Figure 10-11: Overview of morphometric macroinvertebrate data for Buffalo Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

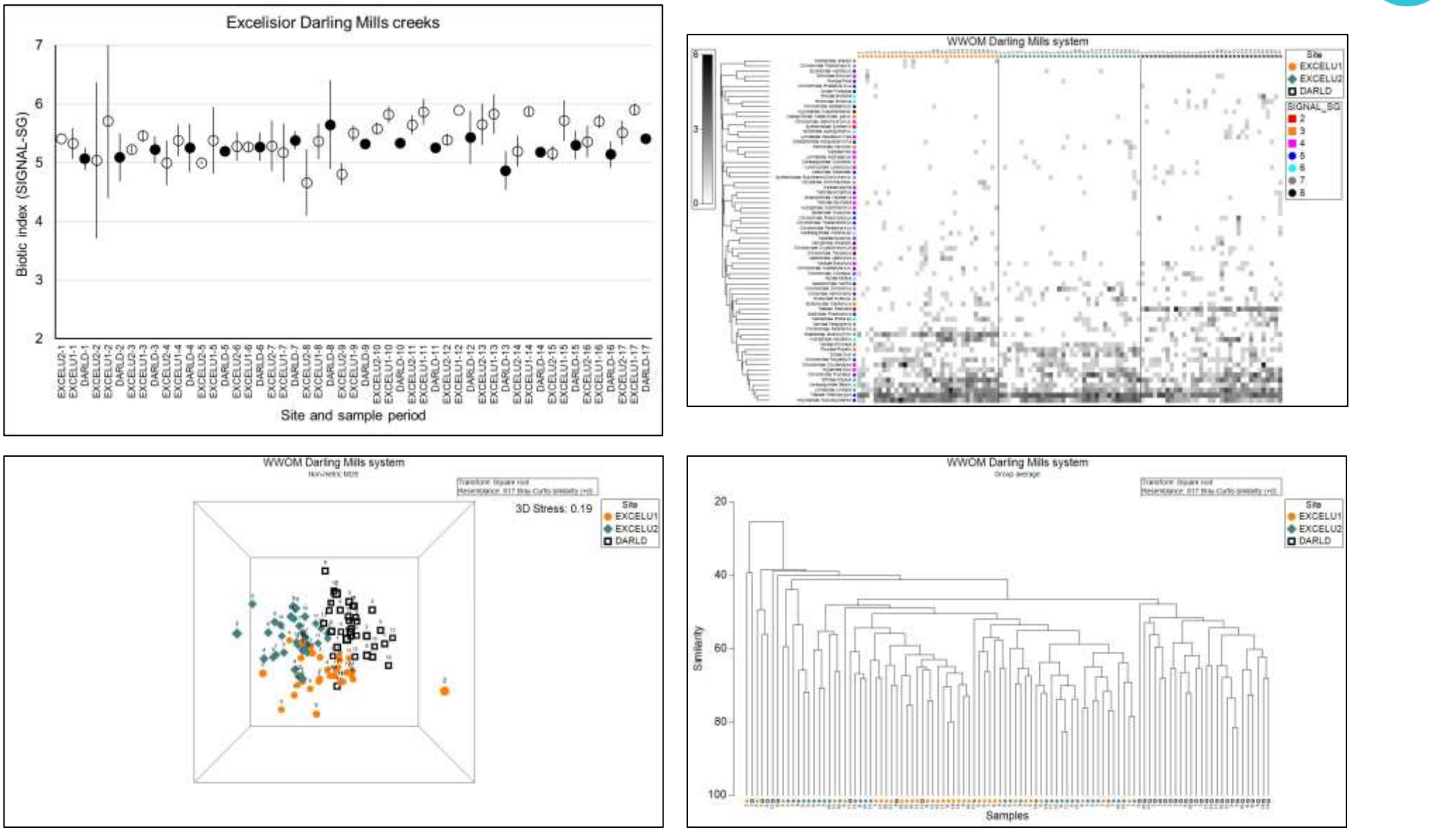


Figure 10-12: Overview of morphometric macroinvertebrate data for Darling Mills creek system

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

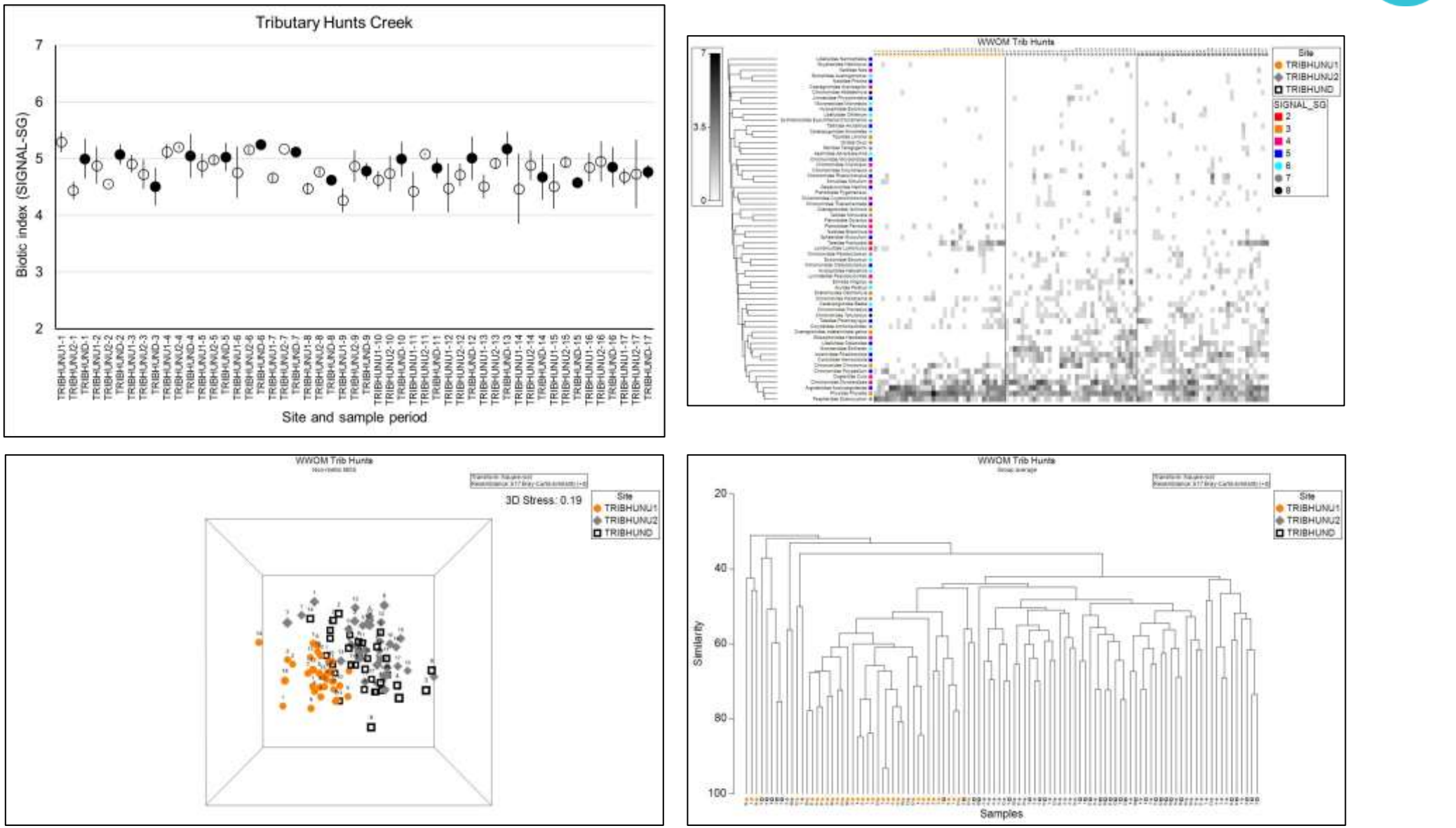


Figure 10-13: Overview of morphometric macroinvertebrate data for Tributary Hunts Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

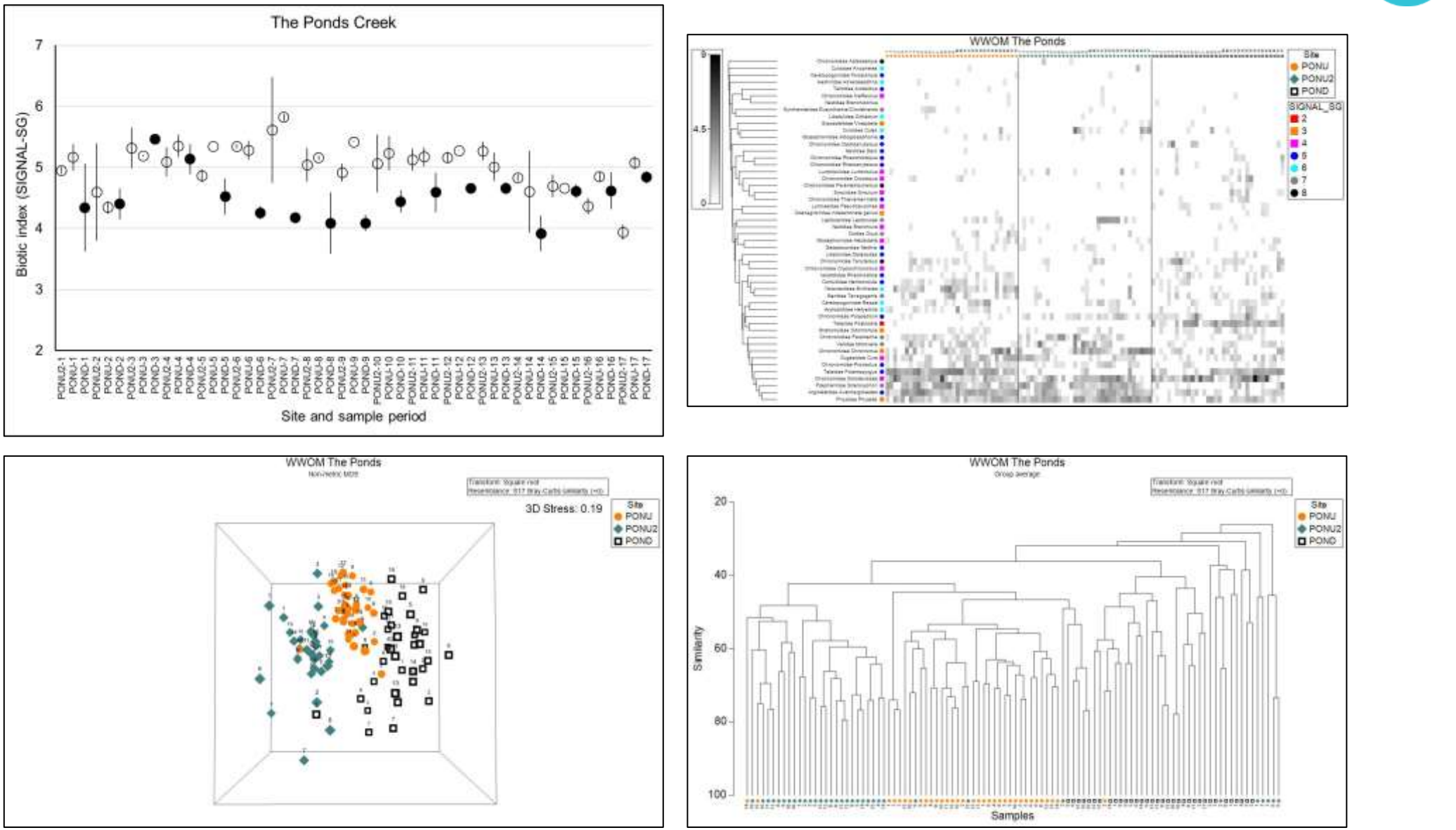


Figure 10-14: Overview of morphometric macroinvertebrate data for The Ponds Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

