Environmental DNA sampling as a tool for monitoring freshwater vertebrates

Emily McColl-Gausden

ORCID ID: 0000-0001-7200-6725

Doctor of Philosophy

February 2022

School of BioSciences

Faculty of Science

The University of Melbourne

Submitted in total fulfilment for the degree of Doctor of Philosophy

Abstract

To monitor biodiversity effectively, accurate and sensitive detection methods that can be implemented over large spatial scales are required. It has been demonstrated that environmental DNA (eDNA) sampling can be a valid alternative to many traditional sampling methods. This thesis explored the data and conclusions which could be drawn from eDNA sampling regimes across large spatial scales for individual species and communities, answering key questions on species occupancy and detection method sensitivity. Adaptations of site occupancy-detection models (SODM) are used throughout.

First, I investigated the relative sensitivity of single- and multi-species detection methods using a consistent methodological framework across multiple datasets with different study designs. I used SODM to evaluate how the molecular method used impacted the probability that a species was detected. I found that qPCR was generally more sensitive at detecting target species but that different methodological decisions impacted its sensitivity.

Second, I used eDNA sampling to conduct an extensive platypus occupancy survey across 37% of the species' estimated distribution. Using a systematic study design, I demonstrated that correlates of platypus occupancy could be revealed using this efficient survey method. I found that platypuses were less likely to occupy sites in areas with a high proportion of surrounding agricultural land or grasslands and were more likely to occupy sites with increased runoff, less zero-flow days, and suitable banks for burrowing.

Third, the 2019/2020 mega-fires impacted south-eastern Australia severely. My pre-fire eDNA data provided the opportunity to investigate if platypus occupancy was impacted by this large-scale event. I developed a SODM extension to incorporate a Before-After Control-impact (BACI) design. After surveying 118 sites in three time periods (pre-fire, 2020 and 2021), I found that platypus occupancy was not significantly impacted by the presence of fire. However, I did find a significant interaction between the proportion of a watershed that was burnt at high severity and rainfall post-fire: platypus occupancy was lower in watersheds with a high proportion of high severity fire that had high rainfall post-fire. This finding is consistent with previous work on the impact of fires on aquatic species.

Lastly, leading on from my work on large-scale eDNA surveys for an individual species, I used eDNA metabarcoding to investigate occupancy patterns among fish communities. I used a

stratified study design across the state of Victoria to investigate how native and introduced fish responded to environmental factors. I found that for the water availability covariates I considered, native and introduced fishes responded similarly. However, for the land use covariates such as the proportion of a contracted catchment covered by agriculture, urban or forested land, different patterns emerged. Native fish responded more positively to these categories, whereas introduced species responded more negatively.

This thesis demonstrated that eDNA sampling is an ideal monitoring tool for individual species or communities over large spatial scales. I showed that SODM and eDNA data can be used to estimate correlates of occupancy efficiently and that eDNA sampling over time enables species responses to major disturbances to be determined across large areas.

Declaration

This is to certify that:

i. This thesis comprises only my original work towards the Doctor of Philosophy, except where indicated in the preface

ii. Due acknowledgement has been made in the text to all other material used; and

iii. The thesis is fewer than 100,000 words in length, exclusive of tables, maps, bibliographies, and appendices

Emily McColl-Gausden

February 2022

Preface

The research presented in this thesis is predominately my own work and was conducted under the supervision of Dr Reid Tingley (The University of Melbourne and Monash University) and Dr Andrew Weeks (The University of Melbourne). I conduced this research as part of the Quantitative and Applied Ecology group, school of BioSciences, at the University of Melbourne. Journal publication has been a priority during my candidature; therefore, my four data chapters have been presented as independent studies in collaboration with co-authors. There is therefore repetition between chapters, and chapters 2 to 5 have their own standalone introductions and discussions. Chapter 1 contains exerts from my published review paper and is cited as McColl-Gausden *et al*. (2020). The pre-print version is available in Appendix A. To acknowledge co-author contributions and to conform with publication standards the collective first-person plural "we" is used throughout chapters 2 to 5.

The chapter contributions of co-authors are outlined below:

Chapter 1

McColl-Gausden EF, Weeks AR, and Tingley R (2020) A field ecologist's guide to environmental DNA sampling in freshwater environments. Australian Zoologist: Vol. 40, No. 4, pp. 641-651. doi: 10.7882/AZ.2019.025

Contributions: EM, RT conceived the research idea. EM wrote the MS with edits and input from RT and AR.

Publication status: Published in Australian Zoologist

Chapter 2

McColl-Gausden EF, Week AR, Coleman R, Song S, Tingley R. Using hierarchical models to compare the sensitivity of metabarcoding and qPCR for eDNA detection

Contributions: EM, RT and AW conceived the research idea. AR and RC facilitated fieldwork and data collection for dataset one. RT and AW obtained funds and EM conducted fieldwork and data collection for dataset two. SS and AW performed laboratory procedures, genetic and bioinformatic analysis. EM and RT designed statistical models and analysis. EM conducted the statistical analysis and wrote the manuscript with inputs from all authors.

Publication status: Unpublished material not submitted for publication

Chapter 3

McColl-Gausden EF, Week AR, Griffiths J and Tingley R. Using eDNA sampling to identify correlates of platypus occupancy across broad spatial scales

Contributions: EM, AW, JG and RT conceived the research idea. EM and JG conducted fieldwork. AW oversaw laboratory procedures and genetic analysis. EM and RT designed the statistical analysis. EM conducted the statistical analysis and wrote the manuscript with inputs from all authors.

Publication status: Unpublished material not submitted for publication

Chapter 4

McColl-Gausden EF, Griffiths J, Collins L, Week AR, Tingley R. The impact of the 2019/2020 mega-fires on platypus occupancy in south-eastern Australia.

Contributions: EM, JG, AR, and RT conceived the research idea. EM and JG conducted the fieldwork with the assistance of other partners. LC produced the fire severity maps. AW and oversaw laboratory procedures and genetic analysis. EM and RT designed the statistical analysis. EM conducted the statistical analysis and wrote the manuscript with inputs from all authors.

Publication status: Unpublished material not submitted for publication

Chapter 5

McColl-Gausden EF, Week AR, Song S, Tingley R. Using eDNA metabarcoding and multispecies occupancy models to investigate freshwater fish distributions over large spatial scales.

Contributions: EM, AR and RT conceived the research idea. EM conducted the fieldwork with the assistance of other partners. AW and SS performed laboratory procedures, genetic and bioinformatic analysis. EM and RT designed the statistical analysis. EM conducted the statistical analysis and wrote the manuscript with inputs from all authors.

Publication status: Unpublished material not submitted for publication

I published a manuscript during my candidature relating to research conducted prior to my PhD:

McColl-Gausden, E. F., Weeks, A. R., Coleman, R. A., Robinson, K. L., Song, S., Raadik, T. A., and Tingley, R. (2021). Multispecies models reveal that eDNA metabarcoding is more sensitive than backpack electrofishing for conducting fish surveys in freshwater streams. *Molecular Ecology* **30**, 3111–3126. doi:10.1111/mec.15644

Results from the research in this thesis has been presented at the following conferences:

2018

Ecological Society of Australia conference, Brisbane, Australia

2019

Environmental Institute of Australia and New Zealand- Special forum on using eDNA for wildlife detection, Sydney, Australia

Ecological Society of Australia conference, Launceston, Australia

2020

International Statistical Ecology Conference, virtual

Ecological Society of Australia conference, virtual

2021

Ecological Society of Australia conference, virtual

Research permits

Fieldwork for chapters 3, 4 and 5 was conducted under research permits for New South Wales (NSW Scientific license #SL102217 and Forestry corporation # RES100002) and Victoria (Parks Permit #10009047).

Funding support

My candidature has been supported by an Australian Government Research Training Program Scholarship. Additional funding was acquired from the San Diego Zoo Wildlife Alliance, the Holsworth Wildlife Research endowment (the Ecological Society of Australia) and the Jasper Loftus-Hills award (The University of Melbourne).

Acknowledgements

First, I would like to acknowledge the Wurundjeri people as the traditional owners of the land on which I lived and worked while this research was being conducted.

There are so many people I want to thank and acknowledge for their support, advice, and friendship over the past few years. My PhD journey was not how I originally envisioned it, more than half of it now being spent under COVID-19 related restrictions and the world's longest lockdowns. My support team, lab group and friends helped me through this crazy time in our history, and a crazy and exciting time in my career.

A huge thank you to my supervisors Reid Tingley and Andrew Weeks, I'm so grateful you supported and trusted me with this project. Thank you especially to Reid for always being up for a chat (or zoom chat) about my study designs, models, and writing. Your support is greatly appreciated. Thanks also to Mick McCarthy and Yung En Chee for being my advisory committee and providing advice and support, especially in the early project development phase, and of course when navigating the complications of covid related extensions. Thanks also to Josh Griffiths for being my go-to platypus and fieldwork guru, as well as the team at enviroDNA for processing my 1000's of samples.

My fieldwork would not have been possible without the help of my lovely volunteers! So, thank you to Lucy Howell, Jessie Moyses, Kelsey Johnson, Matthew Aoto, Alys Young, Erica Marshall, Emily Gregg, Sarah McColl-Gausden, Barbara McColl, Roozbeh Valavi, Jake Barker, Matt Rees, Arabella Eyre, and Rebecca Groenewegen for driving around VIC and NSW with me and taking water samples, we got to go to some pretty amazing places!

I am eternally grateful that I fell into the Quantitative and Applied Ecology group for my master's project. Over the past 6 years, and two degrees, everyone has been so supportive, collaborative and fun. Over lab retreats, after work drinks, ski trips, zoom lab meetings and games nights, it's been so great getting to know everyone and experiencing so much together. I have developed many lasting friendships and connections in this group that I am so grateful to have. Special thanks to my mentor Natalie Briscoe (and everyone's unofficial mentor Pia Lentini) for our chats, advice and just for being such great and inspirational people!

My "cohort" of PhD students (term applied very loosely) has been amazing in their friendship and support! Thanks August, Roozbeh, Martin, Adam, K, Huey, Mat, Bec, Linda, Annalie,

Sarah, Kate and Erica! Extra special thanks to Erica, I'm so glad you decided to move to Melbourne for your PhD. You have been the best friend, support person, fieldwork buddy and housemate anyone could ask for. Your positive energy is infectious, and I couldn't ask for someone better to be locked in a house for 2 years with. Special thanks also to Kate for our Friday night chats, your practical fieldwork advice and friendship over these past few years.

I feel like my housemates need a special mention because of the countless days spent working from home. In 2020, Erica, Callum and Bec, and in 2021 Erica, Callum and Sarah, thanks for the movie nights, takeaways, "Wednesdays", and laughs and for being so patient in such difficult circumstances!

Thank you to my family! My mum Barbara, for always being my cheer squad and excited by what I was doing. Thanks for helping me on fieldwork, I'm so glad you got to be a part of my PhD journey. A huge thank you to my twin sister Sarah. We started our PhDs on the same day and have gone through all the ups and downs of our PhDs together (even when separated by covid restrictions). Thank you for all the support, fieldwork help, Disney videos and travel planning. I'm so grateful we got to go on this PhD adventure together and share so many wonderful experiences, with many more to come, I'm sure!

Thank you to everyone who has supported me over the past almost 4 years!

Table of contents

List of Tables and Figures

Chapter 1
General introduction

1.1 Introduction

Understanding spatial changes in biodiversity patterns is important for informing conservation efforts, but landscape-scale data on species distributions can be logistically and financially difficult to collect. Environmental DNA (eDNA) sampling —the detection of extra-organismal DNA sourced from environmental samples without any obvious biological source material being present (Taberlet *et al.* 2012)— is an emerging survey technique that could, for many species, facilitate efficient and cost-effective collection of landscape-level data. DNA can be shed into the environment from various sources, including skin cells, mucous, faeces, or even individual hairs. Previous studies have shown samples of water, snow, soil, and air can be used to determine species presence or absence at a site (Dalén *et al.* 2007; Andersen *et al.* 2012; Lugg *et al.* 2017; Leontidou *et al.* 2018).

The first step in assessing the suitability of any emerging technology, such as eDNA sampling, is to compare the results it produces against traditional sampling methods (Lahoz-Monfort and Tingley 2018). As with other emerging technologies in ecology, such as drones (Hodgson *et al.* 2016), thermal cameras (Goodenough *et al.* 2017), and detector dogs (Cristescu *et al.* 2015), eDNA sampling in aquatic environments has been compared to traditional survey methods, such as live trapping (Shaw *et al.* 2016), aural surveys (Valentini *et al.* 2016), and electrofishing (Evans and Lamberti 2017). In many aquatic systems tested to date, eDNA sampling has been capable of detecting species or communities more effectively than traditional methods (Fediajevaite *et al.* 2021).

Environmental DNA sampling has many elements that make it a promising survey method in aquatic environments. It allows for large-scale data collection to be efficient for many species without invasive sampling (e.g., capturing), or even sighting animals. Thus, eDNA sampling largely eliminates concerns around animal ethics. Additionally, health and safety concerns for researchers and employers are reduced when using this method. For example, eDNA sampling eliminates high-risk activities, such as entering waterways to check fyke nets for traditional fish and platypus surveys (Serena 1994) or conducting backpack or boat electrofishing surveys for fish (Wilcox *et al.* 2016). Environmental DNA sampling can potentially detect cryptic or rare species that evade detection with other methods and detect species at life stages that were previously difficult to detect or identify between (Dejean *et al.* 2012). It also lends itself well to baseline data collection on species distributions alongside routine monitoring programs that

aim to track changes in species distributions over space and time. If more detailed data on population or individual health (e.g., reproductive output, juvenile recruitment, sex, genetic variation, abundance) are required, eDNA can highlight sampling locations for more traditional, time-intensive capture methods. Moreover, eDNA methods also show some promise in this regard (providing population- and individual-based information), particularly as technologies improve (Sigsgaard *et al.* 2016; Bylemans *et al.* 2017).

Despite these benefits, using eDNA sampling in biodiversity surveys is currently more complex than using most traditional survey methods. This is largely because eDNA sampling is a relatively new monitoring technique and standardised approaches have not been developed (Goldberg *et al.* 2016). There is however a large global interest in the standardisation of eDNA detection methods on a large scale with improvements and recommendations being developed (Bohmann *et al.* 2021).

Single- and multi-species detection methods

Environmental DNA detection methods can be divided into two broad categories: single- and multi-species methods (Figure 1.1). Single-species methods focus on eDNA quantification from a single target species using primers/probes that are specific to the species or population(s) of interest. Most single-species studies use real-time quantitative polymerase chain reaction (qPCR), although traditional PCR (Goldberg *et al.* 2011), Droplet Digital PCR (ddPCR) (Doi *et al.* 2015) and loop-mediated isothermal amplication (LAMP) (Lee 2017) have also been used. Single-species detection methods have been used broadly to detect rare (Laramie *et al.* 2015; Schmelzle and Kinziger 2016; Simpfendorfer *et al.* 2016) and invasive (Hunter *et al.* 2015; Hinlo *et al.* 2017; Klymus *et al.* 2017; Tingley *et al.* 2019) species in aquatic environments. Multi-species detection methods —also known as metabarcoding— take a broader, community-focused approach. High-throughput next generation sequencing (NGS) technologies for DNA sequencing enable all species from one or more target groups (e.g., fish, amphibians, or decapods) to be identified (Taberlet *et al.* 2012).

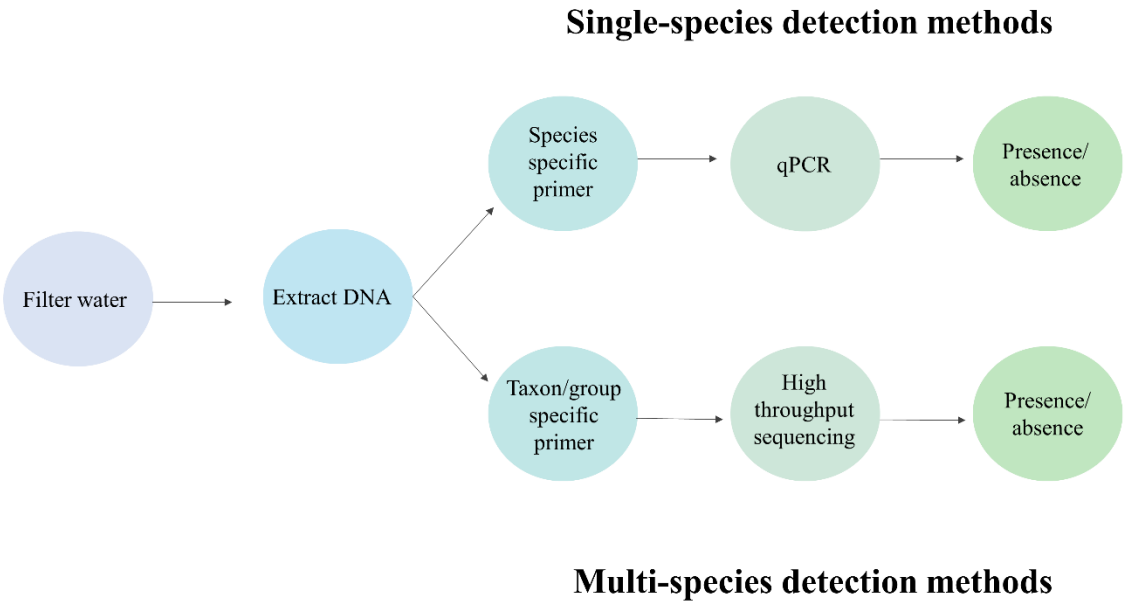

Figure 1.1*: Simplified outline of process for single- and multi-species eDNA detection methods*

Whether to adopt a single- or multi-species eDNA approach depends primarily on the specific study aims. A single-species approach is not appropriate when the scientific or management objective pertains to community-level patterns. But given the rich amount of data afforded by metabarcoding, why would one adopt a single-species approach, even if a study intends to focus solely on a single species? The answer to this question lies primarily in the relative complexity, sensitivity, and cost-efficiency of each eDNA approach.

Single-species approaches require species-specific primers/probes for a target DNA region, whereas metabarcoding requires primers designed to bind to conserved DNA regions across the target group. The ability to distinguish species within the target group then relies on a reference library of known sequences matched to species, with which the sequenced eDNA data can be compared. Publicly available sequence databases, such as GenBank (Clark *et al.* 2016), can be used to generate reference libraries at a broad level, but often Australian native species are poorly represented in these databases due to their uniqueness. Therefore, a local reference library, consisting of genetic sequences of each species from the study area (as in Valentini *et al.* 2016), is recommended; this also helps account for any intraspecific variation within the gene region being sequenced. Once this library has been established, it can be used by future studies; however, its initial development can be expensive and time-consuming relative to the resources needed to design a species-specific primer. Different target gene regions will also provide different levels of species resolution, and this needs to be considered when undertaking metabarcoding surveys, as some gene regions may not differentiate between

closely related species (Bylemans *et al.* 2018). Primer bias can also result in some species not being detected in a sample when their DNA may be present, as some species may amplify more readily than others (Elbrecht and Leese 2017).

Metabarcoding requires bioinformatic pipelines to assign genetic sequences to samples and genetic sequences (or haplotypes) to species (Coissac *et al.* 2012). Once again, these pipelines can be reused once developed, but the bioinformatic tools needed for metabarcoding are much more labour-intensive and require a different level of expertise compared to a single-species analysis, which is much more routine in a molecular laboratory.