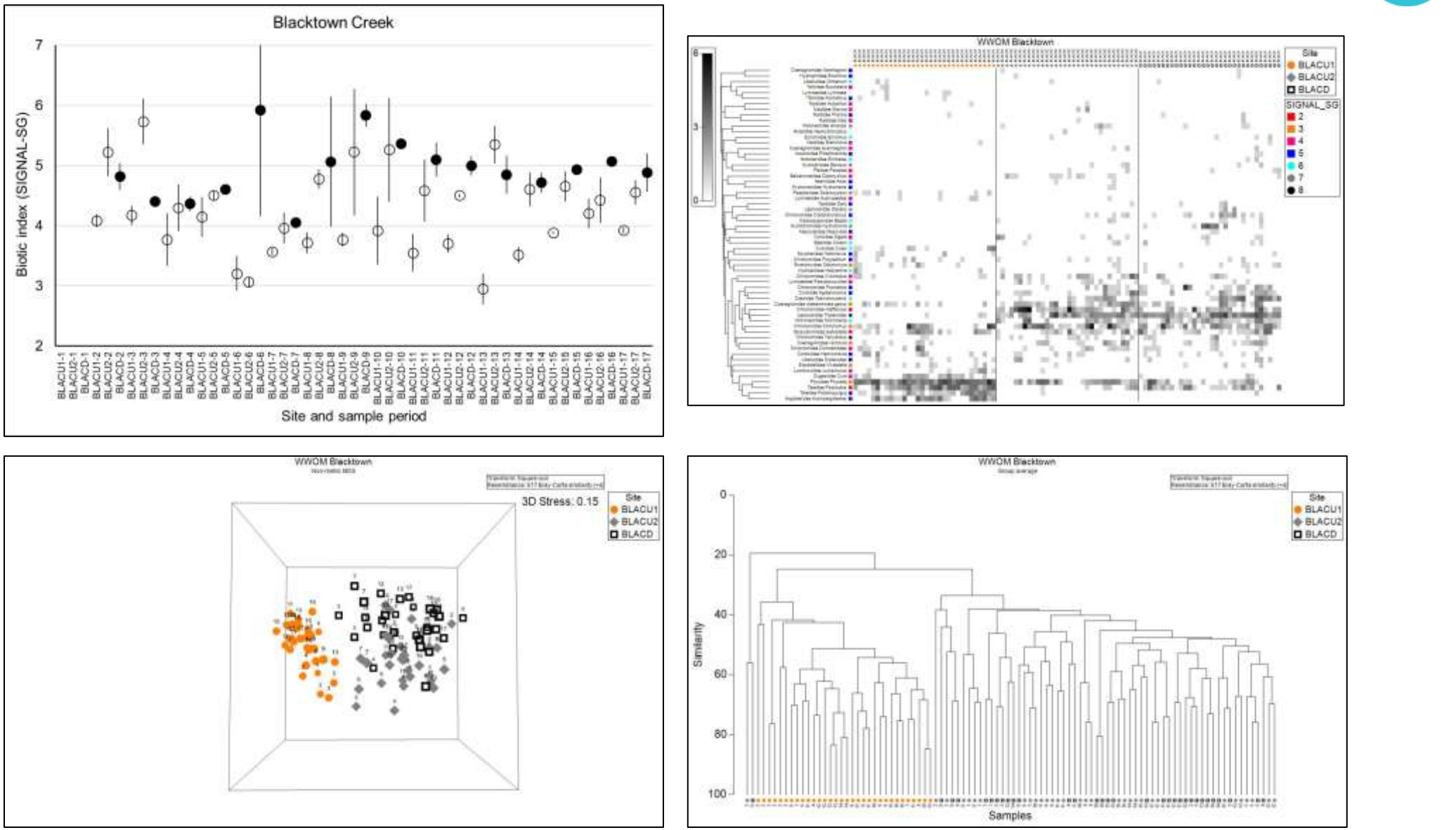
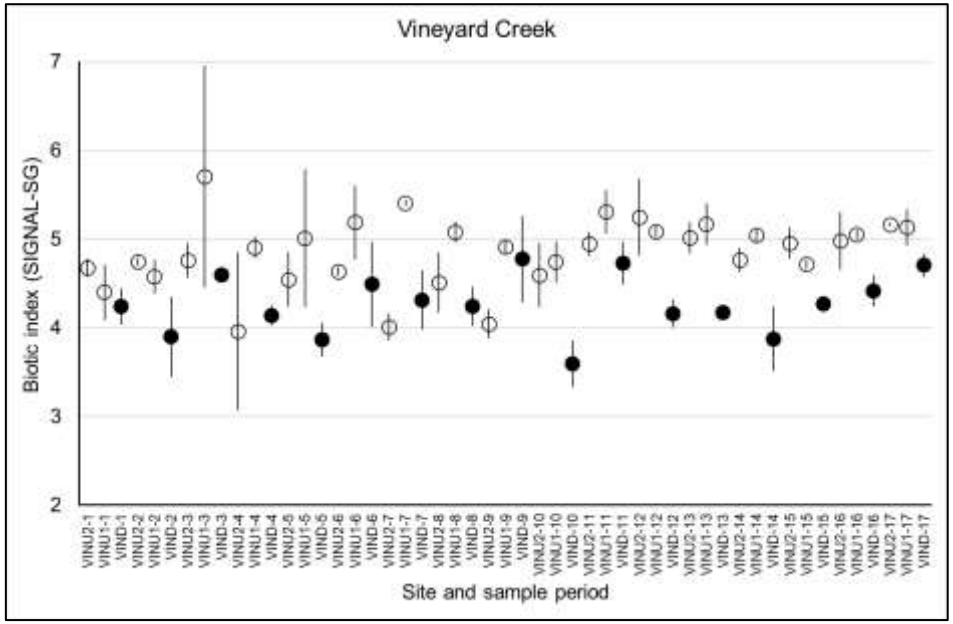
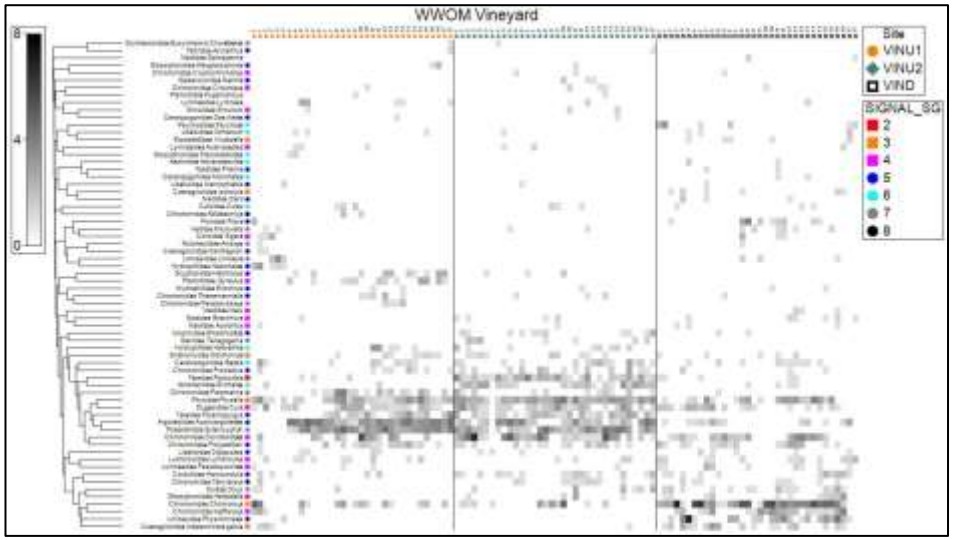
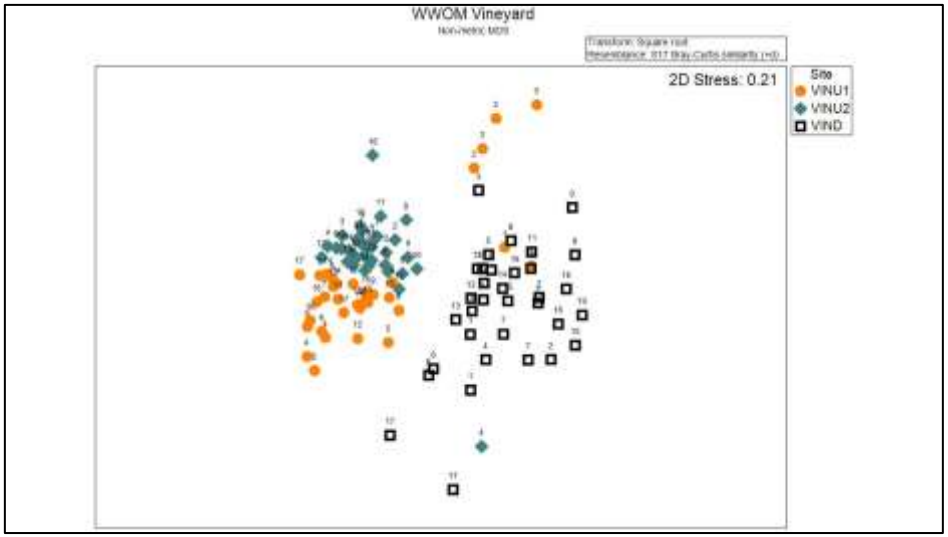


Figure 10-15: Overview of morphometric macroinvertebrate data for Blacktown Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot







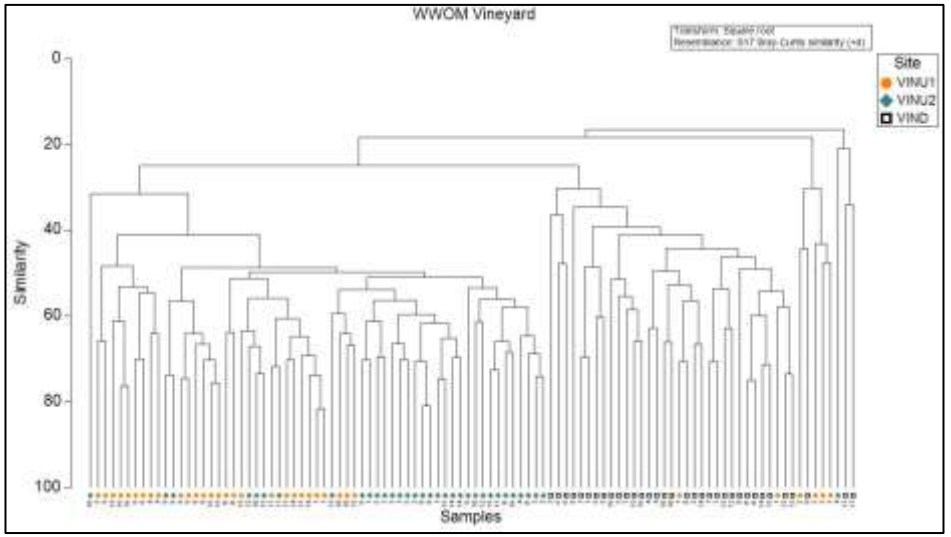


Figure 10-16: Overview of morphometric macroinvertebrate data for Vineyard Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

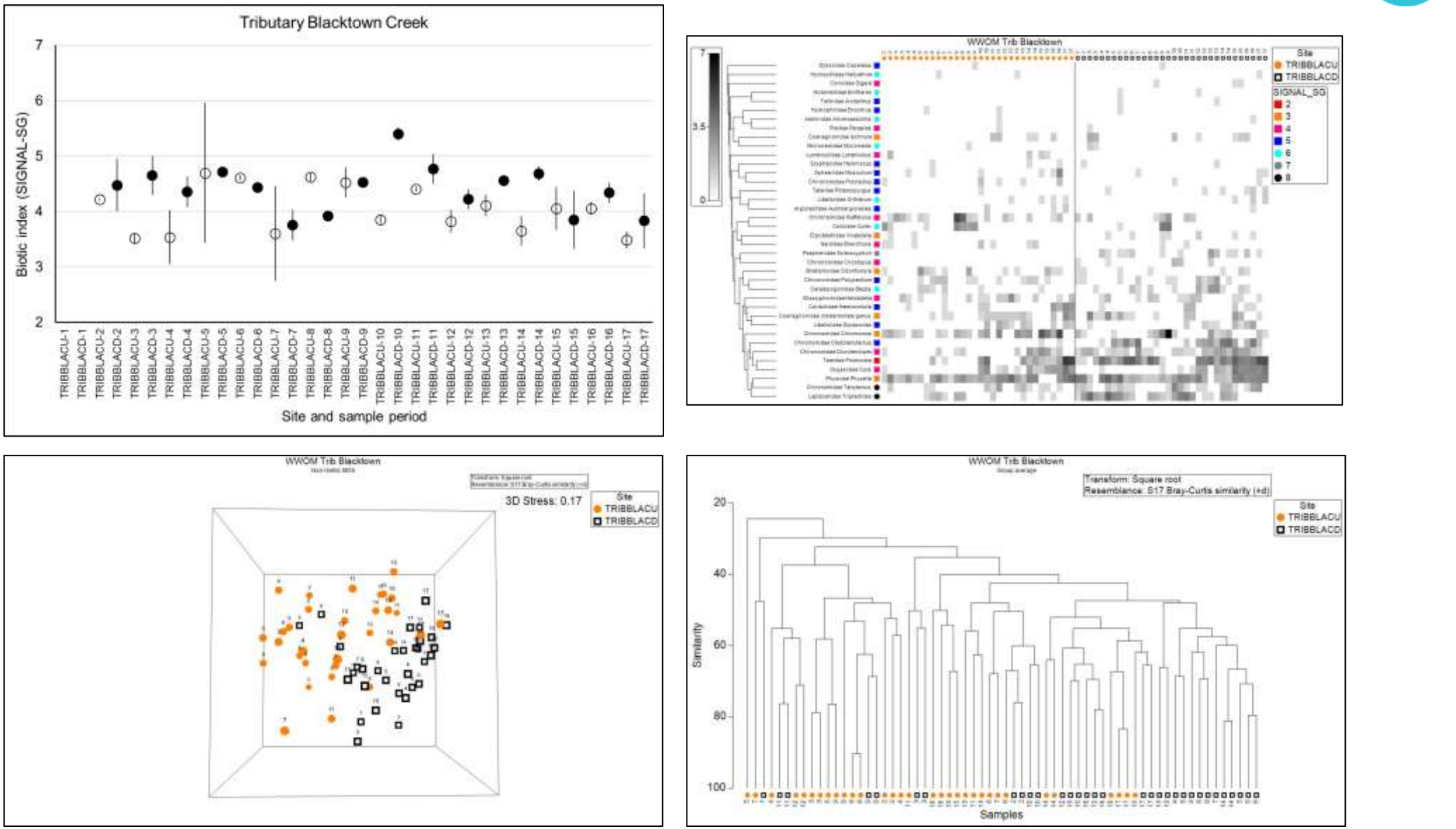


Figure 10-17: Overview of morphometric macroinvertebrate data for Tributary Blacktown Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

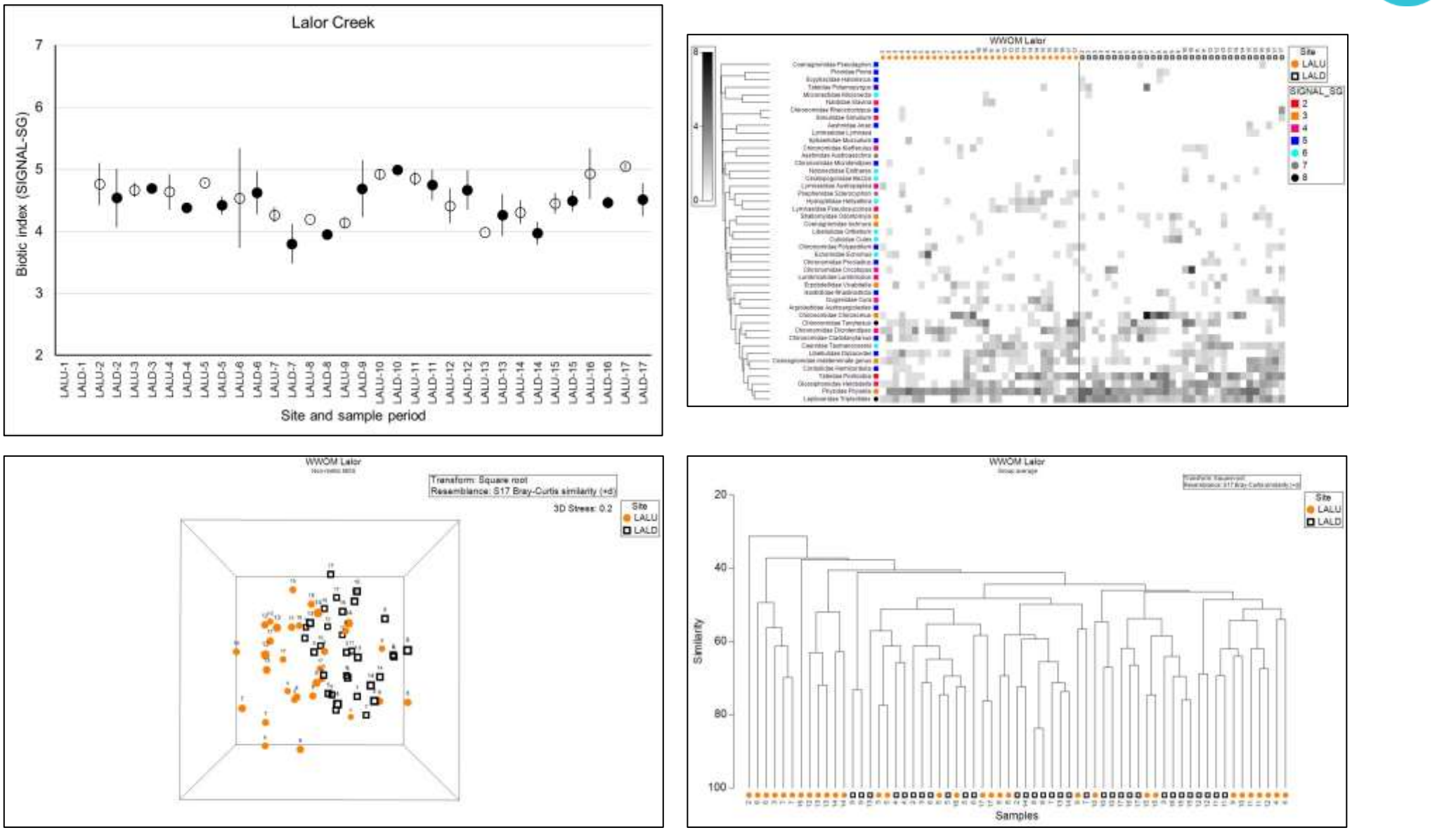
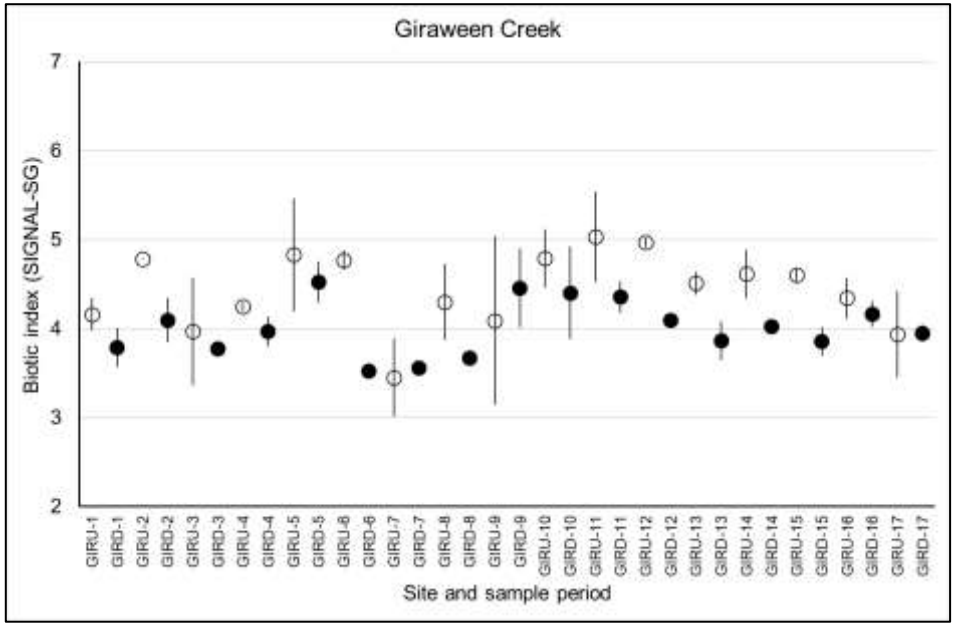
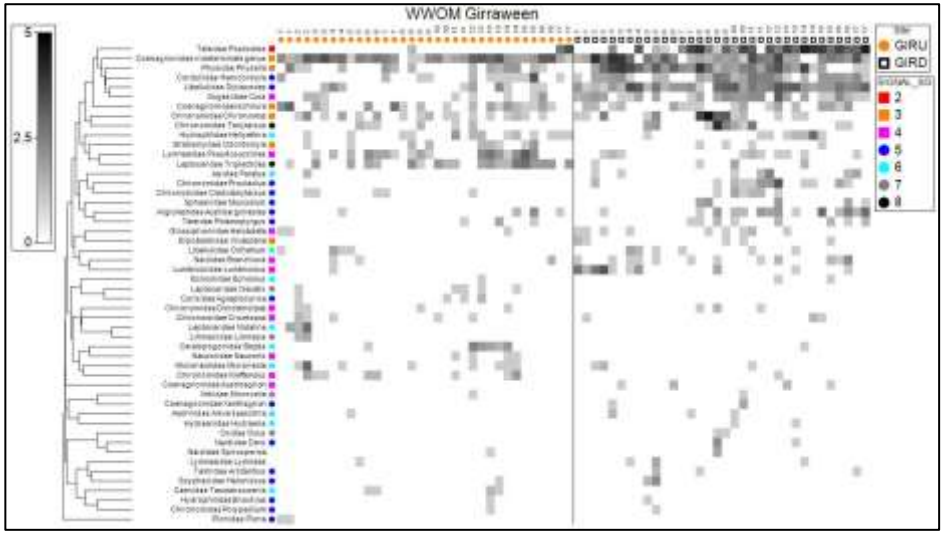
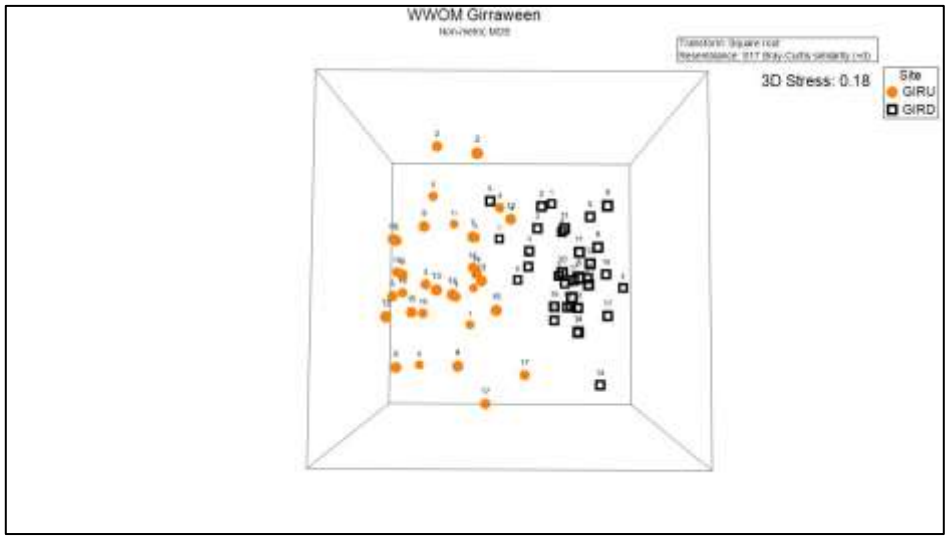