Few studies have directly compared the sensitivity of single- vs multi-species eDNA assays for individual species. Harper *et al*. (2018) demonstrated that qPCR resulted in greater detectability for Great Crested Newts (*Triturus cristatus*) relative to metabarcoding. Wood *et al*. (2019) compared different targeted (singles-species) and metabarcoding (multi-species) approaches for detecting *Sabella spallanzanii* —a Mediterranean fanworm. They demonstrated that targeted single-species detection methods resulted in a higher detection probability compared to the metabarcoding approach, with ddPCR having the highest probability of detection. Bylemans *et al.* (2019) used targeted and metabarcoding methods to detect redfin perch (*Perca fluviatilis*) in an intermittent river system. The targeted survey (qPCR) produced higher detection probabilities compared to eDNA metabarcoding, with metabarcoding failing to detect the species at six sampling sites where qPCR confirmed presence. Nonetheless, any increased sensitivity needs to be evaluated against a study's aims and the nature of the study system. In some cases, accepting a lower detection probability for a target species may be worthwhile if additional data on biotic interactions (e.g., presence of prey, competitors, predators, disease) would change scientific inferences or management decisions. In other cases, such as determining whether an endangered species is present at a site for an environmental impact assessment, a less sensitive metabarcoding assay may not be the optimal choice (Deiner *et al.* 2017). There may be situations where combining approaches (metabarcoding and a single-species assay) is ideal, particularly where sensitivity for the target species is paramount in the survey design, but there is also a need for the additional data on community composition. The advantage here is that sampling and DNA extraction have already been undertaken for one method, significantly reducing cost for the second method.

The reliability of abundance estimates from eDNA samples varies between single- and multi-species detection methods. Single-species eDNA assays using qPCR, ddPCR or LAMP have

revealed relationships between DNA copy number and species abundance within a sample, indicating that the method can be used to estimate abundance at a site (Klobucar *et al.* 2017; Tillotson *et al.* 2018; Yates *et al.* 2019). This provides additional information beyond just presence-absence and is therefore an additional consideration when designing an eDNA study. While, in theory, multi-species eDNA approaches can also estimate a measure of DNA copy number within a sample, this relationship is much more complex in metabarcoding due to the methodology employed, and therefore generally considered a relatively weak indicator of species abundance at a site (Fonseca 2018).

The final consideration as to whether to adopt a single- or multi-species approach —cost-efficiency— has not yet been addressed. Metabarcoding is considerably more expensive than single-species PCR-based approaches when only considering a single target species, but it remains to be seen at what point (e.g., number of species) metabarcoding is more cost-efficient. This will depend on the number of species present in a system and the availability of single-species eDNA assays (Shaw *et al.* 2017).

Large-scale, systematic eDNA studies

Aquatic eDNA sampling is well-suited to surveys of species distributions, which could be at the local, regional, national, or continental scale, depending on the species range. These surveys are important for monitoring purposes, especially for conservation assessments, such as the global International Union for the Conservation of Nature (IUCN) Red List, or for the Australian National Environment Protection and Biodiversity Conservation (EPBC) Act or regional-based listings. They are also essential for understanding and predicting effects of climate or other global changes. Surveillance sampling for potential invasive species, especially over large spatial scales, is another important avenue where eDNA sampling could fill a sampling need (Tingley *et al.* 2019). Environmental DNA sampling is ideal for the collection of broad-scale data, as it can be conducted very quickly at a site, and methods and survey effort can be consistent across the study area. For example, the volume of water may be recorded for each sample, and two samples may be collected at each site, standardising the sampling approach. Large-scale sampling using aquatic eDNA has already been demonstrated for some species. For example, Biggs *et al*. (2015) sampled 75% of the great crested newts' (*Triturus cristatus*) 217,000 km^2 range using a citizen science program. Such broad-scale data can be difficult to collect using traditional sampling methods, such as those that involve capturing individuals. Many traditional sampling methods are logistically challenging and

require a large amount of time and effort to coordinate at larger scales, for example setting up and monitoring traps. Additionally, some species may only be detected at certain times of year, therefore it may be impractical to cover a large sampling area with some traditional methods. Sampling over large spatial scales may then have to occur over multiple years, adding a time cost onto data collection efforts. Alternatively, eDNA sampling can occur over large-spatial scales in a very short space of time. The sensitivity of some traditional sampling methods may also be low. Being able to detect a species when it is present at a site is important. For example, some species may avoid traps or not be captured easily (Griffiths *et al.* 2013), which leads to imperfect detection and can result in inaccurate occupancy estimates for a species (Guillera-Arroita *et al.* 2017).

Large-scale eDNA surveys for generalist species can cover large geographic areas, with variation in climate, environmental gradients, and habitat. During these surveys it is important that sampling is not biased towards specific environments or climates and covers the breadth of potential habitats. Random or stratified sampling approaches are designed to control or reduce sampling variability (Williams and Brown 2019). Stratified random sampling involves dividing the environment up into different strata —e.g., vegetation or land use types— and then taking random samples from within each strata. This ensures that samples are taken across the environmental range where a species may be present.

Platypus occupancy and eDNA

Platypuses (*Ornithorhynchus anatinus*) are an iconic Australian monotreme, endemic to eastern Australia. They are nocturnal and semi-aquatic, feeding in streams and rivers but also occupying burrows dug into riverbanks (Gust and Handasyde 1995). Platypuses feed nocturnally on benthic aquatic macroinvertebrates (Serena *et al.* 2001). They occupy a large area, with their estimated extent of occurrence being ~1.1 million km^2 in eastern Australia (Figure 1.2). Climate varies considerably across the longitudinal extent of their range. In the south they occupy temperate and even alpine Tasmania, and in far north Queensland they occupy tropical environments (Grant and Temple-Smith 2003). Platypuses are classed as vulnerable in the state of Victoria and are classed as Near Threatened by the IUCN (Woinarski and Burbidge 2016). They are predicted to be impacted by the effects of climate change through changes such as increasing temperatures and droughts (Klamt *et al.* 2011; Bino *et al.* 2020; Bino *et al.* 2021).

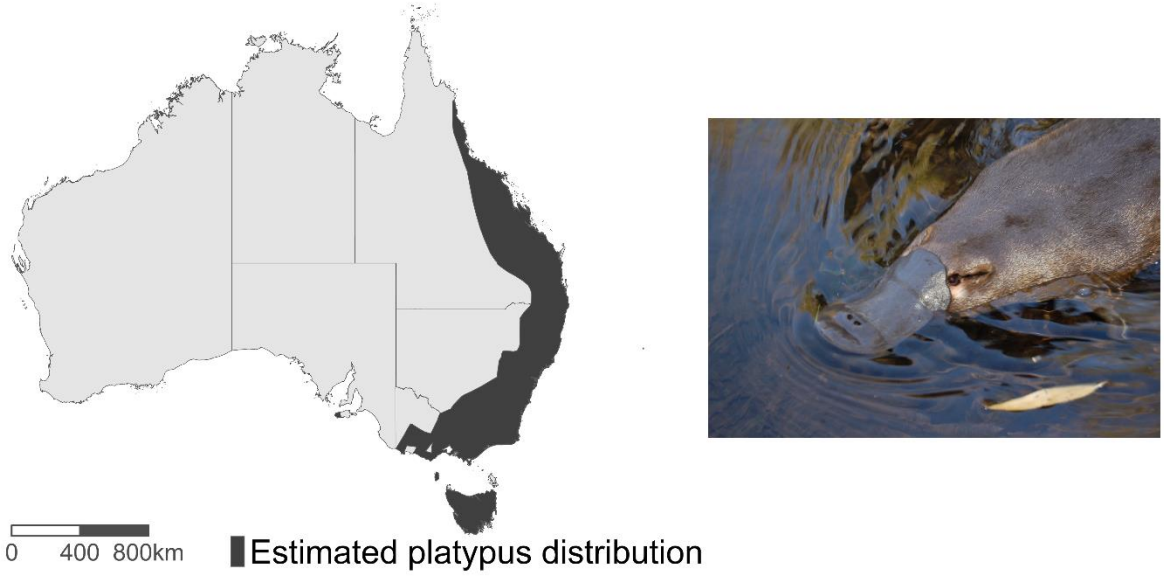

Figure 1.2*: From left to right; the estimated platypus distribution across Australia (Woinarski and Burbidge 2016) and an image of a platypus in water (Ornithorhynchus anatinus). Photo credit Josh Griffiths.*

As platypuses are semi-aquatic and predominately hunt at night, incidental observations of platypuses are rare. Therefore, traditional methods to survey them such as trapping —e.g., fyke or gill nets— are generally used. These nets must be placed in the river before dark and checked regularly throughout the night (Griffiths *et al.* 2013; Hawke *et al.* 2021). This is an intensive method that cannot be used over large spatial scales efficiently due to time and resource constraints. Trapping has other benefits, however, as platypuses can be physically examined and therefore data about their health status, sex, age and movement can be obtained (Hawke *et al.* 2020).

Environmental DNA sampling provides the ideal opportunity to demonstrate efficient large-scale sampling which has not yet been applied to species occupancy surveys. Environmental DNA sampling has already been shown to be more sensitive at detecting platypuses compared to fyke netting (Lugg *et al.* 2017). Lugg *et al.*, (2017) demonstrated the conditional probability of platypus DNA being detected by qPCR in a water sample was between 0.858 and 0.892, whereas the conditional detection probability for a single trapping visit was between 0.219 and 0.470. Environmental DNA sampling was also more cost-efficient, regardless of whether the management objective was to 1) minimize the survey budget needed to achieve a particular asymptotic occupancy estimator variance, or (2) minimize the survey budget needed to detect a change in occupancy over time (Lugg *et al.* 2017). Platypuses also provide the ideal case study for large spatial-scale sampling as they are present over a large extent across different

climatic gradients (Filkov *et al.* 2020). They are also likely to be impacted by climate change and therefore baseline data on their occupancy is critical for understanding how they will respond to changes in the future. Baseline data is lacking for platypuses across their range because of the limitations of current sampling techniques, especially in remote and hard to access areas. The combination of a rapid and efficient detection method and a high detection probability means eDNA sampling is an ideal method for conducting large-scale platypus occupancy studies.

eDNA sampling to estimate impacts of disturbance

Much of the Australian landscape has been shaped by fire and fire regimes. However, the 2019/2020 mega-fires that burnt over 19 million hectares across south-eastern Australia (Filkov *et al.* 2020) caused significant impacts on people, property, and the environment through their unprecedented scale and severity (Collins *et al.* 2021). During these fires, there was the loss of 33 lives and 3,000 houses (Filkov *et al.* 2020). Large and destructive fires are increasing globally, and climate change is driving the climatic conditions conducive for these large fires (Collins *et al.* 2021). The impact of fire and increasingly severe large fires on biodiversity, particularly aquatic biodiversity, is uncertain due to lack of research in this area. For example, it is unclear whether individuals are directly killed by fires or if they are impacted by longer-term post-fire impacts such as food or habitat availability. There is some evidence that the majority of animals are not directly killed by fire, especially under low severity or planned fires (Jolly *et al.* 2022). However, post-fire impacts in aquatic habitats could include lower water quality due to post-fire erosion combined with rainfall, and reduced vegetation cover, as well as the potential for sediment flows and black water (deoxygenation) events (Lyon and O'Connor 2008; Verkaik *et al.* 2014).

There is little known about how platypuses respond to fire, including large-scale high severity fires. Being semi-aquatic, they could be buffered from the direct effects of fire; however, post-fire impacts could include habitat degradation, sediment slugs or decreased water quality (Wilkinson *et al.* 2007) as well as a reduction in prey availability (Vieira *et al.* 2004) all due to post-fire rainfall and erosion. High severity fires can increase the probability of water quality impacts (Wilkinson *et al.* 2007). Aquatic macroinvertebrates, a primary food source for platypuses (Serena *et al.* 2014), can be impacted in terms of their assemblages and richness after fire (Verkaik *et al.* 2014; Verkaik *et al.* 2015; Robson *et al.* 2018). Therefore, any impacts on macroinvertebrates could impact the survival of platypuses after a wildfire. The combination

of high rainfall events and flooding after fire has also been shown to reduce the abundance of macroinvertebrates (Vieira *et al.* 2004), and fish (Shakesby *et al.* 2007).

As eDNA sampling is an efficient survey method for many aquatic and even terrestrial animals, it is an ideal method to rapidly assess biodiversity after an impact or disturbance. For example, the California environmental DNA "CALeDNA" program has been using citizen scientists to take soil samples for eDNA analysis of paired burnt and unburnt sites (Meyer *et al.* 2019). Environmental DNA has also been used to assess benthic community composition associated with deep-sea hard substrata, and recolonisation after an induced disturbance (Cowart *et al.* 2020). Efficient eDNA sampling enables consistent baseline data that can be used to estimate the impact of future disturbances, such as major wildfires. Pre-fire occupancy or distribution data for many species was noted as a limiting factor in prioritising efforts after the 2019/2020 fires (Legge *et al.* 2021; Southwell *et al.* 2021). Pre-impact data (also known as "before" data), especially for rare or cryptic species, can be difficult to obtain, and therefore the true disturbance impact is hard to measure (Christie *et al.* 2019). Indeed, Christie *et al*. (2019) demonstrated that Before-After Control-Impact (BACI) designs performed 3.2–4.6 times better than Control-Impact and 7.1–10.1 times better than After-Impact designs, depending on sample size.

Site occupancy-detection models

Detection is likely to be imperfect with eDNA data. Accounting for imperfect detection, using any survey technique, is critical for ensuring accurate estimates of site occupancy. Fortunately, a rich statistical literature on species occupancy-detection models (SODM) has been developed to account for imperfect detection in wildlife survey data, and these models are a natural framework in which to analyse eDNA-based data (Dorazio and Erickson 2018; Strickland and Roberts 2019). SODM estimate site occupancy whilst accounting for imperfect detection, and thus are ideal for eDNA data, in which replicate samples are collected at a set of sites, and there is potential for both false positive and false negative detections (Schmidt *et al.* 2013; Lahoz-Monfort *et al.* 2016; Guillera-Arroita *et al.* 2017; Lugg *et al.* 2017).

Hierarchical SODM, which account for nested detection processes, are particularly useful for the analysis of eDNA data. This is because multiple PCRs are often nested within multiple water samples taken at each site. For example, Hunter *et al.* (2015), Lugg *et al.* (2017) and Schmidt *et al.* (2013) use a three-level SODM considering (i) the latent presence/absence of a

species, given the occupancy probability (Equation 1); (ii) the probability of the eDNA being contained (or 'available') within the water sample given (i) (Equation 2); and (iii) the probability of detecting the species' eDNA using PCR given (ii) (Equation 3). Using such hierarchical models enables imperfect detection to be incorporated at the different levels of the eDNA sampling process. Guillera-Arroita *et al.* (2017) have shown how these models can be extended to account for false positive detections.

Equation 1. $$z_i \sim Bernoulli(\psi_i)$$

Equation 2. $$a_{ij}|z_i \sim Bernoulli\ (z_i\theta_{ij})$$

Equation 3. $$y_{ijk}|a_{ij} \sim Bernoulli(a_{ij}p_{ijk})$$

SODM can also be applied to multi-species metabarcoding. For example, Valentini *et al*. (2016) used SODM to estimate detection probabilities for multiple amphibian species and (Mcclenaghan *et al.* 2020) used multi-scale, multi-species occupancy models to analyse coastal marine biodiversity eDNA data. Environmental covariates can also be incorporated, enabling occupancy to be compared across different environments.

1.2 Thesis aims

This thesis aims to improve understanding of how eDNA sampling can be used to monitor aquatic vertebrate species and assess threats to their persistence. It evaluates the difference between the two major detection methods in eDNA sampling and uses site occupancy-detection models to demonstrate that the impacts of environmental factors and disturbances on species can be investigated at broad spatial scales.

1.3 Thesis outline

Chapter 2: Using hierarchical models to compare the sensitivity of metabarcoding and qPCR for eDNA detection

In this chapter, I compare the sensitivity of two different eDNA detection methods: eDNA metabarcoding and a single-species detection method. This chapter aims to inform which detection method should be used for eDNA sampling.

Chapter 3: Using eDNA sampling to identify correlates of platypus occupancy across broad spatial scales

Large-scale species occupancy studies can be extremely difficult using traditional methods for some species. In this chapter, I demonstrate the benefits of eDNA sampling as a monitoring tool for vertebrates by conducting large-scale aquatic sampling for platypuses across environmental gradients using a systematic study design. This chapter aims to determine platypus occupancy across south-eastern Australia and uses site occupancy-detection models to investigate environmental correlates of platypus distribution.

Chapter 4: The impact of the 2019/2020 mega-fires on platypus occupancy in south-eastern Australia

Even though platypuses are aquatic, there is the potential for them to be impacted by large-scale wildfires, such as the Australian "black summer" mega-fires of 2019/2020. This chapters aims to investigate whether platypus occupancy changed after the mega-fires using eDNA data collected before and after the fires and site occupancy-detection models. The impacts of post-fire rain (and therefore the potential for sediment flows and water quality impacts) and fire severity are also investigated.

Chapter 5: Using eDNA metabarcoding and multispecies occupancy models to investigate freshwater fish distributions over large spatial scales

Environmental DNA metabarcoding has an advantage over single-species methods as the presence or absence of multiple species' DNA can be determined using the same sample. This chapter aims to investigate how native and introduced fish species respond to environmental factors such as land use or zero-flow events using data collected over broad spatial scales and

a systematic study design. It also investigates whether introduced fish species richness impacts native species occupancy.

Chapter 6: General discussion, synthesis, and future directions

In this chapter I summarise each chapter's findings. I discuss their significance in relation to the current literature and outline future directions for this work.

Appendices

Appendices contain supplementary materials for each chapter (Appendix B-E) as well as the author accepted version of a published review entitled "A field ecologists guide to environmental DNA sampling" which contributed to this general introduction (Appendix A).

Chapter 2
Using hierarchical models to compare the sensitivity of metabarcoding and qPCR for eDNA detection

2.1 Abstract

Environmental DNA (eDNA) sampling —the detection of intra- or extra-cellular DNA in environmental samples— is a rapid and sensitive survey method for detecting aquatic species. The two main detection methods used on eDNA samples target either single- or multiple species. Single-species detection methods (typically based on PCR or LAMP) have been shown to be more sensitive for detecting target species than multi-species detection methods, such as metabarcoding. However, previous studies have generally only compared these two eDNA detection approaches for a single target species and have used different methodological and statistical approaches. Here we present a comparison of single- and multi-species eDNA detection methods, drawing on two published case studies (one fish, one amphibian) and two new extensive datasets on a freshwater mammal (the platypus). To ensure consistent conclusions regarding the sensitivity of each eDNA method, we use the same hierarchical site occupancy-detection model for each dataset, incorporating uncertainty at the site, water sample, and technical replicate level. Overall, qPCR achieved higher detection probabilities than metabarcoding across species and datasets. However, differences in sensitivity between detection methods varied depending on methodological decisions concerning what constitutes a true positive detection (i.e., qPCR and metabarcoding thresholds). The decision as to which eDNA detection method to use should always be influenced by the study aims, but our results suggest that single-species detection methods based on qPCR may be preferable when the aim is to achieve a high detection probability for target species.

2.2 Introduction

Accurate biodiversity monitoring requires sensitive and fit-for-purpose sampling methods. Environmental DNA (eDNA) sampling —detection of intra- or extra-cellular DNA that organisms leave behind in environmental samples such as water, air, or soil— has gained increasing attention in this regard in recent years. Environmental DNA detection approaches can be classified into two broad categories: single-species or multi-species detection methods (McColl-Gausden *et al.* 2020). Both eDNA detection approaches have been shown to be equally (or more) sensitive at detecting species compared to relevant traditional survey methods (e.g. Blackman *et al*. 2020; Lugg *et al*. 2017; Mcdevitt *et al*. 2019; Mcelroy *et al*. 2020; Smart, Tingley, Weeks, Van Rooyen, & McCarthy, 2015; Valentini *et al*. 2016).