Figure 10-18: Overview of morphometric macroinvertebrate data for Lalor Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot







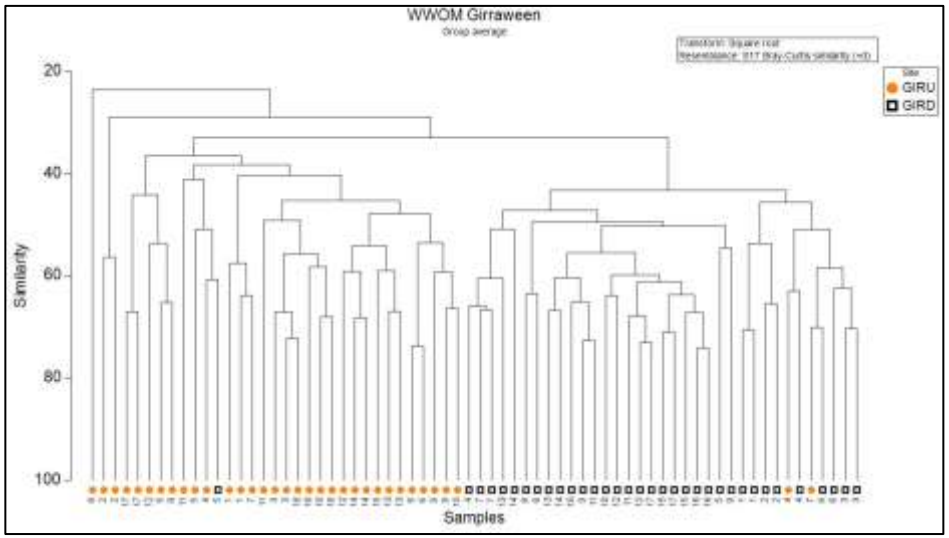
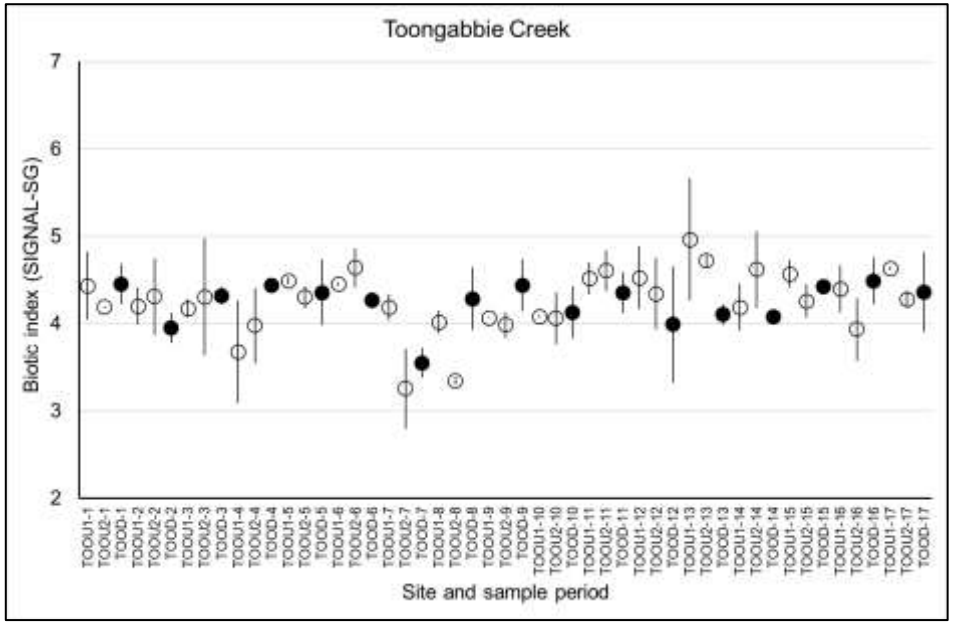
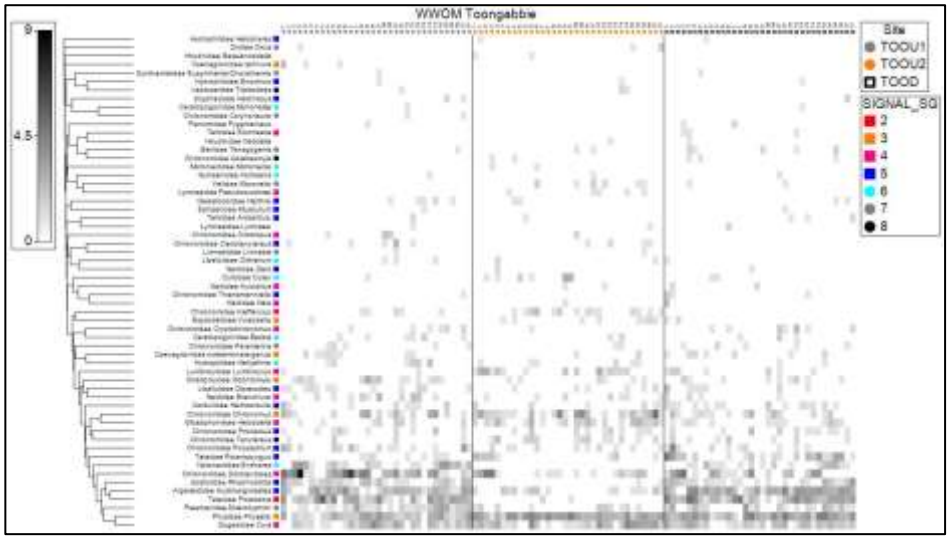
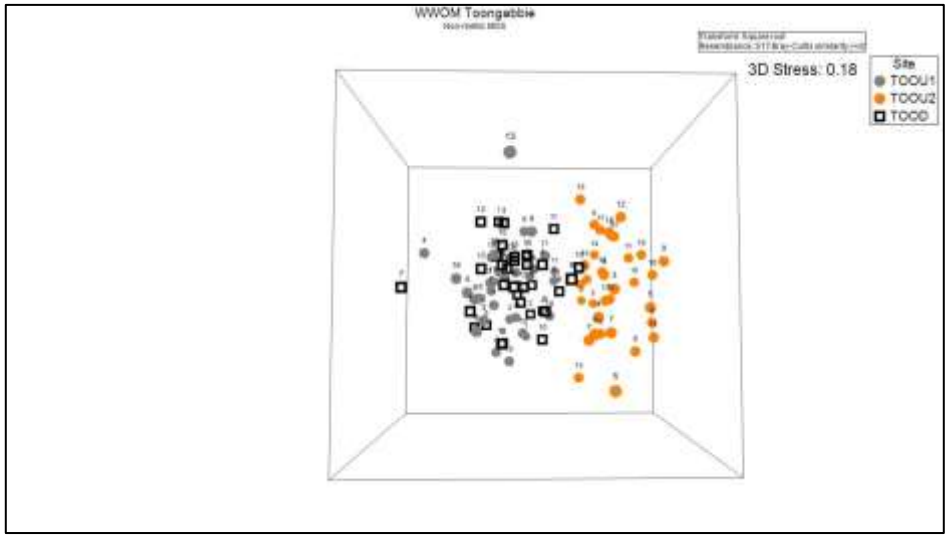


Figure 10-19: Overview of morphometric macroinvertebrate data for Girraween Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot







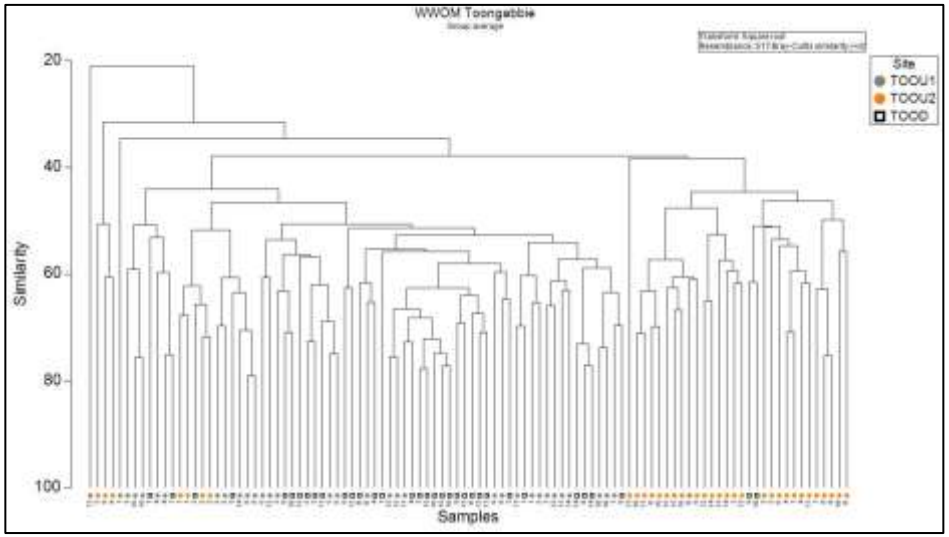


Figure 10-20: Overview of morphometric macroinvertebrate data for Toongabbie Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

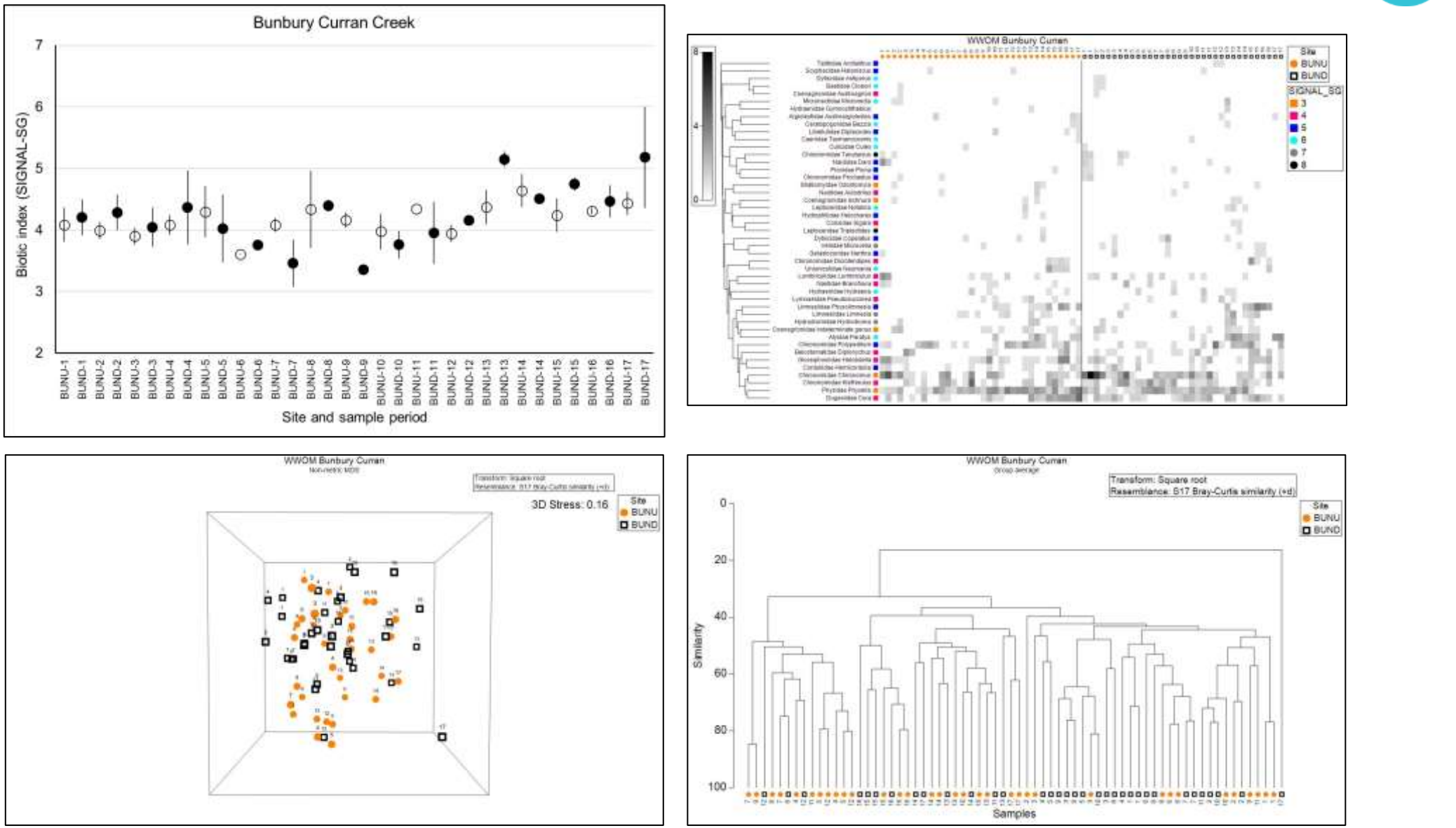


Figure 10-21: Overview of morphometric macroinvertebrate data for Bunbury Curran Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

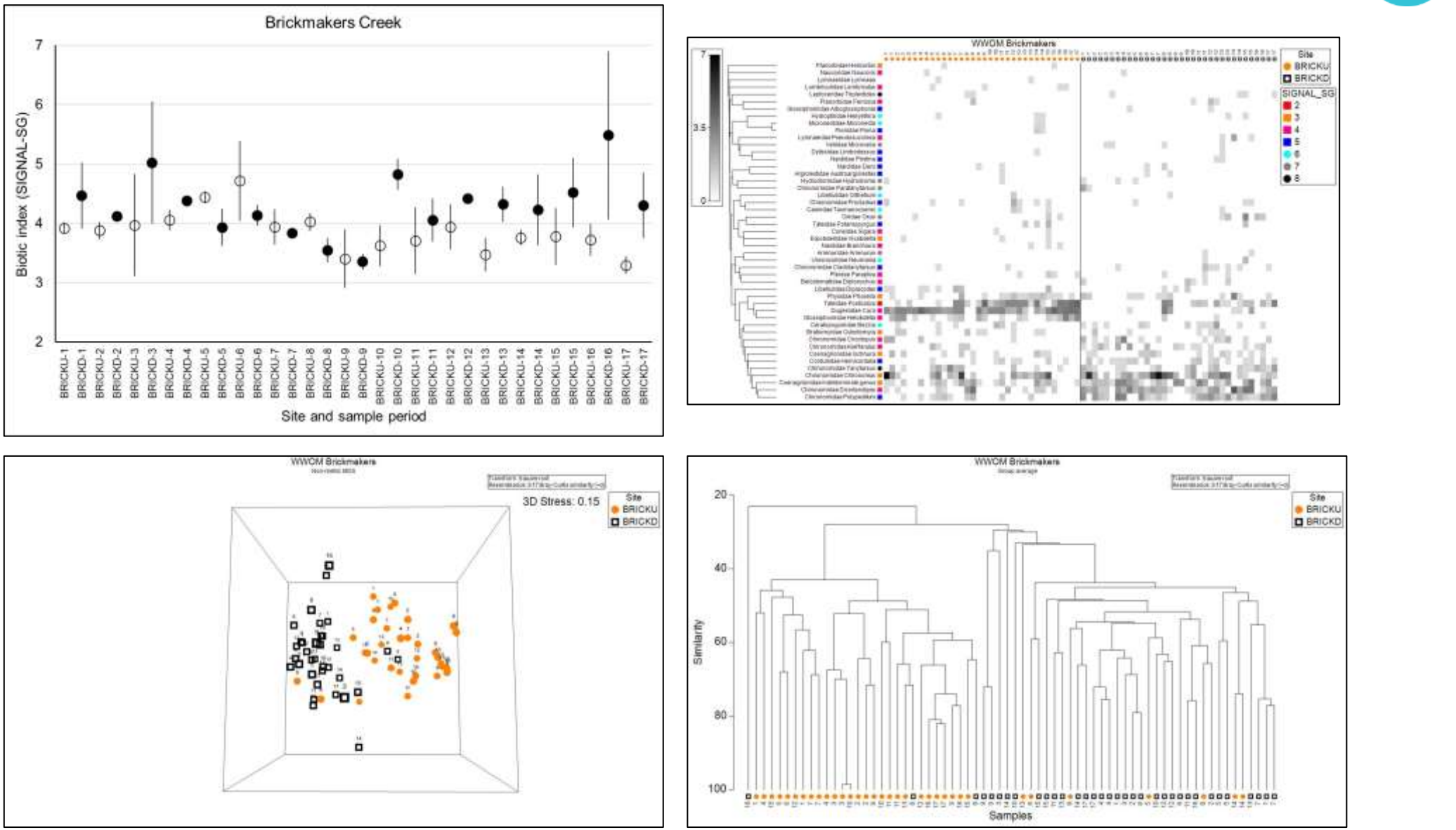
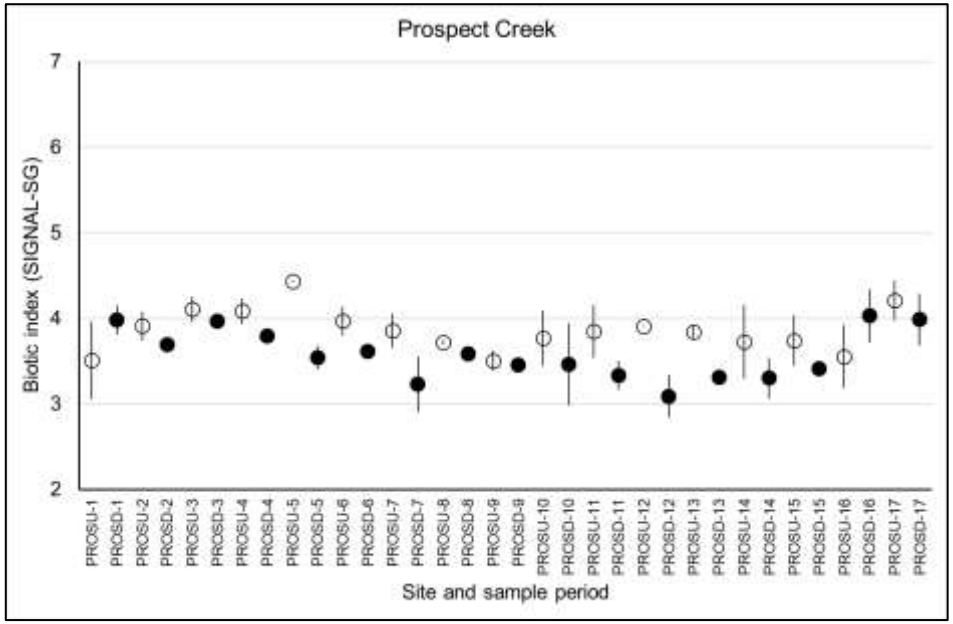
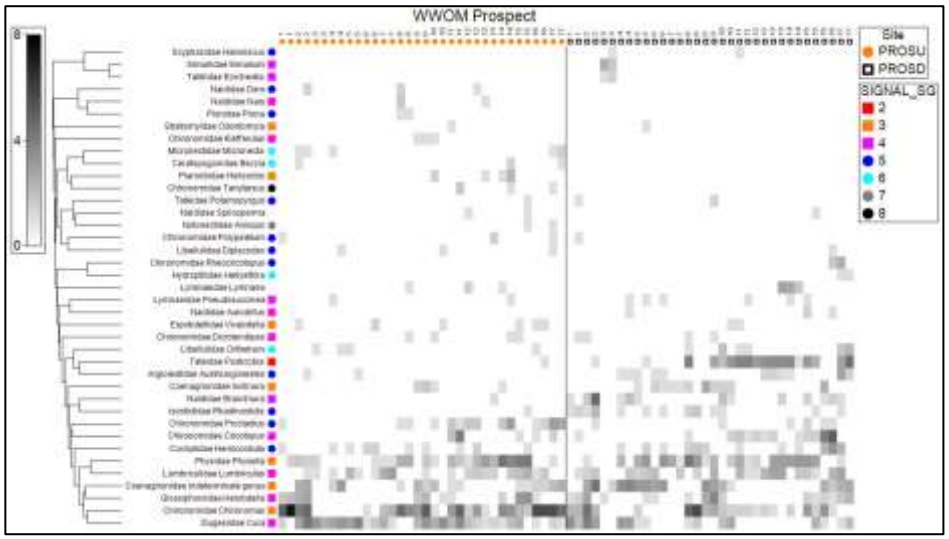
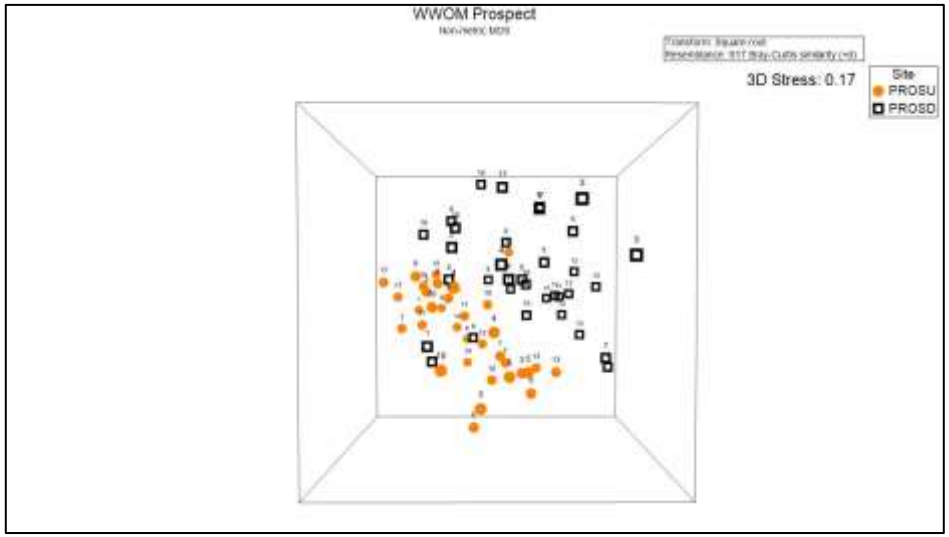