Single-species detection methods were adopted earlier in ecological contexts compared to multi-species methods (Valentini *et al.* 2016). With single-species methods, a target species' DNA is amplified using PCR, or variants thereof, such as quantitative PCR (qPCR), droplet digital PCR (ddPCR) or LAMP (Loop-Mediated Isothermal Amplification), coupled with a primer pair (or pairs in the case of LAMP), and in some cases a probe, that is designed to be specific to the target species. The presence or absence of a species' DNA in one or more replicate environmental samples is then used to infer its presence or absence at a site. Some studies have also found a relationship between the quantity of DNA from qPCR and species abundance to enable comparative estimates between study sites (Tillotson *et al.* 2018; Yates *et al.* 2019; Spear *et al.* 2021). This relationship depends on the species, however, and can be confounded by life stages and environmental conditions (Plough *et al.* 2018; Tillotson *et al.* 2018), such as downstream transport (Rice *et al.* 2018).

Multi-species eDNA detection methods (also known as eDNA metabarcoding) use next generation sequencing (NGS) techniques. With multi-species methods, instead of a species-specific primer pair(s)/probe, universal primers are developed that target an entire taxonomic group e.g., subphylum or class. This technique requires a different level of expertise that is not as routinely used in a molecular laboratory, compared to the expertise required for the single-species approach. For example, decisions need to be made about bioinformatic processing and filtering (Deiner *et al.* 2017; McColl-Gausden *et al.* 2021), OTU selection and removal (Deiner *et al.* 2017), reference libraries (Schenekar *et al.* 2020), and confounding factors, such as primer bias (Kelly *et al.* 2019; Schenekar *et al.* 2020). Despite these complications, eDNA

metabarcoding is becoming increasingly prevalent in biodiversity monitoring, as it enables large-scale monitoring of multiple species that would otherwise be substantially more costly or time consuming using traditional sampling methods (Lacoursière-Roussel *et al.* 2016; Ratcliffe *et al.* 2021; West *et al.* 2021).

The decision to use a single- or multi-species eDNA detection method will depend on several factors, such as the study aims, survey budget, project team expertise, and knowledge of the study system. For example, single-species qPCR will be much cheaper, quicker, and easier to implement than eDNA metabarcoding when the primary aim is to document the distributions of only a few species, whereas eDNA metabarcoding may be more appropriate when the aim is to document patterns of species richness or community composition. Environmental DNA metabarcoding can also be useful for systems in which there is little *a priori* knowledge on species composition. However, the sensitivity and cost of single-species vs multi-species methods must also be carefully considered when choosing between eDNA detection methods, particularly when the presence or absence of specific species is of interest (e.g., threatened, invasive, or keystone species).

Several studies have compared single-species eDNA detection methods to multi-species detection methods. Although sampling characteristics have varied across studies, the consistent conclusion from these studies is that single-species detection methods, such as qPCR, result in greater detection probabilities of target species compared to eDNA metabarcoding. For example, Harper *et al*. (2019) concluded that qPCR resulted in a higher detection rate for the great crested newt (*Triturus cristatus*) than eDNA metabarcoding, but that detection rate depended on the number of positive qPCR replicates used to classify a water sample as positive (1/12 vs 4/12). Wood *et al*. (2019) developed eDNA assays for two different gene regions (COI, 18S rRNA) to compare single-species detection methods (qPCR, ddPCR) and a multi-species method for detecting a Mediterranean fanworm, *Sabella spallanzanii*. They demonstrated that targeted single-species detection methods resulted in higher detection probabilities than metabarcoding, with ddPCR having the highest probability of detection. Schenekar *et al.* (2020) found that European catfish *(Silurus glanis*) and European eels (*Anguilla anguilla*) were detected at more sites with qPCR than with metabarcoding. Finally, (Bylemans *et al.* 2019) used single- and multi-species detection methods for redfin perch (*Perca fluviatilis*) in an intermittent river system, concluding that qPCR produced higher detection probabilities

compared to eDNA metabarcoding, with metabarcoding failing to detect the species at six sites where qPCR confirmed presence.

Previous comparative studies, such as those outlined above, provide important insight into the relative sensitivity of single- and multi-species detection methods. Yet most previous studies have had one or more methodological limitation. First, studies have predominately focused on detecting one or two target species. Numerous studies have shown that the sensitivity of eDNA detection methods can vary considerably among taxa and systems. While individual analyses add to the growing body of knowledge, it is important to establish the generality of those studies across different study systems. Second, previous studies have varied widely in the number of sites sampled, the number of water samples taken at each site, and the number of replicates analysed for each water sample. It remains unclear whether the largely arbitrary nature of most eDNA sampling designs results in sufficient sample sizes to detect whether there is a difference between single- and multi-species detection methods – something that can be resolved via simulation or targeted sampling in closed systems of known community composition. Finally, previous comparative studies have made different methodological assumptions, used different response variables, and/or used different methods of analysis. A prime example is the use of detection thresholds (Peixoto *et al.* 2021). Thresholds have been used in single-species detection methods to classify a water sample as positive or negative based on the number of positive qPCR replicates from a water sample (Peixoto *et al.* 2021), as well as in metabarcoding studies to filter out low abundance reads that represent likely false positives (Deiner *et al.* 2017). Directly comparing the outcomes of studies that use different detection thresholds requires analysing single- and multi-species datasets in a consistent manner. Thus, while several studies have reached similar conclusions regarding the sensitivity of single- and multi-species detection methods, use of a single methodological framework in which to compare different studies with different data structures will enable a more robust consensus.

Hierarchical modelling is increasingly being applied to environmental DNA sampling (Fukaya *et al.* 2022) and provides an ideal method for analysing the sensitivity of single- vs multi-species eDNA methods. This approach has been used to compare eDNA detection methods to traditional sampling methods (McColl-Gausden *et al.* 2021), to analyse community data (Mcclenaghan *et al.* 2020), and to optimise study design efficiency (Fukaya *et al.* 2022), but has not yet been applied to the comparison of single- and multi-species detection methods. Hierarchical site occupancy-detection models allow for imperfect detection to be considered at

the different levels of eDNA detection (sample, technical replicate). Importantly, this approach can also incorporate detection covariates, enabling an assessment of the relative sensitivity of different detection methods. Hierarchical site occupancy-detection models therefore provide a consistent framework in which to compare the results of different studies with different data structures.

Here we compare the detection probabilities of single-species qPCR and eDNA metabarcoding, drawing on two published case studies (fish: Bylemans *et al*. 2019; amphibian: Harper *et al*. 2018) and two new datasets on a freshwater mammal (the platypus *Ornithorhynchus anatinus*) collected in south-eastern Australia. We use hierarchical site occupancy-detection models for this purpose, providing a consistent framework of analysis for drawing inferences across multiple species from different taxonomic groups and environments.

2.3 Methods

Study sites

Two new datasets on the occurrence of platypus DNA in water samples were used in this study. The first, here called the *SE Australia* dataset (Figure 2.1), includes temperate regions across different vegetation types and land uses in south-eastern Australia. Land use varied considerably across these sites, incorporating forested, agricultural, and urban environments. Stream catchment size, velocity and turbidity also varied considerably across these sites. Annual rainfall varied from 600 mm to 1500 mm across the study area. Mean maximum temperatures ranged from 12 °C to 21 °C, whereas mean minimum temperatures ranged from 3 °C to 9 °C (Bureau of Meteorology, 2021).

The second dataset, here referred to as the *Melbourne* dataset (Figure 2.1), focuses on samples collected from around Melbourne, Victoria, as part of Melbourne Water's eDNA 'Aquablitz' and long-term urban platypus monitoring programs. Stream types were varied across a range of land uses including urban, peri-urban, agricultural, and forested. Melbourne is a temperate city, with an average rainfall of 534.9 mm, an average annual maximum temperature of 19.9 °C, and an average annual minimum of 9.6 °C (Bureau of Meteorology, 2021).

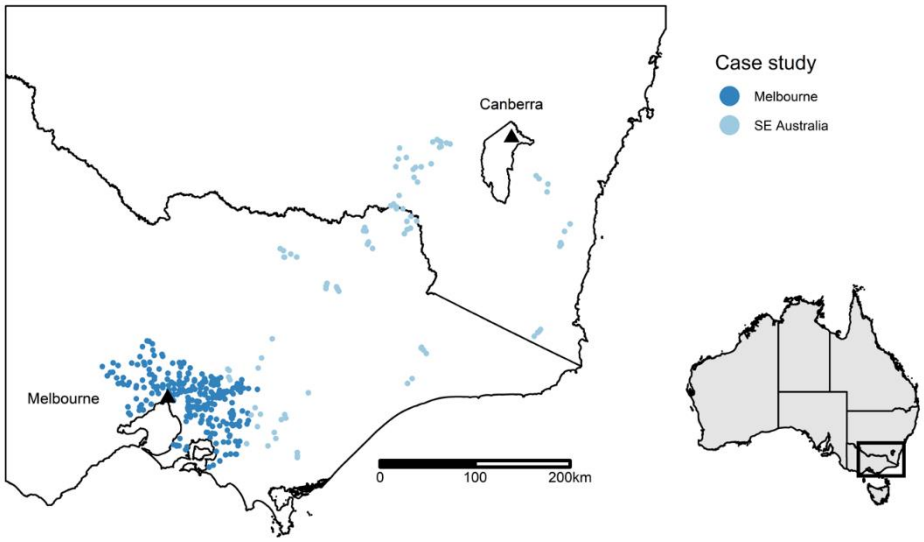

Figure 2.1*: Map of study sites for both case studies. Melbourne dataset is dark blue dots, and SE Australia is light blue dots*

eDNA sampling

Sampling for the *SE Australia* dataset (*n* = 82 sites) occurred between August 2018 and November 2019. No sampling occurred during the summer months of November to February to ensure that sampling was focused when platypus activity is higher in Victoria (e.g., excluding their key breeding period when females are predominately in burrows). At each site, two water samples (filters) were taken using two disposable sterile syringes (60-mL Luer Lock syringes; Hapool medical Technology, Shandong, China). A maximum of 500 ml of water was filtered per water sample by taking multiple 60-mL samples of water at a time and pushing them through the filter. Filtering ceased prior to 500-mL if the filter became clogged. If the water was not accessible by hand, an extendable pole with a sterilised container on the end was used to collect water. The water was then passed through an a detachable 0.22 µm filter by hand (GP 22µm Filter Unit; Sterivex, EMD Millipore corporation, Billerica MA, United States). Sterile syringes and filters, as well as disposable gloves were used, and field equipment was sterilized with 10% bleach solution and dried between sites. Samples were refrigerated for a maximum of three days until frozen, or frozen on site until DNA could be extracted. We ran three qPCRs on each water sample for the targeted, single-species approach, and metabarcoding analysis was conducted on each water sample (two per site) (see Appendix B for details of eDNA analysis).

Samples for the *Melbourne* dataset (*n* = 262 sites) were collected from July-August 2016. Three water samples (filters) were collected at each sampling location. The same methods of sample collection and storage were used as for the *SE Australia* dataset. We conducted three qPCRs on each water sample for the targeted single-species approach, and metabarcoding analysis was conducted on the pooled water samples (one per site).

Threshold scenarios

To explore the impact of applying thresholds to qPCR and metabarcoding data on the relative sensitivity of each method, we included several threshold scenarios. For metabarcoding, we filtered reads at a level that cleared false positive detections (0.0023% for platypuses) sequenced in mock community data (see Appendix B) in the *SE Australia* dataset. We also included a no threshold scenario for the *SE Australia* dataset and a pooled scenario where data from both water samples was combined. We did not include a threshold scenario for the *Melbourne* dataset as no platypus DNA was detected in any controls. Analysis of qPCR data for both datasets included three different scenarios. The first included all qPCR replicates in the model (e.g., for a sample with 3 technical qPCR replicates, the detection history could be detection = 0, 1, 1 (absent, present, present) and all three replicates would be included). For the second scenario, we collapsed the qPCR data so that if a species was detected in any of the three technical replicates, that water sample was labelled as positive (i.e., detection = 1). The third (more stringent) scenario involved collapsing the data so that samples in which a species was detected in two or more of the three technical replicates were labelled as positive (i.e., detection = 1). Threshold scenarios are outlined in Table 2.1.

Additional case studies

Additional case studies were collated from the published literature. Two publicly available datasets that compared a single-species eDNA detection method and a multi-species detection method were used. Bylemans *et al*. (2019) considered the Redfin perch (*Perca fluviatilis)* in streams where eight water samples were taken from eight sites. Three to six qPCR replicates were completed per water sample and there was one metabarcoding sample per water sample. A low abundance read filtering threshold was used for metabarcoding and a threshold of 1/3 qPCRs was used for the single-species method. Harper *et al*. (2018) investigated the great crested newt (*Triturus cristatus*) in ponds where one water sample was taken at 532 sites. Twelve qPCR replicates were conducted per water sample and there was one metabarcoding

sample per site. Two metabarcoding threshold scenarios were used (no threshold and 0.028% —a species specific threshold calculated by ensuring controls were cleared of any target species reads— as well as two qPCR threshold scenarios (1/12 replicates and 4/12 replicates). The studies are outlined in more detail in Table 2.1 (Harper *et al.* 2018; Bylemans *et al.* 2019).

Volume of extracted DNA

The volume of DNA used for qPCR and metabarcoding components differed for all datasets. For the *SE Australia* and *Melbourne* datasets, 6 µl of extracted DNA was used for qPCR per sample (2 µl per qPCR replicate) and 8µl for metabarcoding per sample. Bylemans *et al*. (2019) used 24-48 µl of extracted DNA per sample for qPCR (8 µl per qPCR replicate, with 3-6 replicates/sample) and 4 µl for metabarcoding. Harper *et al*. (2018) used 36 µl of extracted DNA per sample (3 µl per replicate) for qPCR and a total of 6 µl for metabarcoding (2 µl for 3 technical replicates)

***Table 2.1**:Summary of case studies from published datasets and new platypus data*

Case Study	Target species	Environ-ment	Sites	Replicate water samples per site	qPCR replicates per water sample	Metabarcoding replicates per site	Scenarios used in original study	Scenarios used in this study
Bylemans dataset: (Bylemans *et al.* 2019)	Redfin perch (*Perca fluviatilis*)	River/ stream	8	8	3-6	8 (1 per water sample)	Metabarcoding • Autumn 2015 low abundance read filtering threshold • Spring 2015 low abundance read filtering threshold qPCR • 1/3 replicate threshold plus replicates were only considered valid when amplification was observed in at least one assay of generic fish positive control	Metabarcoding • Autumn 2015 low abundance read filtering threshold • Spring 2015 low abundance read filtering threshold qPCR • All replicates included (3 level model) • 1/(3-6) replicate threshold (2 level model) • 2/(3-6) replicate threshold (2 level model)
Harper dataset: (Harper *et al.* 2018)	Great crested newt (*Triturus cristatus*)	Pond	532	1	12	1	Metabarcoding • Threshold (0.028%) • No threshold qPCR • 1/12 replicate threshold • 4/12 replicate threshold	Metabarcoding • Threshold (0.028%) • No threshold qPCR • All replicates included (3 level model) • 1/12 replicate threshold (2 level model) • 4/12 replicate threshold (2 level model)
SE Australia dataset	Platypus (*Ornithor-hynchus anatinus*)	River/ stream	82	2	3	2 (1 per water sample)	NA	Metabarcoding • Threshold (0.0023%) • No threshold • Pooled data across both water samples qPCR • All replicates included (3 level model) • 1/3 replicate threshold (2 level model) • 2/3 replicate threshold (2 level model)
Melbourne dataset	Platypus (*Ornithor-hynchus anatinus*)	River/ stream	262	3	3	1	NA	Metabarcoding • No threshold qPCR • All replicates included (3 level model) • 1/3 replicate threshold (2 level model) • 2/3 replicate threshold (2 level model)

Model description

The observed data collected for this study was binary detection/non-detection observations of a species at site *i* (1,2…*n*). Two model variations were used in this study: a three-level model in which no qPCR thresholds were used (i.e. three levels of data, site, sample and replicate), and a two-level model in which qPCR thresholds were used (i.e. two levels of data, site and sample) (Schmidt *et al.* 2013). Using the *SE Australia* dataset and the three-level model as an example, two independent water samples *j* (1,2) were taken at each site. Four replicate runs *k* (1,2…4), were conducted on each water sample at each site (3 qPCR replicates, and 1 metabarcoding replicate per water sample). Our model is defined by the following equations:

Equation 1. $$z_i \sim Bernoulli(\psi_i)$$

Equation 2. $$a_{ij}|z_i \sim Bernoulli\,(z_i\theta_{ij})$$

Equation 3. $$y_{ijk}|a_{ij} \sim Bernoulli(a_{ij}\omega_{ijk})$$

Equation 4. $$logit(\omega) = \alpha + \beta M_{ijk}$$

Equation 1 describes the occupancy of the species at site *i* ($z_i = 1$ where the species is present, and $z_i = 0$ where the species is absent), given the mean probability of occurrence for that species for each site (ψ_i). Equation 2 describes the presence of the species' DNA in water sample *j* (a_{ij}), conditional on species presence at site *i*, as a function of eDNA availability probability θ (the mean probability that a given water sample contains the species' DNA). Equation 3 describes the observed detection process at site *i* for water sample *j* in replicate *k*, conditional on the probability of detecting the species' DNA (ω), given it is available to be detected (a_{ij}). Finally, Equation 4 shows that the probability of detecting the species' DNA (ω) is a function of an overall intercept α and a coefficient describing the effect of detection method βM_{ijk}, where $M_{ijk} = 0$ for qPCR, and $M_{ijk} = 1$ for metabarcoding. A positive parameter estimate (β) therefore indicates that the species was more likely to be detected with metabarcoding. Likewise, a negative parameter estimate (β) suggests that the species was more likely to be detected with qPCR. Prior distributions were specified on the logit scale where ψ, θ, ω and β were normally distributed with a mean of 0 and a precision of 0.1. We also used the two-level model described in Appendix B.

Model fitting

Models were fit in R version 3.6.3 (R Development Core Team, 2020) using the software JAGS through the package rjags v.4-8. Three model chains were run for 30,000 iterations each. The first 10,000 samples were treated as the burn-in and discarded, and the remaining samples were thinned by a factor of 10. This resulted in 2,000 samples per chain from the posterior distribution. Convergence was measured using traceplots from the jagsUI package v.1.5.0; chains were well mixed. Gelman-Rubin statistic values $\hat{R}$ were below 1.1, suggesting successful convergence of chains. Example JAGS code is presented in Appendix B.

Simulation analysis

To investigate whether the four case studies used here have the study design structure (e.g., site, sample, and qPCR replication) necessary to accurately estimate the parameter of interest (i.e., the effect of eDNA detection method β), we conducted model simulations. For details and results of model simulations, see Appendix B.

2.4 Results

Platypus eDNA was detected in at least 1/6 qPCR replicates (across two water samples with three technical replicates per sample) at 69 of the 82 sites sampled in the *SE Australia* dataset. For most of these detections (61/69), at least 2/3 qPCR technical replicates detected platypus eDNA in at least one water sample. Metabarcoding detected platypus DNA at 46 sites (representing 0.45% of all reads). At 26 of these 46 sites, platypus DNA was detected in both replicate water samples. Where metabarcoding detected platypus DNA, the qPCR assay also detected platypus DNA – the average number of positive qPCRs at each of these sites was 5.2/6 and the most common value was 6/6 qPCRs.

For the *Melbourne* dataset, qPCR detected platypus DNA at 126/262 sites (≥1/9 qPCRs across three water samples each with three technical replicates). Metabarcoding detected platypus DNA at 47 of the 262 sites (representing 0.72% of all reads). Where metabarcoding detected platypus DNA, qPCR also detected platypus DNA, with an average of 2.1/9 positive qPCRs per site.