Figure 10-22: Overview of morphometric macroinvertebrate data for Brickmakers Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot







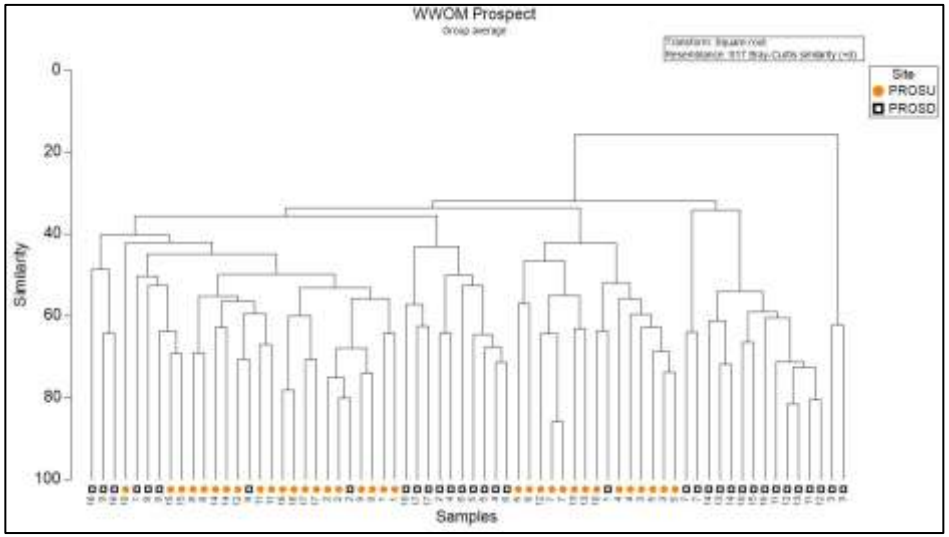


Figure 10-23: Overview of morphometric macroinvertebrate data for Prospect Creek

From top left clockwise: control chart of SIGNAL-SG scores; shade plot; tree diagram; nMDS ordination plot

10.2 Appendix 2 - Study site coordinates

10.2.1 Freshwater monitoring locations

Table 10-2: Freshwater streams monitoring locations for eDNA sediments and macroinvertebrates

Site code	Site Name	Latitude	Longitude
BRICKU	Brickmakers Ck u/s overflow	-33.917297	150.918932
BRICKD	Brickmakers Ck d/s overflow	-33.909258	150.929932
PROSU	Prospect Ck u/s overflow	-33.841405	150.926729
PROSD	Prospect Ck d/s overflow	-33.843400	150.932395
BUNU	Bunbury Curran Ck u/s overflow	-34.006882	150.834765
BUND	Bunbury Curran Ck d/s overflow	-34.006126	150.836665
EXCELU1	Excelsior Ck (& Darling Mills Ck) u/s overflow 1	-33.747665	151.014836
EXCELU2	Excelsior Ck (& Darling Mills Ck) u/s overflow 2	-33.749433	151.013485
DARLD	Darling Mills Ck (& Excelsior Ck) d/s overflow	-33.762655	151.00957
TRIBHUNU1	Trib Hunts Ck u/s overflow 1	-33.778414	151.028244
TRIBHUNU2	Trib Hunts Ck u/s overflow 2	-33.77925	151.033894
TRIBHUND	Trib Hunts Ck d/s overflow	-33.781801	151.022887
VINU1	Vineyard Ck u/s overflow 1	-33.791251	151.033508
VINU2	Vineyard Ck u/s overflow 2	-33.794128	151.038135
VIND	Vineyard Ck d/s overflow	-33.803443	151.031423
PONU	The Ponds Ck u/s overflow	-33.791667	151.049945
PONU2	The Ponds Ck u/s overflow 2	-33.790081	151.048907
POND	The Ponds Ck d/s overflow	-33.809304	151.046695
KITU	Kittys Ck u/s overflow	-33.80586	151.13248
KITD	Kittys Ck d/s overflow	-33.80844	151.135110
BUFFU1	Buffalo Ck u/s overflow 1	-33.814261	151.120310
BUFFU2	Buffalo Ck u/s overflow 2	-33.813413	151.121258
BUFFD	Buffalo Ck d/s overflow	-33.814701	151.122471
DEVU1	Devlins Ck u/s overflow	-33.756765	151.062252
DEVU2	Devlins Ck u/s overflow 2	-33.763762	151.062012
DEVD	Devlins Ck d/s overflow	-33.761557	151.071874
RUDU	Rudder Ck u/s overflow	-33.769150	151.13543

Site code	Site Name	Latitude	Longitude
RUDD	Rudder Ck d/s overflow	-33.772593	151.133525
AVONU	Avondale Ck u/s overflow	-33.75082	151.11647
AVOND	Avondale Ck d/s overflow	-33.76236	151.11858
BBUTU	Blackbutt Ck u/s overflow	-33.76155	151.14152
BBUTD	Blackbutt Ck d/s overflow	-33.77516	151.13682
CONGU	Congham Ck u/s overflow	-33.768996	151.122944
LANEU	Lane Cove R u/s overflow	-33.769818	151.121881
LANED	Lane Cove R d/s overflow	-33.77235	151.129557
HORNU	Hornsby Ck (& Cockle Ck) u/s overflow	-33.69883	151.10687
COCU	Hornsby Ck (& Cockle Ck) u/s overflow	-33.7033	151.11863
COCD	Hornsby Ck (& Cockle Ck) d/s overflow	-33.69475	151.11951
FRENU2	French's Ck u/s overflow 2	-33.731777	151.193319
FREND	French's Ck d/s overflow	-33.729013	151.189368
CARU	Carroll Ck u/s overflow	-33.75002	151.213197
CARD	Carroll Ck d/s overflow	-33.754058	151.207066
COWSU	Cowan Ck, south branch u/s overflow	-33.724735	151.149527
COWSD	Cowan Ck, south branch d/s overflow	-33.714882	151.154974
TOOU1	Toongabbie Ck u/s overflow 1	-33.759305	150.967521
TOOU2	Toongabbie Ck u/s overflow 2	-33.753011	150.9665053
TOOD	Toongabbie Ck d/s overflow	-33.762213	150.963765
LALU	Lalor Ck u/s overflow	-33.761139	150.939288
LALD	Lalor Ck d/s overflow	-33.773794	150.942659
TRIBBLACU	Trib Blacktown Ck u/s overflow	-33.780831	150.935026
TRIBBLACD	Trib Blacktown Ck d/s overflow	-33.777211	150.940677
GIRU	Girraween Ck u/s overflow	-33.799513	150.935372
GIRD	Girraween Ck d/s overflow	-33.784638	150.951104
BLACU1	Blacktown Ck u/s overflow 1	-33.772581	150.918404
BLACU2	Blacktown Ck u/s overflow 2	-33.782548	150.917915
BLACD	Blacktown Ck d/s overflow	-33.773219	150.932555

10.2.2 Estuarine monitoring locations – eDNA sediment samples

Table 10-3: Estuarine monitoring locations for eDNA sediment samples

Site code	Site Name	Latitude	Longitude
SHSED01	Sydney Harbour, Chowder Bay	-33.8398711	151.2539735
SHSED02	Sydney Harbour, Taylors Bay	-33.843894	151.248608
SHSED03	Sydney Harbour, Mosman Bay	-33.836953	151.232436
PRSED01	Parramatta River, Iron Cove	-33.8616196	151.1617482
PRSED02	Parramatta River, Iron Cove	-33.8690836	151.1427953
PRSED03	Parramatta River, Tarban Creek	-33.8364755	151.1402916
PRSED04	Parramatta River, Hen and Chicken Bay	-33.8608946	151.1231847
PRSED05	Parramatta River, Glades Bay	-33.834799	151.119658
PRSED06	Parramatta River near Homebush Bay	-33.820752	151.08162
GRSED01	Georges River, Kogarah Bay	-33.989324	151.122237
GRSED02	Georges River, Boggywell Creek	-33.9826604	151.0522872
GRSED03	Georges River, near Edith Bay	-33.992817	151.044295
GRSED04	Georges River Salt Pan Creek	-33.957956	151.041982
GRSED05	Georges River, near Dhurawal Bay	-33.915814	150.973932
MHSED01	Middle Harbour, Clontarf	-33.807389	151.250709
MHSED02	Middle Harbour, Quaker Hat Bay	-33.816334	151.23841
MHSED03	Middle Harbour, Roseville	-33.774137	151.204697
CRSED01	Cooks River, M5 motorway	-33.94307	151.160223
CRSED02	Wolli Creek, Unwin Street	-33.927227	151.148155
CRSED03	Cooks River, Karool Avenue	-33.914702	151.124869
CRSED04	Cooks River, Wanstead Avenue, Earlwood	-33.9214229	151.1482808
PHSED01	Port Hacking, Yowie Bay	-34.045792	151.109319
PHSED02	Port Hacking, Gymea Bay	-34.049293	151.098189
PHSED03	Port Hacking, Gymea Bay	-34.049869	151.093524
PHSED04	Port Hacking, North West Arm	-34.055774	151.084996
LCSED01	Lower Lane Cove River	-33.821215	151.148156

10.2.3 Estuarine monitoring locations – Passive samplers

Table 10-4: Estuarine passive sampler monitoring locations

Site code	Site Name	Latitude	Longitude
CRPAS01	Cooks River	-33.942917	151.160204
CRPAS02	Cooks River, Wolli Creek	-33.927121	151.147967
CRPAS03	Cooks River	-33.914574	151.122248
CRPAS04	Cooks River	-33.921365	151.148203
GRPAS01	Georges River, Kogarah Bay	-33.991288	151.119481
GRPAS02	Georges River, Boggywell Creek	-33.985637	151.052897
GRPAS03	Georges River, near Edith Bay	-33.993179	151.043891
GRPAS04	Georges River Salt Pan Creek	-33.957552	151.041974
GRPAS05	Georges River, near Dhurawal Bay	-33.915734	150.974189
HCPAS01	Hunts Creek, Northam Drive	-33.778134	151.029564
LCPAS01	Lower Lane Cove River	-33.82699	151.146399
MHPAS01	Middle Harbour	-33.805673	151.25201
MHPAS02	Middle Harbour	-33.816516	151.238761
MHPAS03	Middle Harbour	-33.77408	151.204764
PHPAS01	Port Hacking, Yowie Bay	-34.048201	151.107946
PHPAS02	Gymea Bay	-34.049049	151.09874
PHPAS03	Gymea Bay	-34.049561	151.093225
PHPAS04	Port Hacking, North West Arm	-34.055898	151.085595
PRPAS01	Parramatta River, Rozelle Bay	-33.861729	151.162589
PRPA	Parramatta River, Iron Cove	-33.869851	151.147985
PRPAS03	Parramatta River, Tarban Creek	-33.8364397	151.1412076
PRPAS04	Parramatta River, Hen and Chicken Bay	-33.856479	151.120576
PRPAS05	Parramatta River, Glades Bay	-33.835754	151.118208
PRPAS06	Parramatta River near Homebush Bay	-33.819091	151.072983
SCPAS01	Saltwater Creek, Tripod Street	-33.855217	151.109114
SHPAS01	Sydney Harbour, Chowder Bay	-33.839784	151.253588
SHPAS02	Sydney Harbour, Taylors Bay	-33.843359	151.248891
SHPAS03	Sydney Harbour, Mosman Bay	-33.837349	151.232291