Sensitivity of single- and multi-species detection methods

For the SE Australia dataset, qPCR was more sensitive at detecting platypus DNA compared to metabarcoding across all thresholding scenarios considered. When no metabarcoding threshold was applied, and all replicate qPCRs were included (e.g., the whole detection history), the effect of detection method on detection probability (β) was negative and 95% credible intervals (CI) did not overlap zero, suggesting that platypus DNA is more likely to be detected using qPCR compared to metabarcoding (mean [95% CI] = −1.068 [−1.544, −0.609], Figure 2.2). As expected, applying a threshold to metabarcoding data, but not qPCR, led to a stronger effect of detection method (mean β [95% CI] = −2.176 [−2.639, −1.720], Figure 2.2). Pooling metabarcoding data for each site (combining the two water samples) and not applying a qPCR threshold reduced the effect of detection method on detection probability but qPCR was still the more sensitive method (mean [95% CI] = −0.740 [−1.317, −0.156). When 1/3 or 2/3 positive qPCR replicates were required to label a water sample as positive (and no metabarcoding threshold was applied) the β parameter estimate remained negative and 95% CIs did not overlap zero (1/3 qPCRs: mean [95% CI] = −2.309 [−6.311, −1.573]. 2/3 qPCRs: mean [95% CI] = −1.158 [−1.801, −0.542], Figure 2.2). Thus, for the SE Australia dataset, applying thresholds to qPCR and metabarcoding data, or pooling metabarcoding samples, did not influence our conclusion that qPCR was more sensitive than metabarcoding. See Table B.1 in Appendix B for detailed detection, availability, and occupancy probabilities for all scenarios.

For the Melbourne dataset, no threshold was applied to metabarcoding data. When no threshold was applied to qPCR data (e.g., the whole detection history is included), the effect of detection method on detection probability was negative but uncertain (mean β [95% CI] = −0.298 [−0.729, 0.120]). When 1/3 positive qPCR replicates were required to label a water sample as positive, the effect of detection method was negative and 95% CIs did not overlap zero (mean [95% CI] = −1.322 [−1.745, −0.906]). Using a more stringent threshold of 2/3 positive qPCRs, the direction of the relationship was much more uncertain (mean [95% CI] = 0.061 [−0.419, 0.554]; Figure 2.2). Overall, then, qPCR was more sensitive for the Melbourne dataset, but effects of detection method were more dependent on qPCR thresholds than in the SE Australia dataset.

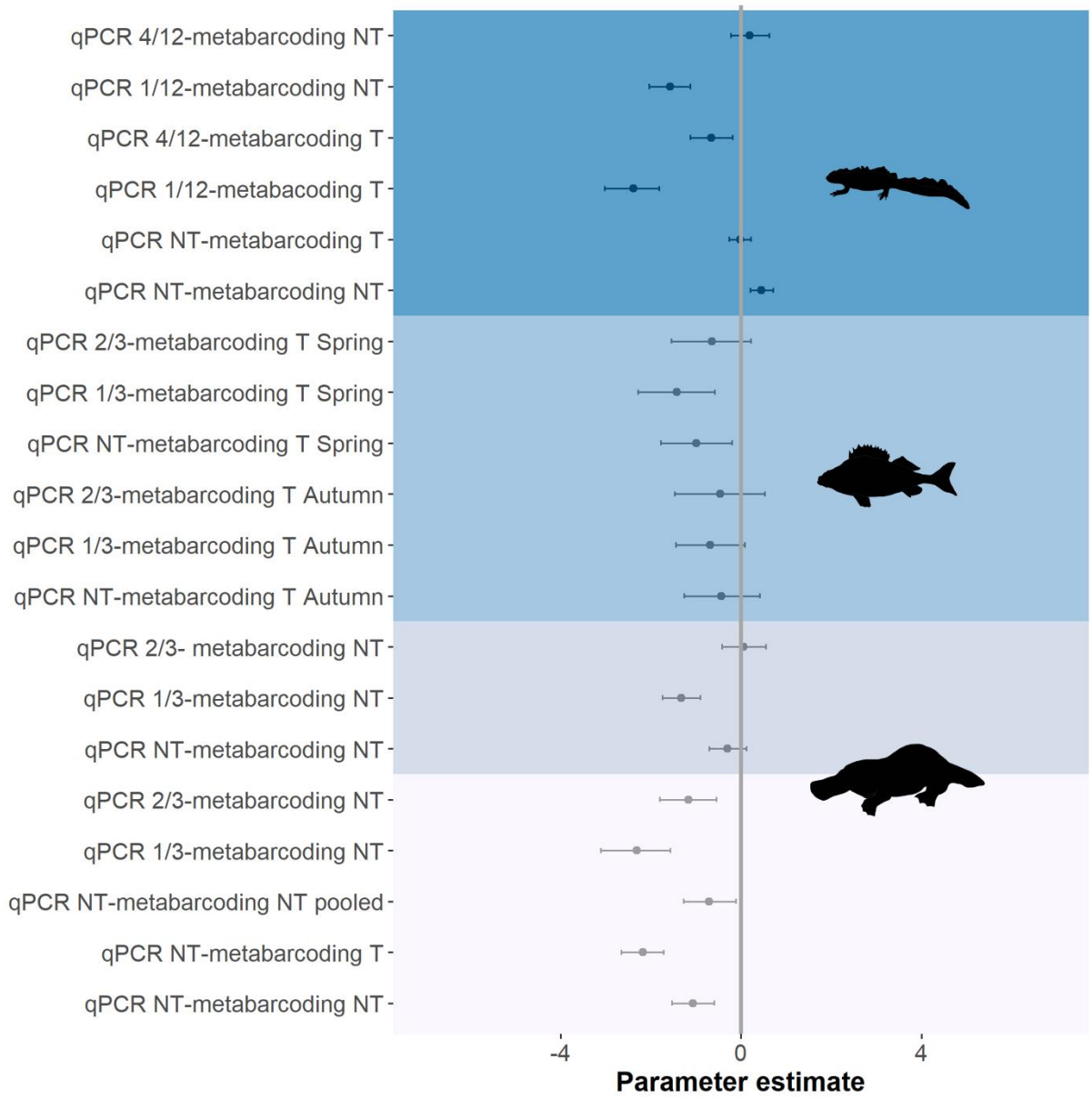

Figure 2.2*:* β parameter estimates (effect of detection method on probability of detection) for each dataset and scenario. Dots represent posterior means and lines represent 95% credible intervals. Datasets are divided by colour block. From bottom to top, datasets are SE Australia, Melbourne, Bylemans et al., (2019) (Autumn, Spring), Harper et al., (2018). NT = No threshold applied, T = Threshold applied, 1/3 = 1 positive qPCR required to class a sample as positive, 2/3 = 2 positive qPCRs required to class a sample as positive, 1/12 = 1 positive qPCR required to class a sample as positive and 4/12 = 4 positive qPCRs required to class a sample as positive.

Additional case studies

We investigated two seasons of data from Bylemans *et al*. (2019). Low abundance read filtering was used on this dataset, therefore all scenarios use this metabarcoding data. When all qPCR

replicates were included using data collected in Spring, the analysis revealed that qPCR was more sensitive than metabarcoding (mean β [95% CI] = −0.983 [−1.788, −0.193], Figure 2.2) at detecting Redfin perch (*Perca fluviatilis*). Applying a 1/3 qPCR threshold resulted in a certain effect of detection method mean β [95% CI] = −1.428[−2.283, −0.587] whilst a 2/3 qPCR threshold overlapped zero and was thus uncertain: mean β [95% CI] = −0.648[−1.551, 0.215], Figure 2.2). The autumn season data, based on the same sites, produced similar results, although effects of detection method were uncertain when no qPCR threshold was applied (mean β [95% CI] = −0.428 [−1.249, 0.439], Figure 2.2). When a threshold of 1/3 qPCRs was applied, β was negative and 95% CIs only slightly overlapped zero (mean β [95% CI] = −0.689 [−1.449, 0.080]). However, when the more stringent threshold of 2/3 qPCRs was applied, the effect of detection method on detection probability was uncertain (mean β [95% CI] = −0.455[−1.471, 0.526], Figure 2.2).

We also analysed data obtained from Harper *et al*. (2018) in six ways: two metabarcoding scenarios, with three qPCR variations. When a metabarcoding threshold was not applied, and all qPCR replicates were included, the great crested newt (*Triturus cristatus*) was more likely to be detected with metabarcoding than with qPCR (mean β [95% CI] = 0.456 [0.203, 0.707], Figure 2.2). Using no metabarcoding threshold, but with a qPCR threshold of 1/12, the effect of detection method was certain and negative (1/12 qPCRs: mean β [95% CI] = −1.566 [−2.040, −1.124]); however, the estimate became uncertain when a 4/12 qPCR threshold was used (4/12 qPCRs: mean β [95% CI] = 0.193 [−0.231, 0.618]). When a species-specific metabarcoding threshold was applied to reduce the occurrence of potential contamination, and all qPCR replicates were used, the effect of detection method became uncertain and was centred around 0 (mean β [95% CI] = −0.022 [−0.271, 0.228], Figure 2.2). Using a metabarcoding threshold and a qPCR threshold of 1/12 or 4/12 positive technical replicates, the β parameter estimate was negative and 95% CIs did not overlap zero (1/12 qPCRs: mean β [95% CI] = −2.384 [−3.023, −1.819]; 4/12 qPCRs: mean β [95% CI] = −0.653 [−1.130, −0.193], Figure 2.2).

2.5 Discussion

Our results illustrate that in most circumstances, single-species eDNA detection methods are more sensitive at detecting target species than multi-species detection methods, although method decisions —particularly qPCR or metabarcoding thresholds— affect the extent of this

disparity. Importantly, these findings were generally robust to different species and study systems.

Site occupancy-detection models

Our study adds to a growing body of literature illustrating the benefits of using site occupancy-detection models to analyse eDNA data. Notably, the hierarchical nature of these models allows for imperfect detection to be incorporated at the different sampling levels inherent in eDNA data (water sample - availability probability, and replicate - detection probability) (Schmidt *et al.* 2013), enabling users to compare the sensitivity of different sampling methods. Using these models, we have compared the sensitivity of single species eDNA detection methods to eDNA metabarcoding and investigated the impact of applying detection thresholds on the sensitivity of both methods, across a range of freshwater study systems in a consistent manner that allows for study design variation to be considered (e.g., different numbers of sites, water samples, technical replicates). Using a single modelling framework to analyse different case studies has allowed us to conduct a more comprehensive and direct comparison between eDNA detection methods than has been previously possible.

Detection thresholds

Applying a threshold to single-species eDNA detection methods, such as qPCR, is common practice (Bohmann *et al.* 2021; Langlois *et al.* 2021). Thresholds can reduce the probability of false positive detections; however, they can also omit rare species or species with low abundance (false negatives) (Ficetola *et al.* 2016; Lahoz-Monfort *et al.* 2016; Deiner *et al.* 2017). In our case studies, applying a low threshold to a single-species detection method (such as 1/3 positive qPCR replicates needed to classify a water sample as positive) resulted in a stronger estimated effect of eDNA detection method (i.e., more negative β estimate) compared to when no qPCR threshold was applied. This is presumably because using a qPCR threshold reduces the number of zeros in a detection history, misrepresenting the uncertainty of the detection process (e.g., a 001 detection history, in which a species was detected in 1/3 qPCR replicates of a water sample, would be assigned a 1). Therefore, an effect may appear to be more extreme —more negative or positive— when a threshold is used compared to the full dataset. This effect can be seen in the *SE Australia, Melbourne,* and *Harper* datasets (Figure 2.2). Using a higher threshold of 2/3 qPCRs resulted in qPCR becoming less sensitive (i.e., a less negative β parameter estimate). In both circumstances (i.e., high and low qPCR

thresholds), using a threshold also resulted in more uncertain estimates (wider 95% credible intervals), as there was less data to estimate the parameters of interest. Replication at multiple levels may be required to accurately estimate parameters in hierarchical models (Guillera-Arroita *et al.* 2017). Thus, reduction in replication at the technical replicate level when qPCR thresholds are used may have a strong impact on parameter estimation.

The effect of applying a threshold to qPCR data was seen most strongly in the *Harper* dataset, which serves as an illuminating case study. When we added the two scenarios in which no threshold was applied to qPCR data, the β parameter estimate became uncertain (95% CIs overlapped zero) or was positive (suggesting metabarcoding was more sensitive). This outcome was a result of the large amount of replication in the single-species detection method (n = 12 qPCR replicates) compared to the metabarcoding method (n = 1 replicate). In this case, applying a threshold to the single-species detection method dramatically reduced the amount of replication, resulting in a single data point instead of 12. While this is an extreme example (12 qPCR replicates vs 1 metabarcoding replicate), the mismatch between the replication for qPCR compared to metabarcoding created a scenario in which the non-detections (zeros) from the large number of qPCR replicates resulted in reduced estimated sensitivity of the single-species detection method. Processing all water samples separately for metabarcoding (not pooling samples) would assist in evening the replication across the two methods and potentially provide a more robust comparison of the eDNA detection methods.

Detection thresholds are also often used in eDNA metabarcoding studies. False positive detections can occur using metabarcoding because of factors such as primer specificity (Deiner *et al.* 2017), primer bias (Elbrecht and Leese 2015), index switching (Esling *et al.* 2015), or contamination (Mcelroy *et al.* 2020). Metabarcoding data can be analysed in at least two ways. First, multiple water samples taken at a site can be pooled into one sample (Deiner *et al.* 2017), reducing replication at the water sample level. Second, metabarcoding data can be analysed separately for each water sample, increasing the amount of replication present in the data. In either case, low abundance read filtering (i.e., removing reads that comprise a small percentage of the total number of reads obtained from a sample) can be used, along with other data preparation steps, to reduce potential false positive detections in eDNA metabarcoding data (Deiner *et al.* 2017). Some common methods include removing a percentage of reads to clear any contamination in negative controls in either a species-specific way (Harper *et al.* 2018), or across a whole study (Olds *et al.* 2016). Thresholds can also be applied to clear mock

communities (Harper *et al.* 2018). Regardless of the precise method used, the current study shows that applying a threshold can affect the perceived sensitivity of eDNA metabarcoding, and that the magnitude of this effect can vary among studies. The general result from the *SE Australia* and *Harper* datasets (which both included data with and without a metabarcoding threshold) was that applying thresholds resulted in qPCR being more sensitive at detecting the target species (i.e., more negative β parameter estimates), compared to when no metabarcoding threshold was applied (Figure 2.2). Applying thresholds altered the direction of our primary finding (i.e., that qPCR is more sensitive compared to metabarcoding) for one study (the *Harper* dataset) but not others considered here.

We investigated the effects of using qPCR and metabarcoding thresholds, as they are common in the literature (including both published case studies used here), but we caution against their use in general, given the availability of models to explicitly account for the hierarchical nature of the eDNA detection process (Lahoz-Monfort *et al.* 2016) and the potential for false positives (Guillera-Arroita *et al.* 2017). Nonetheless, running analyses with and without thresholds enabled us to demonstrate the fact that our overall conclusion – that qPCR was more sensitive than metabarcoding – was relatively robust to this method decision (also see McColl-Gausden *et al.* (2021)).

Study design

We were unable to control for the volume of extracted DNA used in each method for each dataset. The Bylemans *et al*. (2019) and Harper *et al*. (2018) datasets used a larger volume for the single-species targeted approach compared to metabarcoding, whereas the *SE Australia* and *Melbourne* datasets used a larger volume of extracted DNA for metabarcoding compared to the singe-species approach. Bylemans *et al*. (2019) evaluated if the differing volume of extracted DNA impacted results using a subset of three sites where either 4 or 8 µl of template DNA was used for each metabarcoding technical replicate. They found that the volume of extracted DNA used did not significantly affect detection probabilities. The results of especially the *SE Australia* and *Melbourne* datasets support the assertion that qPCR is more sensitive than metabarcoding, as a larger volume of extracted DNA was used for metabarcoding, the less sensitive method. This study demonstrates the large variation in study design and methods used in eDNA detection studies and highlights the need for more standardised methods (Bohmann *et al.* 2021).

Previous studies have shown that primer selection can influence the likelihood of eDNA detection (Wood *et al.* 2019). While we were able to compare single-species detection methods to eDNA metabarcoding across multiple study systems, we were unable to evaluate whether the choice of metabarcoding primer influenced our findings. We used a universal vertebrate primer (Valentini *et al.* 2016), as opposed to a mammal-specific primer, for our two new case studies because of the low number of aquatic mammals present in the study area. This allowed us to detect a total of 167 species in the *SE Australia* dataset and 146 in the *Melbourne* dataset for the purpose of biodiversity screening across the waterways. A more group-specific metabarcoding primer could be used for other taxonomic groups – for example, the "telostei" primer for fish (Riaz *et al*. 2011; Shehzasd *et al*. 2012) or the "batra" primer for amphibians (Valentini *et al.* 2016). Different metabarcoding primers can preferentially sequence different species (Bylemans *et al.* 2019), due to the number of matches between a species' DNA and the primer (Piñol *et al.* 2019), or the concentration of DNA from a particular species in a sample. For example, some species that occur at high abundance can mask rarer species in a sample (Evans *et al.* 2016) but inherently the more specific you make a metabarcoding primer, the less information you will get from the metabarcoding method. It is therefore important to consider these factors when selecting a metabarcoding primer. Future studies could usefully compare multiple metabarcoding primers with varying taxonomic resolution to single-species detection methods to assess the generality of our findings.

An additional factor that is important to consider when comparing single-species detection methods to eDNA metabarcoding, and one that has not been evaluated previously, is whether the amount of site and sample replication used is sufficient to detect a difference between sampling methods. There are evident differences in the data structure of the four datasets used in this study, and this variability in study designs is generally representative of the eDNA literature to date. Accordingly, standardisation of eDNA sampling methods is a topic of increasing interest (Loeza-Quintana *et al.* 2020; Bohmann *et al.* 2021). Standardisation is important not only for quality assurance, but also because it can be difficult to compare or generalise the results of individual studies when they have different data structures (such as the number of sites, samples, and replicates) and different sampling protocols.

Finally, our results support the use of a hybrid sampling approach involving single- and multi-species methods for projects where the detection of a target species is important, but community data may also be useful. In both our platypus datasets (but particularly in the *SE*

Australia dataset), numbers of qPCR detections were high when metabarcoding detected platypus eDNA in either of the two water samples taken, suggesting that metabarcoding did not produce substantial numbers of false positive detections. In addition, results of site occupancy-detection models revealed that metabarcoding is generally more susceptible to false negative detections compared to qPCR. Taken together, these results indicate that metabarcoding could be conducted at all sites for a given survey, and a single-species detection method could be used only at sites that are negative for the target species, especially within its previously known range. Such a hybrid approach reduces the number of sites/samples that need to be processed using both eDNA detection methods. This outcome may assist managers with large-scale project planning and reduce monitoring costs, as both methods don't need to be conducted at every site, and DNA extraction has already taken place. Nonetheless, the decision to use a single- or multi-species detection method still needs to align with the aims of the study. If detection data is needed for target species (e.g., an invasive or threatened species) and false negative detections are detrimental, then the higher sensitivity of a single-species detection method may be warranted. However, if community data is also needed, a hybrid approach, as discussed above, might be a valid option as we have demonstrated for platypuses.

2.6 Conclusions

This study presents a consistent method to compare single-species and multi-species eDNA detection methods across different study designs, and for different aquatic taxa. The consistent result across these case studies is that single-species detection methods can be more sensitive compared to metabarcoding but that the sensitivity of each method will depend somewhat on detection thresholds and some aspects of the study design. Ultimately, the decision to use either a metabarcoding or qPCR approach is a choice between the need for multi-species information (e.g., biodiversity assessments) or individual species information (e.g., threatened or invasive species), where there is a trade-off between detection sensitivity in the former and cost with the latter. It is possible that exploration of finer taxonomic resolution metabarcoding primers (e.g., fish, amphibians), could demonstrate a compromise between the more extreme taxonomic comparison by this study.

Chapter 3
Using eDNA sampling to identify correlates of platypus occupancy across broad spatial scales

3.1 Abstract

Species presence-absence data can be time-consuming and logistically difficult to obtain across large spatial extents. Yet these data are important for ensuring changes in species distributions are accurately monitored and are vital to our understanding of how species respond to global change. Here we demonstrate how environmental DNA (eDNA) sampling can be used to collect species occupancy data rapidly and efficiently across vast spatial domains. We use a widely distributed species, the platypus (*Ornithorhynchus anatinus*), as a test case. We used an environmentally stratified survey to assess the presence-absence of platypus eDNA at 504 sites across 584,292 km2 of south-eastern Australia, representing ~37% of the species' extensive distribution. Site occupancy-detection models were used to analyse how landscape- and site-level factors affect platypus occupancy, enabling us to incorporate uncertainty at the different levels inherent in eDNA sampling (site, water sample replicate, and qPCR replicate). Platypus eDNA was detected at 272 sites and were more likely to occupy sites in catchments with increased runoff and less zero-flow days, and sites with access to banks suitable for burrowing. Platypuses were less likely to occupy sites in catchments with a high proportion of shrubs and grasslands, or agricultural land use. The broad spatial extent of this study makes it the most spatially comprehensive eDNA program undertaken to date using a systematic study design. Our case study shows that such surveys provide an efficient means to understand how environmental characteristics affect species occupancy across broad environmental gradients. The methods employed here can be applied to aquatic and semi-aquatic species globally, providing unprecedented opportunities to understand biodiversity status and change.

3.2 Introduction

An ability to detect where species occur over broad spatial extents is critical to wildlife monitoring (Guisan *et al.* 2013). Without these data, range shifts, contractions, expansions, or invasions can go undetected. Data collected across large spatial extents are also important for informing extinction risk assessments, such as regional or global conservation listings (e.g., the International Union for Conservation of Nature (IUCN) Red List of Threatened Species), and for understanding and predicting the impacts of environmental change on species distributions. Citizen science programs that involve recording wildlife sightings can assist in the collation of large amounts of data over broad spatial domains, but issues such as spatial and temporal sampling biases, uncertain sampling effort, and taxonomic misidentifications, mean that such programs are not ideal for many conservation applications (Tiago *et al.* 2017).

Accurate and efficient detection methods are necessary to ensure wildlife monitoring can feasibly occur over large areas in a systematic way. Many traditional methods, such as trapping, can be time-consuming and typically require multiple sampling events to be confident that a species is absent from a site (Lugg *et al.* 2017). This is important, as false negative detections —failing to detect a species that is present at a site (Guillera-Arroita *et al.* 2010b)— can be detrimental to estimates of site occupancy, and bias inferences regarding the impacts of threatening processes on species distributions (Lahoz-Monfort *et al.* 2014) . Environmental DNA (eDNA) sampling —the detection of intra- or extra-cellular DNA in environmental samples— has been shown to result in high detection probabilities for many species (Wilcox *et al.* 2016; Lugg *et al.* 2017). Environmental DNA sampling not only achieves high detection probabilities but can be extremely efficient. Sampling eDNA at a site can be very quick —less than 10 minutes depending on the water sampling method used. Thus, travel time between sites becomes the limiting factor for large-scale eDNA studies. Recent studies have shown the scalability of this sampling technique. For example, Biggs *et al.* (2015) sampled 75% of the great crested newts' (*Triturus cristatus*) 217,000 km^2 range using a citizen science program; Pont *et al.* (2018) sampled fish along 524 km of river; and Fukumoto *et al.* (2015) sampled for giant salamanders (*Andrias japonicus*) over an area of 1,000 km^2. In marine environments, even larger areas can typically be sampled because of the ease of movement of surveyors. For example, West *et al.* (2021) surveyed bony fish, elasmobranchs and aquatic reptiles at 284 sites along 700 km of coastline in northwest Australia; Closek *et al.* (2019) sampled marine vertebrates within five geographic regions along a 539 km route on the west coast of the USA;

and Fraija-Fernández *et al.* (2020) collected samples for fish across an area of 120,000 km^2. Environmental DNA sampling therefore has the potential to revolutionise large scale (>1,000 km^2) studies of species distributions.

Site occupancy-detection models (SODM) are an ideal method with which to analyse eDNA data and have been used previously in eDNA studies to great effect (Doi *et al.* 2019; Mcclenaghan *et al.* 2020; Bush *et al.* 2020). These models allow for imperfect detection to be incorporated at the different levels of the inherently multi-staged process of eDNA sampling (site, water sample, technical replicate). SODM have been used in single-species (e.g., qPCR) eDNA studies (Schmidt *et al.* 2013; Hunter *et al.* 2018) as well as multi-species (eDNA metabarcoding) studies (Doi *et al.* 2019; Bush *et al.* 2020; Mcclenaghan *et al.* 2020). For example, SODM have been used to compare eDNA sampling to traditional methods (McColl-Gausden *et al.* 2021) and to optimise study design efficiency (Fukaya *et al.* 2022).

An additional benefit of using SODM with eDNA data —one that has received little attention in the eDNA literature to date— is that effects of covariates can be estimated at the site, sample, or technical replicate level. Including covariates in SODM enables researchers to estimate the impacts of landscape- or site-level factors on occupancy or detection. For example, Mcclenaghan *et al.* (2020) used SODM to estimate the impact of water depth at the site level, and sequencing depth at the technical replicate level, finding that mean occupancy probability decreased with increasing water depth. Because eDNA sampling is particularly suited to rapid, large-scale surveys, it is ideal for understanding species distributions along broad environmental gradients. Environmental DNA sampling can therefore inform species' habitat requirements, as well as shed light on species responses' to global change, such as the effects of extreme droughts or fire events (Ward *et al.* 2020).

In this chapter, we demonstrate the large spatial scale over which eDNA data can be efficiently collected for a freshwater monotreme endemic to eastern Australia —the platypus (*Ornithorhynchus anatinus*). This iconic species has been recently classified as *vulnerable* in the state of Victoria, Australia and is classified as Near Threatened by IUCN (Woinarski and Burbidge 2016). However, there remains substantial uncertainty regarding the extent of the species' decline, as well as the threats that have led to its imperilment (Hawke *et al.* 2019). Platypuses are cryptic and largely nocturnal, foraging in rivers and streams, and occupying burrows in riverbanks (Serena *et al.* 1998). Incidental observations of platypuses are therefore uncommon, although targeted citizen science programs are increasing in popularity. Moreover,

targeted platypus surveys, which commonly use fyke or gill nets (Serena *et al.* 2014), are time-consuming. Traps must be checked regularly, and sometimes continuously (Serena *et al.* 2014), and thus trapping can only be deployed over small spatial scales without extensive coordinated effort. In addition, repeated trapping surveys (≥5) are needed to achieve a high probability of platypus detection (Lugg *et al.* 2017). These factors are compounded by the fact that platypuses occur across a very large area of eastern Australia (Woinarski and Burbidge 2016), making it difficult to monitor changes in the species' distribution and status. A detection method is therefore needed that can be used to quickly and accurately monitor the species' distribution and investigate the impacts of threatening processes. Previous research has shown that eDNA sampling is more sensitive and cost-efficient than trapping for detecting platypuses (Lugg *et al.* 2017), and thus may be suitable for this purpose.

We use eDNA surveys, based on a stratified sampling design, across 584,292 km^2 of south-eastern Australia —encompassing ~37 % of the estimated platypus distribution— to determine how landscape- and site-level factors affect platypus occupancy. To the best of our knowledge, this work represents the largest systematic eDNA sampling program conducted to date, shedding insight into factors that impact the species distribution and providing a benchmark for evaluating their conservation status.

3.3 Methods

Study area and site stratification

This study focused on a 584,292 km^2 area of south-eastern Australia, encompassing the states of Victoria and New South Wales, and the Australian Capital Territory (Figure 3.1A) and approximately 37% of the platypuses' distribution (Figure 3.1B). This area includes temperate, alpine, and grassland regions across 12 Interim Biogeographic Regionalisation for Australia (IBRA) bioregions (see Appendix C for a list of bioregions). Rainfall, temperature, and vegetation vary significantly across the study region. Annual rainfall varies from 1500 mm in alpine areas to 200 mm in the west. Mean maximum temperature varies from 24 °C to 9 °C and mean minimum temperature varies from 9 °C to 0 °C (Bureau of Meteorology 2021).

Contracted catchments (Commonwealth of Australia 2012) were chosen as sampling units. These catchments are calculated from a nine-second digital elevation model with a contracted node —a feature in the landscape such as a significant monitoring location or the confluence

of two major rivers— at the outflow point. Only contracted catchments that were within a buffered 5 km area around major rivers (Commonwealth of Australia 2012) were considered. This subset of contracted catchments was then used to develop a random stratified design. Contracted catchments were stratified by vegetation state, defined using the 5-km national-level Vegetation Assets, States and Transitions (VAST) dataset for Australia (Lesslie *et al.* 2010). VAST classifies vegetation according to its degree of anthropogenic modification from a natural state. The classes (from least human modification to most) consist of residual bare, residual, modified, transformed (native), replaced and removed. Contracted catchments were assigned a VAST class based on the predominant class across the catchment. Catchments were then selected at random to produce an even number (where possible) of catchments within each VAST class. This resulted in the selection of 215 contracted catchments across the study area (see Appendix C, Figure C.1). Ten contracted catchments were selected in the removed (urban) class (as there were less contracted catchments classified as urban available to select); all other selected catchment classes varied between 42 and 56 selected catchments. Extensive wildfires (Ward *et al.* 2020) during sampling prevented some catchments from being visited, and thus 35 additional contracted catchments were added to the sampling program. These additional catchments were in the same stratification classes as those affected by fire. Within each of the selected catchments, the number of sites sampled increased with waterway length, such that between three and 12 sites were sampled within a catchment (see Appendix C for details). Site were selected based on accessibility within the selected contracted catchment. There was also an attempt to sample sites further than 7 km away from each other to assist with site independence from downstream transport (Fremier *et al.* 2019; Shogren *et al.* 2019).

eDNA sampling

Environmental DNA sampling took place from August 2018 to April 2021, with no sampling occurring during the summer months of November to February when females and offspring spend large amounts of time in burrows. Sampling was focused, where possible, on the peak platypus activity times of spring and autumn, with some opportunistic sampling occurring in Winter of 2020 due to sampling restrictions associated with the COVID-19 pandemic. At each site, two independent water samples were taken (Lugg *et al.* 2017). Two disposable syringes (60-mL Luer Lock sterile syringes; Hapool medical Technology, Shandong, China) were used to draw water from different microhabitats along the water's edge. A maximum of 500 ml of water was then pushed through a detachable 0.22 µm filter by hand. by collecting multiple 60-

mL samples at a time. The maximum volume was sometimes not achieved if the filter clogged (GP 22µm Filter Unit; Sterivex, EMD Millipore corporation, Billerica MA, United States). Samples were stored in a car fridge for a maximum of three days until frozen, or frozen on site until DNA could be extracted.

Environmental variables

The dataset was analysed in two different ways as a potentially influential dataset (a zero-flow index) was only available for the state of Victoria, and not across the whole sampled region. For the first analysis using all available data, at each site, several environmental variables were collected. The following variables were collected on an ordinal scale of 1-5, with 1 being very poor and 5 being excellent (see Appendix C Table C.1): bank erosion, bank vegetation (immediately adjacent [< 2 m] to water), verge vegetation (riparian zone), banks suitable for burrowing, instream complexity, and channel complexity. Stock access and human activity (e.g., litter, presence of campfires) were scored as present or absent.

We also estimated environmental attributes of each site using a variety of spatial layers. Land use was estimated using the Dynamic Land Cover dataset (Geoscience Australia, 2017); original categories were condensed into five categories (mines and urban areas, waterbodies, pasture and cropping [agriculture], grasslands and shrubs, and forested). We used this refined dataset to calculate the proportion of each contracted catchment covered by each land use category. The only exception was urban land use, which was modelled as a binary variable due to the high prevalence of zeros. Total runoff for each contracted catchment from 2018-2019 was calculated as a proxy for water availability, using runoff data from the Australian Water Resources Assessment Landscape model (Frost *et al.* 2018).

For the second analysis, we analysed a dataset consisting of 309 sites across Victoria, Australia (Appendix C, Figure C.2) to incorporate a spatial dataset, exclusive to the state of Victoria, containing an index of days of zero-flow annually, a potentially important correlate of platypus occupancy. In this zero-flow dataset, a higher index value corresponds to fewer zero-flow days annually, compared to other river systems in Victoria. The Victorian eDNA analysis incorporated water samples from all Victorian sites from the dataset described above, as well as samples from an additional 58 sites around Melbourne that were collected using the same methods as described above. Covariates in this second analysis included land use (Dynamic Land Cover dataset) and runoff across the catchment (as above), in addition to the annual zero-

flow data obtained from The Victorian Department of Environment, Land, Water and Planning. Annual zero-flow data were in point form and therefore the nearest data point to the sampling location following the river network was obtained using QGIS (QGIS Geographic Information System 2021).

DNA extraction and qPCR

DNA was extracted from Sterivex filters using Qiagen DNeasy Blood & Tissue Kits (Spin-column protocol) in a room that is dedicated to low-quantity DNA sources with qPCR setup undertaken in a sterilised laminar flow hood. Into each filter unit, 540 μL of ATL buffer and 40 μL of proteinase K was added and then each filter was sealed and incubated for 3 hours at 56 °C with constant agitation. The lysis solution was transferred into new 2 mL tubes. The manufacturers protocol was followed for the rest of the DNA extraction with some minor adjustments; 500 μL AL buffer, 500 μL ethanol, and a final elution step of 100 μL AE buffer was used for each sample. Included in every batch of DNA extractions (~24) was a negative DNA extraction control (that replicated the process using a sterile Sterivex filter).

Species-specific primers and TaqMan® probe targeting the platypus *Cytb* mitochondrial control region used in this study were from Lugg *et al.* (2017). Sequences for the primers/probe were; forward primer OAcr_F CAGCAATACCCTAGACAAGG, reverse primer OAcr_R CGCTTCAATGGCTGCGC, and MGB probe OAcr_MGB CGAACCCCATGAGTAGAAAAT. Assays were from Life Technologies (Thermo Fisher Scientific) as a custom TaqMan® gene expression assay.

Using a Roche LightCycler 480 system in a 384-well format, Real-time TaqMan® PCR assays were conducted in 10 μL reaction volumes containing: 5 μL of 2 Qiagen multiplex PCR Master Mix (Qiagen), 0.5 μL 20× TaqMan® Gene Expression Assay, 2.5 μL ddH_2O and 2 μL of DNA were prepared in triplicate. In each 384-well assay plate, control reactions containing 10, 100, 1000, 10000, 100,000 copies of *O. anatinus Cytb* synthetic oligonucleotide (Integrated DNA technologies, Baulkham Hills, NSW, Australia). A negative qPCR control with no DNA template and all DNA negative controls extracted at the same time as samples were also included. Amplification occurred under the following conditions: 15 min at 95 °C, 15 s at 95 °C and 1 min at 60 °C for 50 cycles. Crossing point (Cp) values were determined from the

amplification profiles of PCR using the Absolute Quantification module of the LightCycler® 480 software package. To test for the presence of PCR inhibitors, a TaqMan® Exogenous Internal Positive Control VIC probe was also run for each sample. If inhibition was detected, samples were diluted (1:10) and the qPCR analysis was repeated. The efficiency of all qPCR reactions was >98% and all negative controls (extraction or qPCR) were negative.

Model description

The observed data were binary detection/non-detection observations of platypus eDNA in three qPCR replicates *k* (1,2,3) per water sample *j* (1,2) at site *i* (1,2…*n*).

Equation 1. $$z_i \sim Bernoulli(\psi_i)$$

Equation 1 describes the occupancy of the species at site *i* (z_i=1 where the species is present, and z_i=0 where the species is absent), given the probability of occurrence of the species at site *i* (ψ_i).

Equation 2. $$logit(\psi_i) = \psi.intercept + \beta_n * cov_n + \gamma_i + \gamma_{catchment}$$

Equation 2 shows that probability of occurrence at site *i* (ψ_i) is a logit function of an overall intercept ($\psi.intercept$) and one or more covariates. Covariates are either in the form of ordinal sequences from one to five, continuous values between 0 and 1, or binary presence/absence (See Appendix C Table C.1). A positive β estimate for a continuous predictor variable (e.g., runoff) indicates platypuses are more likely to occupy sites with higher values of that covariate, whereas a positive β estimate for a categorical variable (e.g., presence of stock access) indicates that platypuses are more likely to occur when that variable is present. For ordinal covariates, a positive β estimate suggests that platypuses are more likely to occupy a site with the characteristics of that level compared to the baseline (very poor) level. This equation also incorporates random effects for catchment ($\gamma_{catchment}$) for the first broad-scale analysis, and a random effect for site (γ_i) for the Victorian analysis.

Equation 3. $$a_{ij}|z_i \sim Bernoulli\ (z_i\theta_{ij})$$

Equation 3 describes the presence of the species' DNA in a water sample ($a_{ij}|z_i = 1$ where a species DNA is present, and $a_{ij}|z_i$ =0 when it is absent, at site *i*, for water sample *j*). This is a function of the occurrence of the species' eDNA as well as the availability probability θ_{ij}.

Equation 4. $$y_{ijk}|a_{ij} \sim Bernoulli(a_{ij}p_{ijk})$$

Equation 4 describes the detection process ($y_{ijk}|a_{ij}$=1 where the species is detected, and $y_{ij}|z_i$=0 where the species is not detected), at site *i* for water sample *j* and in replicate *k*. This is a function of the occurrence (z_i) of the species at site *i* as well as the probability of detecting the species' eDNA (p_{ijk}) at site *i* in water sample *j* in replicate *k*, given it is available to detect (a_{ij}). Prior distributions were specified on the logit scale, where ψ, p and β_p were normally distributed with a mean of 0 and a precision of 0.1.

Model selection

To select which covariates to use in the final model for both datasets, we used the waic function in the R package msocc (Stratton *et al.* 2020). We ran all model combinations for 1000 iterations each and selected the model with the lowest waic value. The selected model was then evaluated via inspection of Dunn–Smyth residuals (Warton *et al.* 2017) to determine whether covariates should be included as quadratic terms (Appendix C, Figures C.3-C.9).

Model fitting

We fit the model using the package rjags v.4-8 in R version 3.6.3 (R Development Core Team, 2020) and model code is provided in Appendix C. Three model chains were run for 350,000 iterations each. The first 50,000 samples were discarded as the burn-in, and the remaining samples were thinned by 1000, resulting in a posterior sample of 300 samples per chain. Convergence was checked visually using trace plots from the jagsUI package v.1.5.0, which indicated that chains were well mixed. Gelman-Rubin statistic values $\hat{R}$ were below 1.1, indicating successful convergence of chains.

Simulations

To ensure our model was able to accurately estimate effects of covariates on occupancy probability, we conducted model simulations. We simulated 500 datasets using the data

structure from our broad-scale sampling (504 sites, 2 water samples per site, 3 qPCR replicates per sample). We considered an ordinal covariate ranging from 1-5 (e.g., erosion), and a numeric covariate with a value between 0 and 1 (e.g., the proportion of a catchment covered in a particular land use type). Availability and detection probabilities were set at 0.9 and 0.8, respectively. The occupancy probability intercept (ψ) was varied between 0.2, 0.5 and 0.8. The β coefficients for the ordinal and numeric covariates were -1, -0.5, 0, 0.5 and 1. Each of the 500 simulated datasets was used to fit the model described above (Equations. 1-4), and the mean estimate of each β coefficient was retained, resulting in 500 estimates of each β coefficient for each ψ and β coefficient combination (e.g., ψ=0.2 and β=-1).

3.4 Results

Between August 2018 and May 2021, we visited 777 sites across the study area. We took water samples (total n = 1,008) from 504 of these sites spread across 159 contracted catchments, due to a lack of safe access or water availability at some sites. All negative DNA extraction and qPCR controls were negative. For the broad-scale dataset, platypus eDNA was detected in at least one qPCR replicate in one water sample at 272/504 sites and was detected in at least one qPCR replicate for both water samples at 241/504 sites (Figure 3.1A). For the Victorian data, platypus eDNA was detected in at least one qPCR in one sample at 150/289 sites, and in at least 1 qPCR in both samples at 128/289 sites (Appendix C, Figure C.2).