10.2.4 Autosampler monitoring locations

Table 10-5: Autosampler monitoring locations

Site code	Site Name	Latitude	Longitude
GBAUTO_W	Gymea Bay West	-34.049543	151.09321
GBAUTO_E	Gymea Bay East	-34.049313	151.099059
GBAUTO_SEW	Gymea Bay Sewer Carrier	-34.042637	151.111384
LCAUTO_US	Upstream Overflow Lane Cove River	-33.814871	151.142934
LCAUTO_DS	Downstream Overflow Lane Cove River	-33.827025	151.146401
BCAUTO_DS	Downstream Overflow Buffalo Creek	-33.81462	151.12257
BCAUTO_SEW	Sewer Buffalo Creek Carrier	-33.81441	151.12058
DMAUTO_US	Darling Mills Upstream Overflow	-33.754383	151.020353
DMAUTO_DS	Darling Mills Downstream Overflow	-33.769042	151.00538
DMAUTO_SEW	Darling Mills Sewer Carrier	-33.769042	151.00538
VCAUTO_US	Vineyard Upstream Overflow	-33.802388	151.03466
VCAUTO_DS	Vineyard Downstream Overflow	-33.803681	151.031161
VCAUTO_SEW	Vineyard Sewer Carrier	-33.802769	151.032177

10.2.5 Sewer gauge locations

Table 10-6: Sewer gauge monitoring locations

RTU ID number	ERS number (AGN*)	Site Name	Latitude	Longitude
SG0121	2613265	Parker St, Northbridge 2063	-33.805553	151.216639
SO0014	8723124	Bells Rd, Oatlands 2117	-33.802656	151.032289
SO0016	1396981	Kissing Point Rd, Dundas 2117	-33.803019	151.032556
SO0017	1374307	62 Kissing Point Rd, Oatlands 2117	-33.803092	151.032492
SO0035	1355524	Morella Rd, Mosman 2088	-33.83825	151.253556
SO0045	1195800	34 Edensor St., Epping 2121	-33.766564	151.081286
SO0046	1283734	Wilson St., North Ryde 2113	-33.788261	151.114292
SO0049	1165861	Weston St., Fairfield 2165	-33.872722	150.960703
SO0054	1395911	Gloucester Ave Fire Trail, West Pymble 2073	-33.771575	151.125064
SO0058	1367301	Hopkins St., Constitution Hill 2145	-33.799528	150.978669
SO0059	1342314	Lady Game Drv., West Pymble 2073	-33.770011	151.141303
SO0061	1379022	End of Station Ave, Concord West 2138	-33.846703	151.082756
SO0062	1299696	Marina Rd., Baulkham Hills 2153	-33.761989	150.964967
SO0070	1302439	Woodlawn Drive, Toongabbie 2146	-33.78165	150.955028
SO0071	1303330	Peter Pde., Old Toongabbie 2146	-33.785214	150.976981
SO0073	1298380	Sierra Place reserve, Baulkham Hills 2153	-33.765958	150.961142
SO0077	1299724	Burrandong Cres., Baulkham Hills 2153	-33.757428	150.965336
SO0078	1302444	Seven Hills Rd reserve, Baulkham Hills 2153	-33.758731	150.964778
SO0079	1162809	Fairfield Rd., Yennora 2161	-33.860367	150.958019
SO0081	1298935	Cecilia St., Toongabbie 2146	-33.785325	150.950081
SO0123	1197326	19 Plympton Rd., Carlingford 2118	-33.763367	151.064542
SO0125	1336092	49 The Comenarra Parkway, West Pymble 2073	-33.753897	151.118986
SO0126	1393776	224 Midson Rd., Cheltenham 2119	-33.7612	151.068144
SO0128	1360672	Dublin St., Smithfield 2164	-33.842128	150.929664
SO0157	8125282	Leisure Close, Macquarie Park 2113 (aqueduct)	-33.771833	151.124583
SO0161	1201970	16 Morona Ave Fire Trail, Creek 1, Wahroonga 2076	-33.74085	151.092286
SO0162	1329913	16 Morona Ave Fire Trail, Creek 3, Wahroonga 2074	-33.748469	151.098519

RTU ID number	ERS number (AGN*)	Site Name	Latitude	Longitude
SO0166	1354940	The Esplanade on the beach., Mosman 2088	-33.8206	151.250969
SO0167	1352672	38 Monash Cres., Clontarf 2093	-33.807278	151.252344
SO0180	1076162	Stanwell Cres., Ashcroft 2168	-33.9205	150.90125
SO0187	2594033	21 Candowie Cres., Baulkham Hills 2153	-33.753911	151.008711
SO0188	1336048	72 Wallalong Cres Fire Trail, West Pymble 2073	-33.760631	151.119569
SO0196	1370134	Hammers Rd, Northmead 2152	-33.790506	150.987614
SO0197	1301434	110 Best Rd RSL carpark., Seven Hills 2147	-33.778536	150.937125
SP0098	8615839	Sheffield St, Auburn 2144	-33.845549	151.017076
SG0116	9241446	Garrigal National Park, East Lindfield 2070	-33.759842	151.186967
SO0039	1279041	12 Dan St., Marsfield 2122	-33.789197	151.107047
SO0053	1405974	Unwin St. SWC, Earlwood 2206	-33.926661	151.147861
SO0056	1291610	Millwood Ave, Chatswood West 2067	-33.792183	151.161969
SO0066	1354766	45 Avenue Rd., Mosman 2088	-33.834258	151.235022
SO0069	1374946	Parramatta Rd., Auburn 2144	-33.835911	151.023181
SO0075	1394373	Leisure Cl., Macquarie Park 2113	-33.774922	151.126978
SO0076	1299871	Tucks Rd., Seven Hills 2147	-33.780192	150.955778
SO0156	1214419	31 Lady Davidson Cct, Forestville 2087	-33.751958	151.210906
SO0159	1395913	Lady Game Drv aqueduct., West Pymble 2073	-33.772306	151.138117
SO0160	1395908	Yanko Rd Fire Trail, West Pymble 2073	-33.772289	151.132975
SO0164	1398964	53 Moodie St., Rozelle 2039	-33.864125	151.165769
SO0169	1348167	39 Euroka St., Northbridge 2063	-33.808131	151.214103
SO0172	1395351	64 Park Rd., Baulkham Hills 2153	-33.759264	151.010206
SO0174	1165884	Fairfield STP creek, East Pde, Fairfield 2165	-33.882497	150.950342
SO0181	1077190	42 Maxwells Ave., Ashcroft 2168	-33.915394	150.902853
SO0184	1353051	25 Battle Blvd, Seaforth 2092	-33.801558	151.244803
SO0189	8253013	55 Carlyle Rd., Roseville Chase 2069	-33.76975	151.195094
SO0190	1344267	67 Burraneer Ave., St Ives 2075	-33.748761	151.177153
SO0191	1344855	Garrigal National Park. East Lindfield 2070	-33.765656	151.180667
SG802081	2612604	Bestic St Cycleway, Arncliffe 2205	-33.940939	151.157325
SG802153	9194824	Myrtle St, Oatley 2223	-33.982008	151.077036

RTU ID number	ERS number (AGN*)	Site Name	Latitude	Longitude
SG802154	8287486	Acacia St, Oatley 2223	-33.975139	151.067961
SG820685	10330869	Makinson St, Gladesville 2111	-33.831425	151.133836
SO0047	1284398	35 Kent Rd., North Ryde 2113	-33.790953	151.116192
SO0051	1165527	Scrivener St., Warwick Farm 2170	-33.921519	150.937339
SO0055	1251181	Finlays Ave bush track, Earlwood 2206 (on Wolli Creek)	-33.931183	151.131414
SO0057	1216070	Lower Gibbes St, Chatswood 2067 - at Scotts Creek	-33.784631	151.199547
SO0065	1308863	Chiswick Rd., South Granville 2142	-33.859119	151.0135
SO0074	1367809	Neilson St., Granville 2142	-33.8455	151.01605
SO0088	1288806	Reid Drive walking track, Chatswood West 2067	-33.798814	151.161867
SO0098	1205586	12 Alma Rd., Padstow 2211	-33.961272	151.039697
SO0122	1295641	58 Jean St., Seven Hills	-33.772719	150.927
SO0124	1395948	Eastern Arterial Rd., St Ives 2075	-33.751264	151.170125
SO0129	1166102	Gallop St., Warwick Farm 2170	-33.911869	150.928833
SO0131	1395935	22 Stone Pde Fire Trail, Davidson 2085	-33.730308	151.19805
SO0132	1216131	Ferguson St Aqueduct, Forestville 2087	-33.754419	151.207767
SO0155	1345744	55 Carlyle Rd., East Lindfield 2070	-33.769989	151.195181
SO0163	1111719	37 Dobroyd Pde., Haberfield 2045 (No. 37)	-33.871467	151.140569
SO0165	1354086	6 Bay St., Mosman (O/F in No.6)	-33.816761	151.240194
SO0175	1376412	Rumsey Cres., Dundas Valley 2117	-33.796303	151.052333
SO0193	1348581	20 Lodge Rd. mansion, Cremorne 2090	-33.818228	151.230564
SG802080	1403338	Bestic St Cycleway, Arncliffe 2205	-33.942443	151.160018
SG802080N	1403334	Bestic St Cycleway, Arncliffe 2205	-33.942175	151.159663
SG802080S	1406066	Bestic St Cycleway, Arncliffe 2205	-33.942631	151.159986
SG820274	2608776	Babbage Rd, Roseville Chase 2069	-33.774381	151.204553
SO0023	1270675	Peter Place, Gymea Bay 2227	-34.045544	151.0873
SO0024	1269723	56 Ellesmere Rd, Gymea Bay 2227	-34.050992	151.092142
SO0025	1268331	35 Darryl Place, Gymea Bay 2227	-34.046208	151.088814
SO0027	1270308	Forest and Alkaringa Rds, Gymea Bay 2227	-34.042989	151.093828

RTU ID number	ERS number (AGN*)	Site Name	Latitude	Longitude
SO0029	1271600	10 Calypso Place, Yowie Bay 2228	-34.047419	151.097725
SO0030	1157319	8 Maroopna Rd, Yowie Bay 2228	-34.049953	151.099958
SO0050	1285000	10 McCallum Ave, East Ryde 2113	-33.814169	151.136386
SO0052	1162799	Scrivener St., Warwick Farm 2170	-33.921569	150.937228
SO0064	1356955	Iluka Rd., Mosman 2088	-33.842719	151.249340
SO0095	1167934	39 Riverside Drive, Chipping Norton 2170	-33.917397	150.973047
SO0100	1265350	Griffin Pde., Illawong 2234	-33.993	151.043781
SO0113	1106765	McGrath Ave., Five Dock 2046	-33.863136	151.127511
SO0114	1105568	James St., Five Dock 2046	-33.863092	151.123611
SO0115	1105863	Maple Cl., Canada Bay 2046	-33.865669	151.117372
SO0116	1107145	Cnr of Lyons Rd W and Udall Ave, Five Dock	-33.864342	151.128067
SO0117	1104212	William St golf course, Five Dock 2046	-33.866267	151.120131
SO0118	1106196	William St Leisure Centre, Five Dock 2046	-33.867917	151.11811
SO0119	1395946	108 Regatta Rd., Canada Bay 2046	-33.864061	151.117611
SO0120	10490415	East St and 451 Lyons Rd., Five Dock 2046	-33.864358	151.129569
SO0121	1104817	West St and 463 Lyons Rd., Five Dock 2046	-33.8645	151.128836
SO0136	1297216	End of Endeavour St., Seven Hills	-33.775164	150.923664
SO0152	1378306	Andrew St., West Ryde 2114	-33.815619	151.078392
SO0182	1401313	Lawrence Hargrave Rd., Warwick Farm 2170	-33.911475	150.928853
SO0200	1105620	Hutton St riverside walkway, Earlwood 2206	-33.914297	151.125375
SO106	1109448	Premier St, Marrickville 2204	-33.91872	151.149039
SP0187	1165235	Intersection of Church St and Sussex St	-33.903308	150.93548
SG0062	2612415	Henry Lawson Drive, Chipping Norton 2170	-33.917567	150.972922
SO0018	1405851	Robertson Avenue, Seven Hills 2147	-33.769756	150.940631
SO0026	1270699	11 Bunarba Rd, Gymea Bay 2227	-34.042608	151.086897
SO0028	1270320	5 Alkaringa Rd, Gymea Bay 2227	-34.043511	151.093678
SO0037	1284736	Magdala Rd reserve, North Ryde 2113	-33.801661	151.135381
SO0048	1283126	98 Higginbotham Rd., Ryde 2112	-33.8146	151.121992