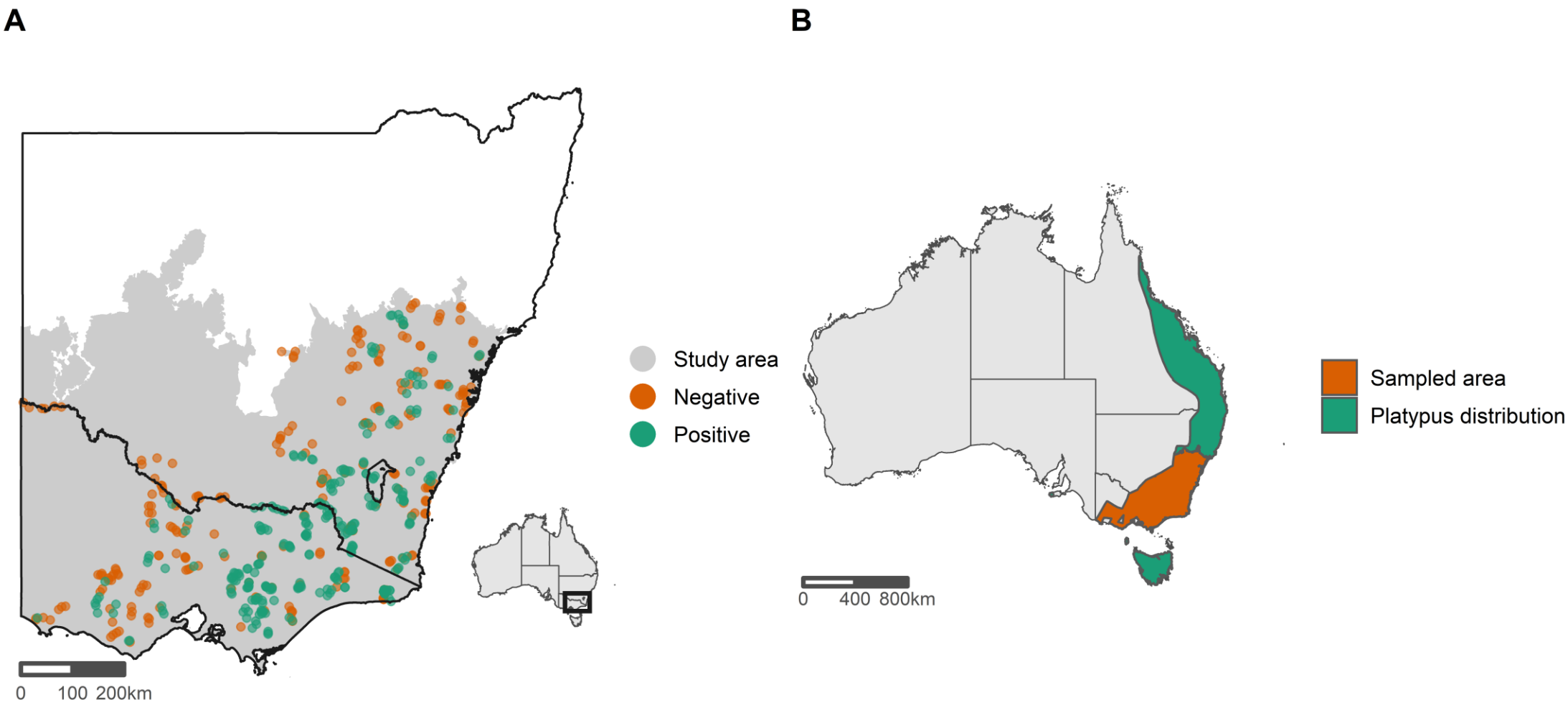

Figure 3.1: *A: Map of qPCR results from sampled sites. Orange=site was negative in all water samples and all qPCRs for platypus eDNA. Green= site was positive for platypus eDNA in at least 1 qPCR replicate in at least 1 sample. Grey=study area. B: Estimated platypus distribution (Woinarski and Burbidge 2016) and the overlap of estimated platypus distribution with the study area.*

Model selection

The model for the broad-scale analysis with the lowest waic value included 8 covariates: human disturbance, erosion, bank vegetation, burrowing banks, agriculture (cropping and pasture), grasslands and shrubs, forests, and runoff. Diagnostic plots based on Dunn–Smyth residuals suggested that grasslands and shrubs, forests, and runoff should be incorporated as quadratic terms (Appendix C, Figures C.3-C.6).

For the Victorian analysis, the model with the lowest waic value included three covariates: urban areas, runoff, and annual zero-flow days. Diagnostic plots based on Dunn–Smyth residuals of occupancy suggested that none of the covariates should be incorporated as quadratic terms (Appendix C, Figures C.7-C.9).

Occupancy modelling

The mean probability of occupancy across sites used in the broad-scale analysis was 0.55 (95% credible intervals (CI):0.20, 0.86), the availability probability was 0.90 (95% CI: 0.87, 0.93), and the detection probability was 0.78 (0.75, 0.80). The estimated number of occupied sites was 275.2/504. A confusion matrix is provided in Appendix C Table C.2 to compare the observed occupancy and estimated occupancy for each site. A threshold of 0.5 was used to convert the predictions to binary values. A plot of random effects is provided in Appendix C Figures C.10 and C.11. For the Victorian dataset, the mean probability of occupancy was 0.54 (95% CI: 0.26, 0.81), the availability probability was 0.89 (95% CI: 0.87, 0.93) and the detection probability was 0.78 (95% CI: 0.75, 0.81). The estimated number of occupied sites was 157.6/289.

For the broad-scale analysis, effects of covariates on occupancy probability were varied and predominately uncertain (Figure 3.2). Occupancy probability increased with less erosion and more suitable bank vegetation at the site scale, although effects were uncertain across all condition levels (β estimates overlapped zero). Overall, the quality of burrowing banks at the site scale had a positive effect on platypus occupancy (Figure 3.3). Using *very poor* as a reference category, coefficients were uncertain for the *poor* and *excellent* categories, although coefficients for both categories were, on average, larger than for preceding categories. Coefficients for the *fair* and *good* categories were the most certain (*fair* mean = 1.41, 95% credible interval (CI) = 0.64, 2.24; *good* mean = 1.93, CI = 0.84, 3.09) (Figure 3.2).

At the catchment scale, the coefficient for the effect of runoff was positive and 95% credible intervals did not overlap zero, suggesting that probability of occupancy increased with increasing runoff (mean = 0.71, 95% CI = 0.06, 1.37) (Figure 3.2). The linear and quadratic terms for grasslands and shrubs also did not overlap zero (linear mean = -0.85, 95% CI = -1.46, -0.37; quadratic mean = 0.19, 95% CI = 0.0004, 0.40). The coefficient describing the effect of agricultural land use on occupancy was negative and did not overlap zero (mean = -0.34, 95% CI = -0.70, -0.009). Thus, at the catchment scale, probability of platypus occupancy increased with water availability, and decreased with an increase in the proportion of a catchment covered by grasslands and shrubs, and agriculture (Figure 3.3).

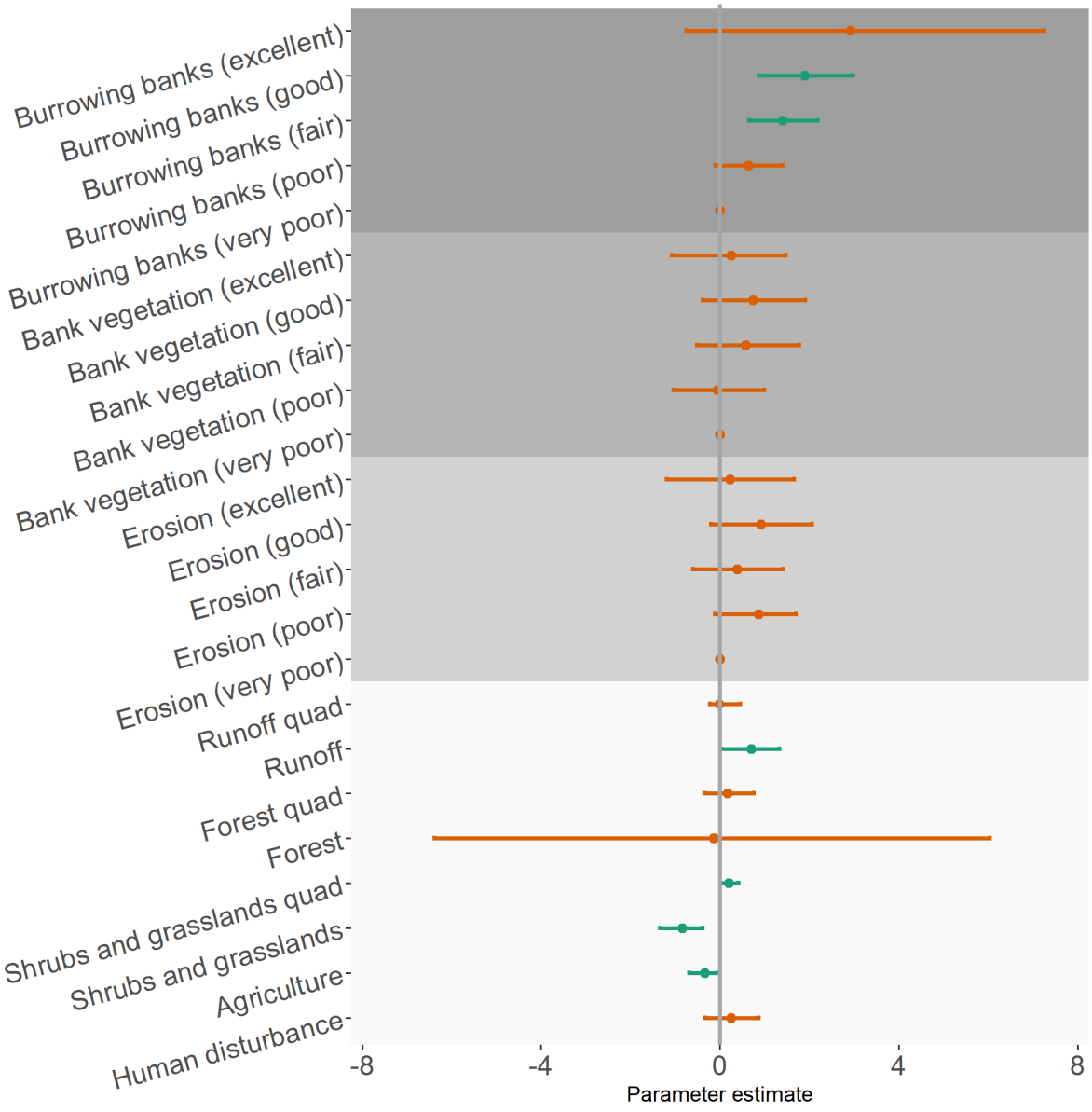

***Figure 3.2**: β coefficients describing effects of environmental covariates on platypus occupancy. Points are the mean and lines are the 95% credible intervals. Green lines do not overlap zero, orange lines do overlap zero. Bank vegetation and burrowing banks covariates are in comparison to a reference category (very poor).*

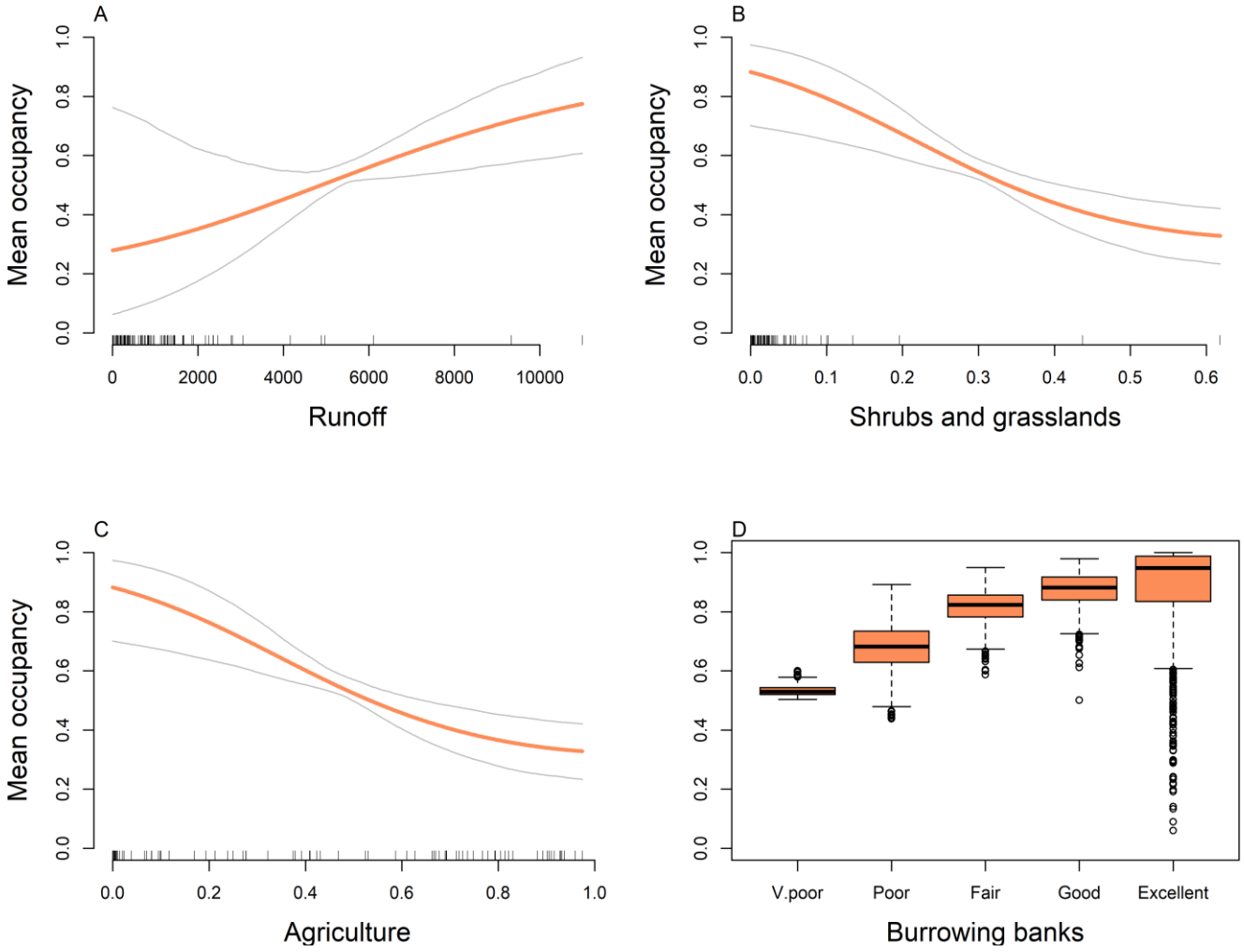

***Figure 3.3**: Platypus occupancy probability as a function of the proportion of the amount of runoff in a contracted catchment, the proportion of a catchment covered by agriculture or shrubs and grasslands, and the condition of banks for burrowing (reference category=very poor) (Broad scale analysis)*

For the Victorian dataset, the coefficient for the effect of urban land-cover on occupancy was negative but uncertain (Figure 3.4). Probability of platypus occupancy increased with runoff in the catchment (mean = 0.53, 95% CI = 0.18, 0.94), and with fewer annual zero-flow days (mean = 0.60, 95% CI = 0.247, 1.03) (Figure 3.5). See Figures C.12-C.19 (Appendix C) for plots of predicted occupancy as a function of all other covariates.

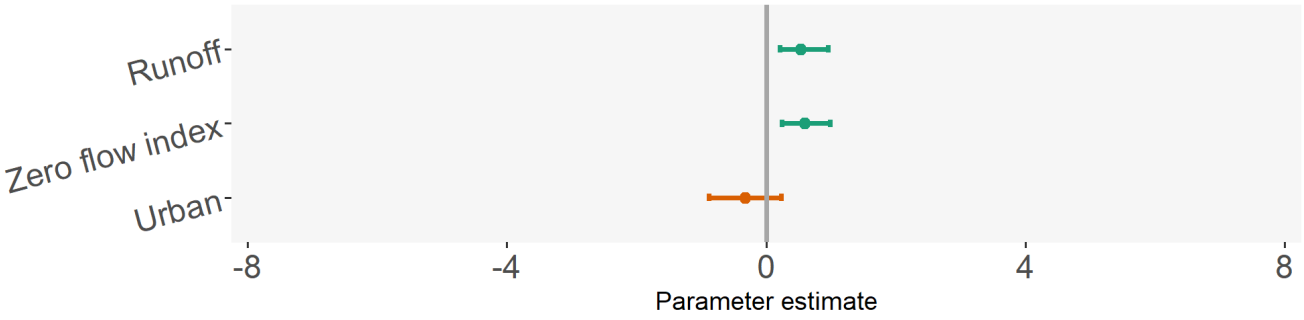

***Figure 3.4**: β parameter estimates for Victorian subset of data. Points are the mean; lines are the 95% credible intervals. Green lines do not overlap zero, orange line overlaps zero*

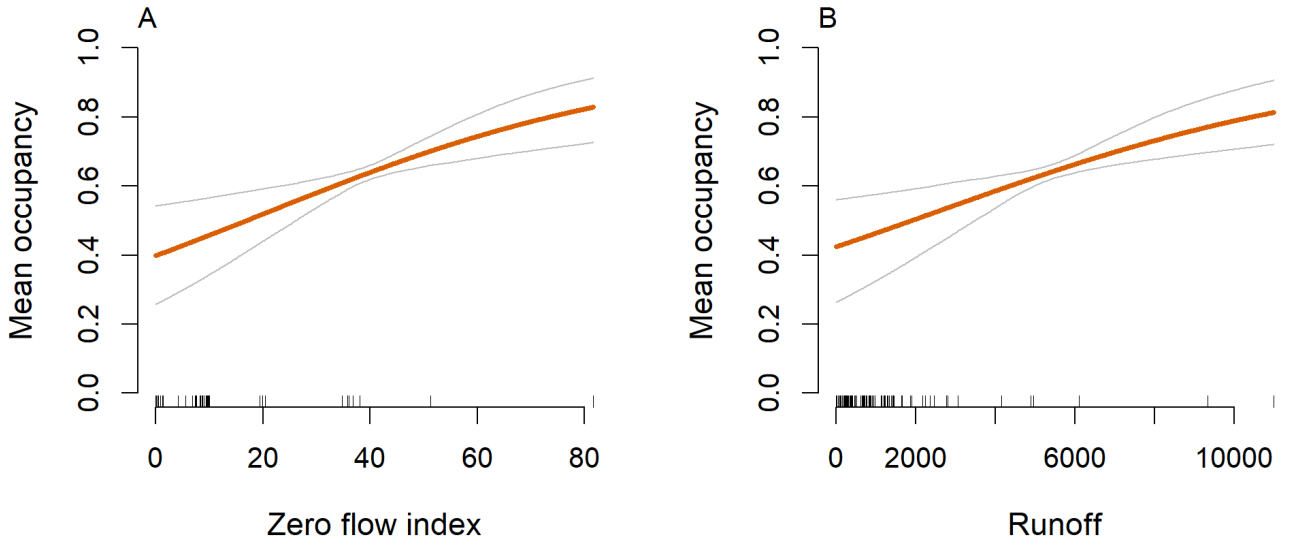

***Figure 3.5**: Response of occupancy probability to zero-flow index and runoff (Victorian data dataset)*

Simulations

Our model design was generally able to correctly estimate the β parameters we set, whether they were on the ordinal scale or numeric. For example, when β was set at 0.5, the estimated β from the model was between 0.50 and 0.52 when ψ=0.2, between 0.53 and 0.56 when ψ=0.5 and between 0.61 and 0.64 when ψ=0.8 (Appendix C, Figures C.20-C.34).

3.5 Discussion

Predicting and mitigating the impacts of global change on biodiversity requires an understanding of species contemporary distributions, and the factors that limit those distributions. Our results demonstrate that eDNA sampling can be used to rapidly obtain species distribution data over vast spatial extents, providing an efficient means to understand the effects of environmental covariates on occupancy. Below, we use our case study as an exemplar to describe the benefits of eDNA for surveying and modelling widespread species, such as the platypus, before discussing the implications of our findings for the monitoring and conservation of this iconic species.

Using eDNA to monitor and model species distributions

The platypus is an iconic Australian monotreme that is widely distributed across eastern Australia. With this wide distribution comes challenges for monitoring occupancy and distribution over space and time. Large spatial extents and remote sites necessitate efficient survey methods, so that species distributions can be sampled appropriately. This challenge is not specific to platypuses, but applies to many aquatic species across Australia, and globally (Bohmann *et al.* 2014; Deiner *et al.* 2017). Our case study demonstrates that eDNA sampling shows great promise in this regard. We undertook eDNA sampling across 584,292 km^2, encompassing approximately 37% of the platypus' distribution. To the best of our knowledge, this is the largest-scale use of aquatic eDNA sampling to date using a systematic study design. This comprehensive study will allow future surveys to understand whether distribution contractions are occurring through time, especially in the marginal areas of the platypus distribution.

The efficiency of eDNA sampling is further evident from the amount of time taken to sample this vast area. Environmental DNA samples can be taken quickly and efficiently in the field; sites can be sampled within a matter of minutes. Indeed, all eDNA sampling for this study occurred over 86 sampling days. Importantly, the efficiency of eDNA sampling does not come at the cost of reduced sensitivity; in fact, numerous studies have demonstrated higher sensitivity of eDNA sampling compared to traditional methods in aquatic environments (Piggott *et al.* 2021; Plante *et al.* 2021; Fediajevaite *et al.* 2021; McColl-Gausden *et al.* 2021), including studies of platypuses (Lugg *et al.* 2017). It is also important to note that all sampling was conducted predominately by two people. With a larger number of people working

synchronously, vast areas could be sampled, enabling widescale monitoring over a short period of time. Citizen science programs are one example where this type of synchronous sampling could be achieved, as has been demonstrated for the great crested newt in the United Kingdom (Biggs *et al.* 2015).

Our demonstration of the scalability of eDNA sampling is conservative, in that the large-scale and systematic nature of this study resulted in some sites being visited and not sampled. Sites were selected based on remote layers of watercourses, which in some circumstances can be inaccurate. In other cases, streams were dry or inaccessible because of current environmental conditions. The timescale of the study meant that we experienced a large range of environmental impacts on our sampling, from droughts to floods and widespread catastrophic wildfires. In these circumstances where selected catchments were impacted by environmental events, they were not sampled or not included in this analysis to reduce external variation in the data, as these factors were not accounted for in our analysis. However, the broad-scale sampling from our dataset demonstrated another use. We were able to provide extensive pre-fire data on platypus occupancy across the large area burnt by the 2020/2021 mega-fires in south-eastern Australia.

Monitoring efficiency could be further improved using eDNA metabarcoding, which enables synchronous monitoring of many species from different taxonomic groups, including those that are native, threatened, or invasive. This community-level approach could be easily incorporated into large-scale monitoring programs, as different survey techniques are not required for different species or groups (e.g., electrofishing for fish; acoustic surveys for anurans; trapping for platypus).

Our case study also demonstrates that eDNA sampling is amenable to sampling remote locations, which are likely to be encountered when monitoring widespread species. Environmental DNA sampling can be undertaken relatively safely (Crookes *et al.* 2020), as no direct access to the water is needed (sampling poles can be used if necessary). Furthermore, specialist training is not necessary (Ji *et al.* 2013), and sampling requires little equipment, which is in stark contrast to most traditional methods for monitoring aquatic species (e.g., trapping or electrofishing). In field preservation techniques are also improving, reducing the need for refrigeration or freezing of samples in remote locations (Spens *et al.* 2017).

The above considerations show that eDNA sampling unlocks the possibility of widescale monitoring of species distributions. It enables collection of baseline data on species distributions to be established, to monitor the effects of global change more accurately. This is a vital step in understanding and predicting how biodiversity will change in the future.

Effects of environmental covariates on platypus occupancy

Site occupancy-detection models are being increasingly used to analyse eDNA data, as they can account for the hierarchical nature of eDNA data (site, sample, and qPCR) and allow for imperfect detection to be incorporated at each of these levels. These models also allow for covariates of interest to be included at any of the levels. In this study, we investigated how landscape and site-level covariates impacted the occupancy of our target species. However, these models can be adapted so that effects of covariates on detection or availability probabilities can also be investigated.

Using our dataset, a general positive trend emerged, whereby platypuses were more likely to occupy sites with more suitable site characteristics, such as more suitable banks for burrowing compared to less suitable (very poor) sites. It has been shown in previous studies that platypuses prefer sites with overhanging vegetation to stabilise banks for burrowing and to provide cover for burrows (Grant and Temple-Smith 1998; Serena *et al.* 1998). Therefore, suitable bank vegetation and erosion levels are beneficial for platypus occupancy. Land uses such as grasslands and shrubs and agriculture showed a negative effect. These land use categories may not support the habitat requirements for platypuses at a site level. For example, they may be less likely to have bank or verge vegetation to reduce erosion of banks used for burrowing.

South-eastern Australia is a highly modified landscape. Approximately 44% of Australian forests and woodlands have been cleared since European settlement and 39% was cleared prior to 1972 (Metcalfe and Bui 2016). In recent years, land clearing has stabilised in the states of Victoria and NSW, where this study takes place, but increasing development in urban and peri-urban areas is putting pressure on remaining native vegetation and flood plains (Jackson *et al.* 2016). The historic distribution of platypuses is difficult to estimate due to a lack of historic data, and therefore declines are difficult to estimate. Recent estimates suggest however, that platypuses have declined from some catchments they used to inhabit (Hawke *et al.* 2019), largely attributed to a historical fur trade, river regulation and habitat destruction. Our data

suggests that currently, platypuses are less likely to occupy sites impacted by land clearing such as those with less suitable banks for burrowing, and in areas with grasslands and shrubs or agriculture. This supports previous studies considering the habitat preferences of platypuses (Grant and Temple-Smith 2003; Bino *et al.* 2020).

Our analysis also suggests that platypuses were more likely to occupy sites with a higher amount of runoff across the contracted catchment. Runoff in this study was used as a proxy for how much water was available in the system. The Victorian dataset also showed that platypuses were more likely to occupy sites in contracted catchments with a higher level of runoff and less zero-flow days (higher index value), highlighting the need for platypuses to have adequate flows. Platypuses feed predominately on aquatic benthic macroinvertebrates (McLachlan-Troup *et al.* 2010). Droughts and zero-flow days exacerbated by water regulation and extraction (Bino *et al.* 2019) could therefore have a negative impact on platypus occupancy due to prey availability or abundance (Marchant and Grant 2015). This has been evidenced in western Victorian where platypus populations dropped dramatically after a 10-year drought during the early 2000's (Griffiths and Weeks 2013; Griffiths *et al.* 2016; Griffiths and Weeks 2018).

The availability of water in the future, and how it relates to platypus occupancy is uncertain. There has however been medium agreement between climate models that there will be a substantial decrease in spring and winter rain in eastern Victoria and the Murray basin, and high model agreement for a substantial decrease in spring and winter rain for western Victoria (Hope *et al.* 2018). Summer and autumn rainfall predictions are uncertain because of meteorological processes. A decrease in rainfall could negatively affect platypus distribution into the future, given their occupancy is positively correlated with increased runoff and less zero-flow days. Drought as well as habitat clearing have been shown to negatively impact platypus metapopulation estimates with the combined effects of fragmentation, habitat destruction, and drought, under climate change projections platypuses are predicted to decrease metapopulation occupancy by between 35.6% to 55.5% (Bino *et al.* 2020).

Across both analyses, there were environmental covariates for which the relationship between occupancy and the covariate was uncertain. This could be because there is no effect of that covariate on the occupancy of this species, that we need more data to tease apart these relationships. For example, there were only 45/504 sites which had an excellent level in the erosion category. Whilst stratification occurred on the larger scale land use categories, fine-scale site-level variables could not be stratified in this study. However, a dataset such as this

can be built-on in the future, to provide more data which could help increase the sample size for some categories.

Despite some uncertainty in some selected land use or site categories, our results support previous research into the habitat requirements of platypuses. This demonstrates that large-scale, systematic eDNA surveys can be used to gather data on factors affecting the occupancy of species. These analyses and datasets are a starting point for larger scale data collection on platypus occupancy and can be used as a baseline structure for large-scale studies of other species.

3.6 Conclusions

This study represents the largest scale use of eDNA sampling using a systematic study design to date. We demonstrate that eDNA sampling can be conducted efficiently over large spatial scales in south-eastern Australia and that this data can be used to understand how landscape or site characteristics affect a species occupancy. These methods can be applied to aquatic species monitoring globally.

Chapter 4
The impact of the 2019/2020 mega-fires on platypus occupancy in south-eastern Australia

4.1 Abstract

Fire can play an important role in ecosystems, but extreme mega-fires are increasing the area burnt in forested regions globally. The 2019/2020 mega-fires in south-eastern Australia were unprecedented with respect to the amount of area burnt and the spatial extent of high severity fire. Yet, there is limited knowledge regarding the impact of these mega-fires on biodiversity, especially aquatic fauna. Here we investigate the impact of the 2019/2020 mega-fires on the distribution of a semi-aquatic monotreme thought to be in decline: the platypus (*Ornithorhynchus anatinus*). Whilst platypuses are buffered from many direct effects of fire, they may be impacted post-fire by vegetation loss, and increased run-off and debris. To investigate impacts of the 2019/2020 mega-fires, we leveraged extensive pre-fire environmental DNA (eDNA) sampling across south-eastern Australia to conduct a Before-After Control-Impact study. We sampled 118 sites pre-fire (2018/2019) and sampled the same sites again at two timepoints post-fire (2020 and 2021). We used site occupancy-detection modelling to estimate the probability of platypus occupancy across burnt and unburnt sites during the different sampling periods. We also investigated the extent to which rainfall post-fire interacted with the proportion of a watershed burnt at high severity. We did not detect a significant effect of fire presence on platypus distribution: occupancy probability increased similarly post-fire in both burnt and unburnt sites. However, we did detect a significant interaction between rainfall post-fire and the spatial extent of high severity fire: platypus occupancy was predicted to be lower at sites where there was high rainfall post-fire, and a large proportion of the watershed burnt at high severity. This finding is consistent with previous studies of fire impacts, whereby the greatest threat to aquatic fauna is high rainfall post-fire. With area burnt increasing globally, and predictions of more extreme rainfall events in south-eastern Australia, the impact of fire on aquatic fauna needs greater consideration in post-fire assessments and biodiversity management more generally.

4.2 Introduction

The 2019/2020 "black summer" mega-fires were unprecedented in scale and severity (Boer *et al.* 2020; Bowman *et al.* 2020; Collins *et al.* 2021). Thirty-three people were killed, 3,000 houses were destroyed, and almost 19 million hectares were burnt (Filkov *et al.* 2020). The fires occurred near the conclusion of the warmest and driest year on record for both Australia, and the area affected (south-eastern Australia) (Abram *et al.* 2021). Averaged across south-eastern Australia, annual mean rainfall was the lowest on record in 2019 (Abram *et al.* 2021). The 2019/2020 fire season resulted in a cumulative fire extent that was greater than in previous seasons, and the absolute area of high severity fire was larger than previously observed (Collins *et al.* 2021). However, these mega-fires were not an isolated incident. There is evidence that forest fires are becoming larger in extent and more frequent in many biomes across Australia (Nguyen *et al.* 2020). Yet the impact of fire and changing fire regimes on biodiversity, particularly aquatic fauna, remains largely uncertain (Verkaik *et al.* 2014).

Fire severity is defined as the loss or change in organic matter resulting from fire (Keeley 2009). In Australian forests and woodlands, fire severity is typically quantified as the degree of scorch and consumption of the foliage (Hammill and Bradstock 2006). Fire severity can be important for determining the effects of fire on species (Smucker *et al.* 2005; Chia *et al.* 2016), as severity determines both direct and indirect effects via changes to vegetation and habitat suitability (Benyon and Lane 2013; Collins 2020). For example, the indirect effects of wildfire on aquatic fauna may depend on interactions between fire severity and post-fire rainfall. Erosion and sedimentation should increase in areas with a high degree of foliage removal (i.e., high severity fire), and in the presence of high rainfall events. However, these effects may be short-lived in resilient ecosystems such as eucalypt forests, owing to the prevalence of species that can resprout shortly after fire (Clarke *et al.* 2015).

The platypus is an iconic, semi-aquatic monotreme that was highlighted as potentially impacted by the 2019/2020 mega-fires by the Australian Commonwealth (Legge *et al.* 2020). Platypuses feed on macroinvertebrates and use the banks of rivers and streams for burrowing (Serena *et al.* 1998). The species is widespread, ranging from Tasmania in Australia's temperate south, to Queensland in mainland Australia's tropical north (Woinarski and Burbidge 2016). Historical data on platypus distribution is sparse, but the species was previously assumed to be common. However, there is increasing evidence to suggest the platypus is under threat, particularly in

urban and agricultural landscapes (Woinarski and Burbidge 2016). Indeed, available evidence suggests platypus populations may have been extirpated or are currently declining in some catchments, such as those in Queensland and western Victoria (Bino *et al.* 2020). Yet the impact of fire on platypuses remains poorly understood. Wildfires were identified as a major factor in the extirpation of an isolated platypus population near Melbourne, Victoria (Serena and Williams 2004). In addition, population viability analyses suggest that the synergistic effects of drought and fire could impact platypus recruitment and density (Bino *et al.* 2021). However, environmental, or human induced flows could potentially ameliorate the impacts of fire on platypuses. At a site in western Victoria, consistent flows were released from an upstream impoundment after an intense but localised fire, and effects of the fire on platypus populations were limited (Griffiths and Weeks 2014). Nonetheless, our current understanding of fire impacts on platypuses largely stems from anecdotal observations or studies encompassing few sites. To date, there have been no rigorous landscape-scale studies of platypus responses to fire.

As a semi-aquatic species, the platypus is largely buffered from the primary impacts of fire (e.g., direct mortality). However, secondary impacts, including habitat degradation and reduced water quality, may pose a significant threat. Recently-burnt areas are often susceptible to erosion, which, in combination with heavy rainfall post-fire, can result in debris flows (Nyman *et al.* 2020) or the movement of suspended sediment "slugs" and nutrients downstream (Shakesby *et al.* 2007). For example, two weeks after the containment of the 2020 wildfire in Green Wattle Creek, New South Wales, Australia, 276 mm of rain fell in 72 hours (Neris *et al.* 2021), resulting in sediment and ash loads being transported into a reservoir used for the urban drinking supply of Sydney. Fish abundance can decline as a result of fire-induced sediment slugs moving downstream, and impacts on abundance can be evident even after 12 months (Lyon and O'Connor 2008). Furthermore, aquatic macroinvertebrate assemblages can be altered by fire (Verkaik *et al.* 2014; Verkaik *et al.* 2015; Robson *et al.* 2018) and abundances can be reduced, especially when high rainfall events and flooding occur after fire (Vieira *et al.* 2004). Aquatic macroinvertebrates are a primary food source for platypuses (Serena *et al.* 2014), and thus impacts on macroinvertebrates may impact platypus survival post-fire. While the mechanisms outlined above remain speculative, they suggest that there are multiple pathways through which platypus occupancy may be impacted by wildfire.

As wildfire occurrence is difficult to predict, most studies of wildfire impact on fauna have relied on space-for-time study designs (e.g., Bino *et al.* (2021), Verkaik *et al.* (2015)). However, to rigorously establish whether a species' distribution has been impacted by a wildfire event, even in the short term, data on species occupancy prior to the event is crucial. As platypuses are predominately nocturnal, incidental observations are not particularly common. Furthermore, traditional sampling methods (fyke and gill nets; Serena *et al.* 2014) are time-consuming to deploy and check, prohibiting routine landscape-scale monitoring. Environmental DNA (eDNA) sampling is a technique that has been used more recently to efficiently and accurately survey platypus occupancy across large areas (Lugg *et al.* 2017), helping to alleviate the issue of a lack of baseline monitoring data.

Here, we investigate the impact of the 2019/2020 Australian mega-fires on platypus occupancy in south-eastern Australia using an extensive dataset of pre-and post-fire eDNA samples. We use this dataset to address the following questions: (i) does platypus occupancy decline in burnt sites more than in control sites post-fire? and (ii) does the interaction between rainfall and fire severity impact platypus occupancy?

4.3 Methods

Site selection

The study took place in south-eastern Australia. Sites were selected based on the presence of pre-fire eDNA sampling and proximity to burnt areas (Figure 4.1). A total of 471 sites were selected as candidates to sample post-fire. We used two different definitions to classify a site as burnt: (i) whether fire was present in the watershed of a site; and (ii) whether a site was directly inside the burnt extent. Analyses were repeated with both definitions.

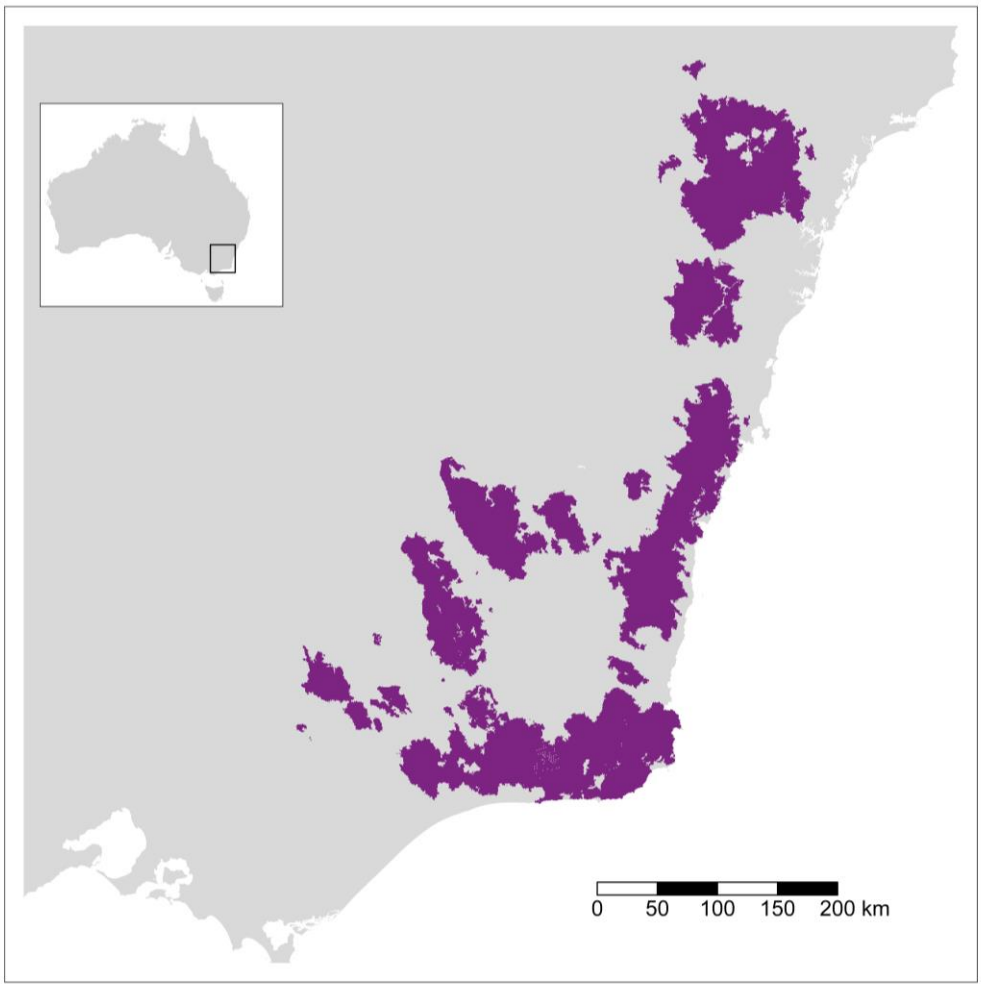

Figure 4.1*: Study area in south-eastern Australia with fire extent in purple*

eDNA sampling

Sampling occurred during three time points, hereafter called *pre-fire*, *post-fire 1* and *post-fire 2*. Pre-fire eDNA sampling took place in 2018 and 2019. Post-fire 1 sampling took place in March-November 2020 and post-fire 2 sampling in March-June 2021. Sampling was focused, where possible, on the peak activity times of platypuses (i.e., spring and autumn), with some opportunistic sampling occurring in winter of 2020 due to sampling restrictions imposed by the COVID-19 pandemic. No sampling occurred during the summer months of November to February when females and offspring spend large amounts of time in burrows. At each site, two independent water samples were taken as per Lugg *et al.* (2017). Two disposable syringes (60-mL Luer Lock sterile syringes; Hapool medical Technology, Shandong, China) were used to draw water from different microhabitats along the water's edge. Multiple draws of 60-mL were used to push 500 ml through a detachable 0.22 µm filter by hand (GP 22µm Filter Unit;

Sterivex, EMD Millipore corporation, Billerica MA, United States). The maximum volume was not always achieved due to the filter becoming clogged. Samples were stored in a car fridge for a maximum of three days until frozen, or frozen on site until DNA could be extracted in the lab.