RTU ID number	ERS number (AGN*)	Site Name	Latitude	Longitude
SO0072	1351570	154 Eastern Valley Way, Castlecrag 2068	-33.799061	151.209656
SO0080	1164994	Homepride Ave., Warwick Farm 2170	-33.913211	150.925972
SO0093	1280229	214 Buffalo Rd, Ryde 2112	-33.810767	151.111422
SO0108	1109420	Thornley St walkway, Marrickville 2204	-33.921044	151.148294
SO0134	1249638	Westfield St., Earlwood 2206	-33.922828	151.114561
SO0135	1104975	Fuller Ave., Earlwood 2206	-33.920083	151.116492
SO0140	1395182	Cnr Williamson Rd & Campbelltown Rds, Ingleburn 2565	-34.006611	150.836319
SO0141	1248253	2-4 Cadia St., Kogarah 2217	-33.959792	151.131528
SO0153	1286788	25 Melba Drv., East Ryde 2113	-33.807569	151.133169
SO0170	1282706	50 Ross St., Gladesville 2111	-33.834025	151.121064
SO0171	1251650	67 Planthurst Rd., Carlton 2218 (rear of No.67)	-33.978689	151.113464
SO0176	1197683	10A Shari Ave., Picnic Point 2213	-33.962997	151.010711
SO0183	1345521	Gaza Rd., Naremburn 2065	-33.812742	151.201267
SO0198	1207055	22 Robyn St., Lugarno 2210	-33.975683	151.053028
SP0138	8138562	Tallawarra Ave, Padstow 2211	-33.956132	151.040972
SO0148	1107278	2a Wellbank St, Concord 2137	-33.855419	151.106558
SG0074	2613296	Greenhaven Rd, Grays Point 2232	-34.049644	151.074547
SG8INT19	1044430	Palmer St, Woolloomooloo 2011	-33.870594	151.218022
SO0036	1042790	Tokanue Pl., St Ives Chase 2075	-33.707536	151.154567
SO0038	1038464	Burns Rd., Turramaurra 2074	-33.716183	151.138064
SO0040	2612856	Matong Place, Gymea Bay 2227	-34.048211	151.090728
SO0082	1265021	304 Forest Rd., Kirrawee 2232	-34.040056	151.074289
SO0083	1270330	Rulwalla Pl., Gymea 2227	-34.040936	151.079386
SO0084	1153257	333 Burraneer Bay Rd., Caringbah South 2229	-34.048467	151.112364
SO0085	1156836	Forest Rd., Yowie Bay 2228	-34.045364	151.106753
SO0096	1154029	17 Tallong Pl., Caringbah South 2229	-34.056386	151.115022
SO0097	2612723	18 Norfolk Place, North Rocks 2151	-33.778036	151.029297

RTU ID number	ERS number (AGN*)	Site Name	Latitude	Longitude
SO0099	2609947	27 William Place, North Rocks 2151	-33.781433	151.014031
SO0101	2609979	14 Northam Drive, North Rocks 2151	-33.778869	151.027311
SO0109	1156756	4 Sherwood Ave., Yowie Bay 2228	-34.052439	151.106372
SO0110	1394840	60 North West Arms Rd, Gymea Bay 2227	-34.043786	151.078542
SO0111	1266636	22 Bilga St., Kirrawee 2232	-34.038217	151.071433
SO0112	1263197	7 Kyogle Place, Grays Point 2232	-34.056333	151.076831
SO0127	1337851	Bedford Ave., North Turramurra 2074	-33.719744	151.150469
SO0133	1115480	Cecil Rd., Hornsby 2077	-33.696956	151.1165
SO0137	1036056	7 King Rd., Asquith 2077	-33.697661	151.10855
SO0138	1393834	145 King Rd SWC Aqueduct, Hornsby 2077	-33.694897	151.119678
SO0142	8512639	145 King Rd SWC Access Chamber, Hornsby 2077 (2x sites)	-33.696861	151.118075
SO0173	1222351	Ben Love Pl carpark., Beacon Hill 2100	-33.750017	151.257367
SO0199	1339727	Merrivale Lane reserve., Turramurra 2074	-33.725775	151.14895
SO0206	1270912	22 Alkaringa Rd Gymea Bay 2227	-34.045361	151.09595
SG8INT16	8122531	Stanley Lane, Darlinghurst 2010	-33.87532	151.214635
SG8INT18	-	Palmer St, Darlinghurst 2010	-33.879556	151.216123
SO0032	7438673	Redgum Drive, Padstow 2211	-33.959292	151.039698
SO0033	1198180	Bottlebrush Place, Alfords Point 2234	-33.988403	151.022516
SO0034	1374920	3 Jillong St Rydalmere 2116	-33.80582	151.04709
SO0006	1185644	Statham Ave, North Rocks 2151	-33.783518	151.023653

* AGN = Asset Generated Number

10.2.6 Rain gauges

The Sydney Water rain gauges listed below provided valuable data for the characterisation of conditions that result in respective overflow discharge.

Table 10-7: Rain gauge locations

Asset code	Location	Latitude	Longitude
5CPS02	RG @ Belmore BC	-33.91771543	151.0924417
566008	RG @ North Ryde Golf Club, access via Twin Rd	-33.80052894	151.1201834
566017	RG @ Chatswood Bowling Club	-33.80140001	151.1803
566018	RG @ Cronulla WRP	-34.0308	151.1643
566020	RG @ Belfield Bowling Club	-33.90049145	151.08777
566022	RG @ SPS41, Homebush	-33.8568767	151.081332
566026	RG @ Marrickville Bowling Club	-33.91034706	151.1639318
566027	RG @ Taronga Zoo, Mosman	-33.84354373	151.2385078
566031	RG @ Revesby Bowling Club	-33.95478792	151.019478
566032	RG @ Army Barracks, Paddington	-33.88677822	151.2248642
566036	RG @ Potts Hill Reservoir	-33.8893	151.0368
566037	RG @ West Ryde Pumping Station	-33.80855532	151.0907091
566038	RG @ Vaucluse Bowling Club	-33.85786831	151.2785294
566047	RG @ Mortdale Bowling Club	-33.97459962	151.080927
566049	RG @ Liverpool WRP	-33.92155357	150.9384468
566053	RG @ Hornsby Heights WWTP	-33.66719934	151.1047443
566056	RG @ Yarrawarrah	-34.0574	151.034
566062	RG @ Bexley Bowling Club	-33.9424	151.1101
566064	RG @ Concord Bowling Club	-33.85587628	151.1065618
566065	RG @ Lilyfield Bowling Club	-33.87606949	151.1615369
566066	RG @ SPS65, Five Dock	-33.85490159	151.1435004
566068	RG @ Dee Why Bowling Club	-33.73869733	151.2811979
566069	RG @ Bankstown Trotting Club	-33.93116471	151.007944
566071	RG @ Belrose Bowling Club	-33.7340152	151.2200075
566072	RG @ Kyle Bay Bowling Club	-33.98828382	151.1023272

Asset code	Location	Latitude	Longitude
566073	RG @ Pymble Bowling Club	-33.74058972	151.1393478
566076	RG @ Pennant Hills Bowling Club	-33.7355	151.0735
566078	RG @ Sth Cronulla Bowling Club	-34.0703145	151.1513923
566080	RG @ Harbord Bowling Club	-33.76823054	151.2841767
566081	RG @ Carlingford Bowling Club	-33.7827	151.0496
566082	RG @ Auburn RSL Bowling Club	-33.860267	151.0190048
566083	RG @ North Epping Bowling Club	-33.75377443	151.091886
566084	RG @ Police Driving School, North St Ives	-33.70908627	151.1877517
566085	RG @ East Lindfield Bowling Club	-33.76319843	151.194809
566087	RG @ Gladesville Bowling Club	-33.82304099	151.1296336
566091	RG @ Kyeemagh RSL Club	-33.946777	151.1608479
566092	RG @ Sutherland Bowling Club	-34.02963247	151.0711536
566098	RG @ Caringbah Bowling Club	-34.036061	151.121117
566100	RG @ North Head WWTP	-33.80800227	151.3018984
566112	RG @ Ashfield Park Bowling Club	-33.8849	151.1348
566113	RG @ Canterbury Racecourse	-33.9068	151.1176
566175	Menai Reservoir (Replacement for 566100 / 566108)	-34.02227568	151.0126101
566176	RG @ Asquith BC	-33.69006055	151.1012551
566178	RG @ Waringah Golf Course	-33.7800802	151.2698699
567076	RG @ Castle Hill WRP	-33.7111	150.9842
567077	RG @ Fairfield WWTP	-33.880781	150.950362
567078	RG @ Glenfield WWTP	-33.982628	150.908039
567079	RG @ Guildford (Water Trunk Main Pipehead)	-33.84679239	150.9687777
567083	RG @ Prospect Reservoir	-33.819328	150.91265
567089	RG @ Prestons SWC Depot	-33.93902712	150.8708105
567102	RG @ Dural Reservoir	-33.697034	151.027943
567104	RG @ Northmead Bowling Club (formerly a UPRCT Gauge)	-33.7817	150.9958

Asset code	Location	Latitude	Longitude
567112	RG @ North Parramatta (Burnside Homes) (formerly a UPRCT Gauge)	-33.791675	151.017966
567115	RG @ Constitution Hill Reservoir	-33.79794536	150.9694984
567145	RG @ Baulkham Hills Reservoir	-33.74413744	150.9871944
567146	RG @ Greystanes Golf Course (formerly a UPRCT Gauge)	-33.8214	150.9416
567148	RG @ NSW Soccer Federation, Kings Langley	-33.74047	150.945903
567149	RG @ Cumberland State Forest, West Pennant Hills (formerly a UPRCT Gauge)	-33.74529278	151.0381455
567151	RG @ Toongabbie Bowling Club (formerly a UPRCT Gauge)	-33.784327	150.952063
567153	RG @ Pinegrove Memorial Park, Minchinbury	-33.8014	150.8384
567154	RG @ Cabramatta Bowling Club	-33.894731	150.945904
567157	RG @ Blacktown BC	-33.79394148	150.8957671
567164	RG @ Raby Reservoir	-33.97179985	150.8276
567169	RG @ Calmsley Hill City Farm, Abbotsbury	-33.869	150.8594
567171	RG @ Seven Hills Reservoir, Summit Place	-33.76118399	150.9586628
568156	RG @ Camden Golf Course	-34.04839351	150.7300568
568174	RG @ Eagleview Rd Reservoir, Minto	-34.0309	150.8614
568175	RG @ Kenny Hill Trunk Receiving Stn	-34.0585	150.7808
568177	RG @ Ingleburn Dam, Varroville	-33.99594596	150.8025197
568179	RG @ Campbelltown Bowling Club	-34.0658	150.8166