DNA extraction and qPCR

DNA was extracted from Sterivex filters using Qiagen DNeasy Blood & Tissue Kits (Spin-column protocol) in a room that is dedicated to low-quantity DNA sources with qPCR setup undertaken in a sterilised laminar flow hood. Modifications to the manufacturers protocol were the same as in Lugg *et al.* (2017). Included in every batch of DNA extractions (~24) was a negative DNA extraction control (that replicated the process using a sterile Sterivex filter).

Species-specific primers and TaqMan® probe targeting the platypus *cyt b* mitochondrial control region used in this study were from Lugg *et al.* (2017) with the same assay conditions including custom TaqMan® gene expression assay from Life Technologies (Thermo Fisher Scientific). Amplification conditions were the same as in Lugg *et al.* (2017). A negative qPCR control with no DNA template and all DNA negative controls extracted at the same time as samples were also included. Crossing point (Cp) values were determined from the amplification profiles of PCR using the Absolute Quantification module of the LightCycler® 480 software package. To test for the presence of PCR inhibitors, a TaqMan® Exogenous Internal Positive Control VIC probe was also run on each sample. If inhibition was detected, samples were diluted (1:10) and the qPCR analysis was repeated.

Environmental variable selection

Three environmental covariates were included in one of our analyses; fire severity, rainfall and the volume of water collected (Figure 4.2). Fire severity and rainfall were calculated at the watershed scale to incorporate potential impacts of sediment flows and erosion post-fire, which could impact aquatic organisms such as platypuses or their prey (Vieira *et al.* 2004). Watersheds were calculated for each sampling point using the watershed tool from Whitebox Tools (Lindsay 2014) in QGIS (QGIS Geographic Information System 2021).

Fire severity was mapped for forests, woodlands and shrublands across wildfires within each watershed using Landsat imagery and a Random Forest classifier (Collins *et al.* 2018). The

random forest classifier has been extensively trained and validated using data from wildfires (n=44) across south-eastern Australia and has been found to have very high classification accuracy (~88%) (Collins *et al.* 2020; Collins *et al.* 2021). Five severity classes were mapped as: including unburnt areas, low canopy scorch (<20% scorch), moderate canopy scorch (20-80%), high canopy scorch (>80%) and canopy consumption. Severity predictions were reclassified as either high or low severity based on impacts to the canopy foliage in a vegetation type (Appendix D, Table D.2). The high severity class incorporated defoliated canopy in forest, woodland and shrubland areas, as well as burnt grasslands or pasture. Low severity included canopies which were intact or partially defoliated in forest, woodland or shrublands. These severity classes were targeted as areas with canopy defoliation are more prone to sediment flows during heavy rainfall following fire (Dragovich and Morris 2002; Blake *et al.* 2020). The proportion of each watershed covered by each of these severity classes (30 m resolution) was used in the analysis of platypus occupancy described below.

Total daily rainfall (ml) across each watershed for the 60 days post-fire was calculated using gridded rainfall data from the Australian Bureau of Meteorology (Bureau of Meteorology 2021) and Moderate Resolution Imaging Spectroradiometer (MODIS) burn date data (Giglio *et al.* 2015). The last date of a fire burning in a watershed was used to determine the commencement of the 60-day post-fire period. For example, if a fire in a watershed was extinguished on January 10th, daily rainfall for the 60 days after January 10th was summed across the watershed. If a watershed did not have available date data for a particular fire, the latest date within a 500 m buffer was used. All covariates were scaled and centred prior to analysis.

The volume of water collected in each sample was included, as water volume could impact eDNA availability (Mächler *et al.* 2016). A summary of water volume data is presented in Appendix D, Table D.3. The mean water volume collected across samples and visits varied between 235.1 ml and 300.7 ml.

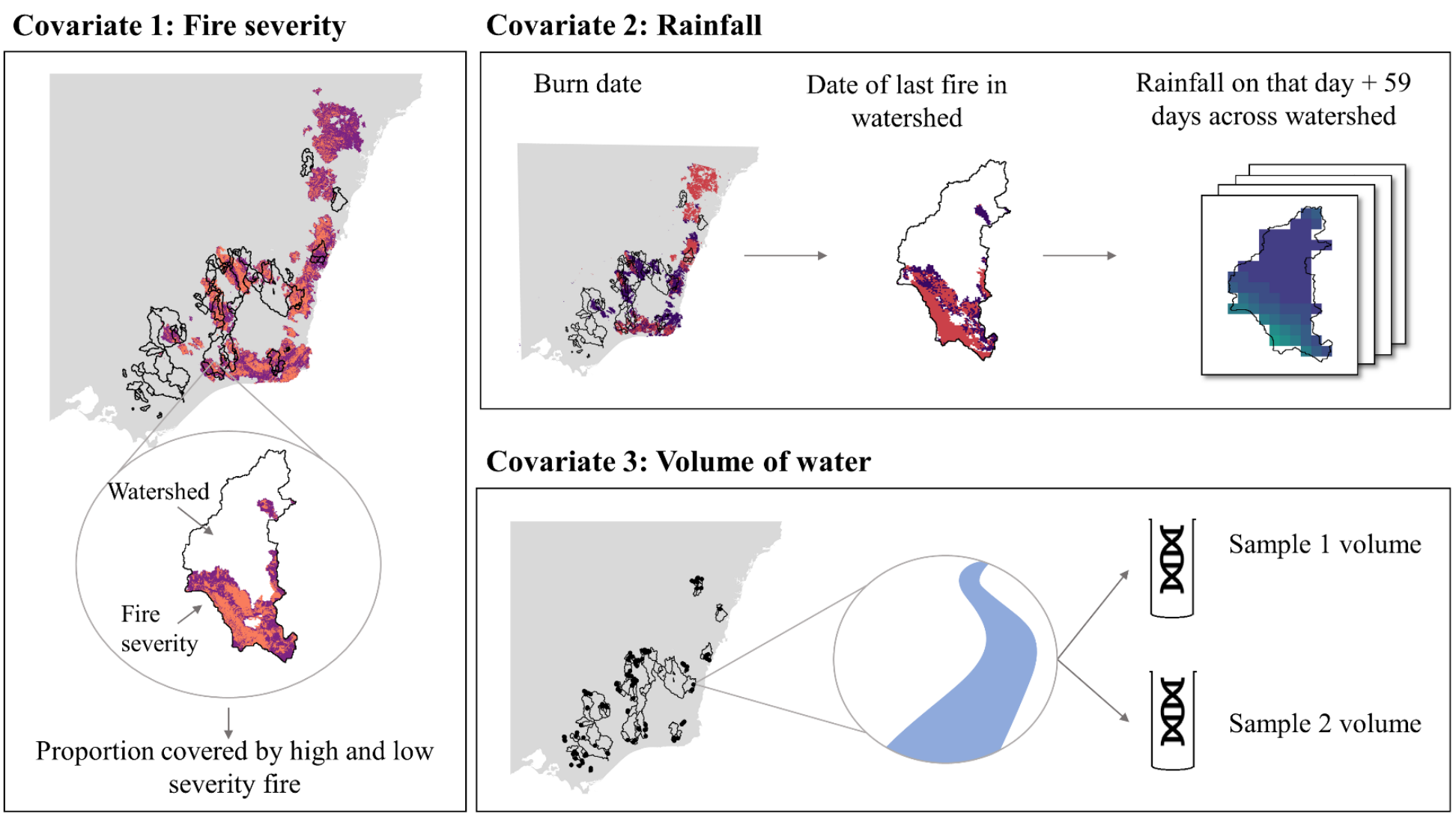

***Figure 4.2**: Framework for calculating environmental covariates for use in the model. Covariate 1 shows fire severity classes across the landscape in orange (high severity) and purple (low severity) and the watersheds used for each site. The proportion of each watershed covered by each severity type was then calculated. Covariate 2 shows how rainfall was calculated for each site. The date of the last fire for each watershed was used to sum the rainfall across a watershed for the 60 days post-fire. Covariate 3 shows that the volume of water for each sample at each site was recorded and used in the model.*

Model description

Two models were used in this study. The first model considers wildfire to be a binary variable, whereas the second model considers the finer scale variability of wildfire by incorporating fire severity and post-fire rainfall across a watershed. For model 1, observed data was in the form of binary detection/non-detection observations of platypus eDNA during visit *i* (1,2) at site *j* (1,2…*n*) in water sample *k* (1,2) and qPCR replicate *l* (1,2,3).

Equation 1 $$z_{ij} \sim Bernoulli(\psi_{i,j})$$

Equation 1 describes the latent occupancy of the species at site *j* during visit *i* ($z_{i,j}$=1 where the species is present, and $z_{i,j}$=0 where the species is absent), given the probability of occurrence of the species at site *j* during visit *i* ($\psi_{i,j}$).

Equation 2 $$logit(\psi_{ij}) = logit(intercept.\psi_j) + \beta_{burnt,i,j} + \beta_{after1,i,j} + \beta_{after2,ij} + \beta_{after1burnt,i,j} + \beta_{after2burnt_{,i,j}}$$

Equation 2 shows that the probability of species occurrence at site *j* during visit *i* is affected by a sites pre-fire occupancy probability ($intercept.\psi_j$), whether the site is burnt or unburnt during the 2019/2020 wildfire season ($burnt, i, j$, a binary indicator) and time since fire ($after1, i, j$ and $after2, i, j$, both binary indicator variables with 'pre-fire' as the reference category). An interaction between burnt status and time since fire ($\beta_{after1burnt,i,j}$, $\beta_{after2burnt,i,j}$) is also included.

Equation 3 $$a_{ijk}|z_{ij} \sim Bernoulli\ (z_{ij}\theta_{ijk})$$

Equation 3 describes the latent presence of the species' eDNA in a water sample during visit *i* at site *j* for water sample *k* ($a_{ijk}|z_{ij} = 1$ when a species' eDNA is present, and $a_{ijk}|z_{ij} = 0$ when it is absent), which is a function of eDNA presence z_{ij} and the probability that eDNA is detected in a water sample θ_{ijk}.

Equation 4 $$cloglog(\theta_{ijk}) = cloglog(intercept.\theta) + \beta_{volume} * volume_{i,j,k}$$

Equation 4 shows that the probability of the species' eDNA being available for detection in water sample *k* is affected by the volume of water collected during visit *i* at site *j* in water

sample *k*. A cloglog link function that rescales the linear predictor to the interval [0, 1] was used (Smart *et al.* 2015).

Equation 5 $$y_{ijkl}|a_{ijk} \sim Bernoulli(a_{ijk}p_{ijkl})$$

Finally, Equation 5 describes the observed detection data ($y_{ijkl}|a_{ijk}$=1 where the species' eDNA is detected, and $y_{ijk}|z_{ij}$=0 where it is not), during visit *i* at site *j* for water sample *k* and qPCR replicate *l*. This is a function of the occurrence of the species' eDNA in sample *k* at site *j* during visit *i* (a_{ijk}), as well as the probability of detecting eDNA via qPCR (p_{ijkl}).

Environmental variables used in model 2 are described in Appendix D Table D.1. Model 2 used the same basic structure as model 1 and is described in the supplementary material. Occupancy covariates investigated in model 2 were: (i) post-fire 1 and (ii) post-fire 2, as above; proportion of a watershed classified as (iii) high severity fire and (iv) low severity fire; (v) rainfall in the watershed post-fire; and interactions between (vi) rainfall and high severity fire, (vii) rainfall and low severity fire, (viii) low severity fire and post-fire 1, (ix) low severity fire and post-fire 2, (x) high severity fire and post-fire 1, (xi) high severity fire and post-fire 2, (xii) rainfall post-fire and post-fire 1, (xiii) rainfall post-fire and post-fire 2. The volume of water used for each sample was used as an availability probability covariate. Prior distributions were specified on the logit scale, where βs were normally distributed with a mean of 0 and a precision of 0.1 and $intercept.\psi_j$, and $intercept.\theta$ were uniformly distributed across the range of 0-1.

We fit the models using the software JAGS through the package rjags v.4-8 in R version 3.6.3 (R Development Core Team, 2020). JAGS models are provided in Appendix D. Three model chains were run for 30,000 iterations each. The first 10,000 samples were discarded, and the remaining samples were thinned by a factor of 10, resulting in 2,000 samples per chain from the posterior distribution. Convergence was measured using trace plots from the jagsUI package v.1.5.0, and Gelman-Rubin statistic values $\hat{R}$ were below 1.1, also indicating successful convergence of chains.

4.4 Results

We sampled a total of 285 unique sites, spanning approximately 133,249 km^2. However, limited post-fire access at some sites coupled with COVID-19 restrictions meant that not all sites were sampled in all time periods (i.e., some sites were sampled pre-fire but not post, or

were sampled post-fire but not pre-fire). A total of 118 sites were sampled in all three time points and will be used in the subsequent analysis. Histograms and scatter plots of the environmental data (proportion of watershed covered by low severity fire, and high severity fire, cumulative rainfall in a watershed post-fire) are presented in Appendix D, Figures D.1-D.4. Platypus DNA was detected at 95 sites prior to the fires, 89 sites in post-fire 1, and 92 sites in post-fire 2. Between pre-fire sampling and post-fire 1, 8 sites changed from platypus not detected to platypus detected and 14 sites changed from platypus detected to not detected. Between post-fire 1 and post-fire 2, 13 sites changed from platypus not detected to platypus detected and 10 sites changed from platypus detected to not detected. For the following analysis, a site was classified as burnt if fire was present in its watershed. Using this definition, 57 sites were burnt and 61 were unburnt (Figure 4.3).

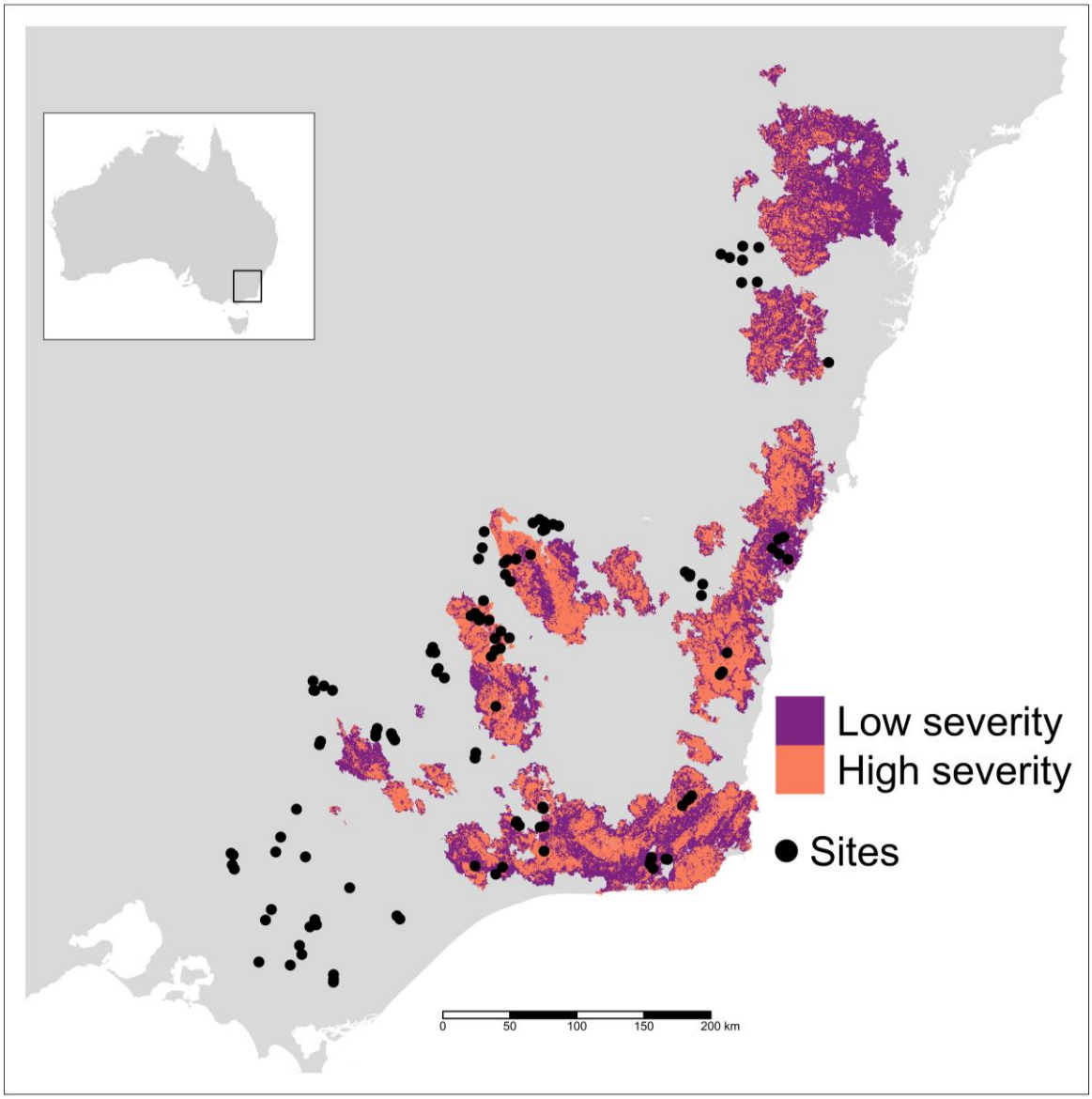

Figure 4.3*: Map of 118 sites sampled in all three time periods including fire extent with fire severity classifications*

Model 1

Using model 1, and when a site was classified as burnt when there was fire present anywhere within the watershed, the predicted number of occupied sites was relatively similar and also stable between the two groups (burnt and unburnt) across time (Figure 4.4). For burnt sites, the predicted number of occupied sites decreased from 44.2 pre-fire to 40.9 post-fire 1 and

increased to 44.9 in post-fire 2. For unburnt sites, the predicted number of occupied sites decreased from 51.2 pre-fire to 49.7 in post-fire 1 and to 48.9 in post-fire 2. A confusion matrix for each sampling period is provided in Appendix D Table D.4 to compare the observed occupancy and estimated occupancy for each site. A threshold of 0.5 was used to convert the predictions to binary values.

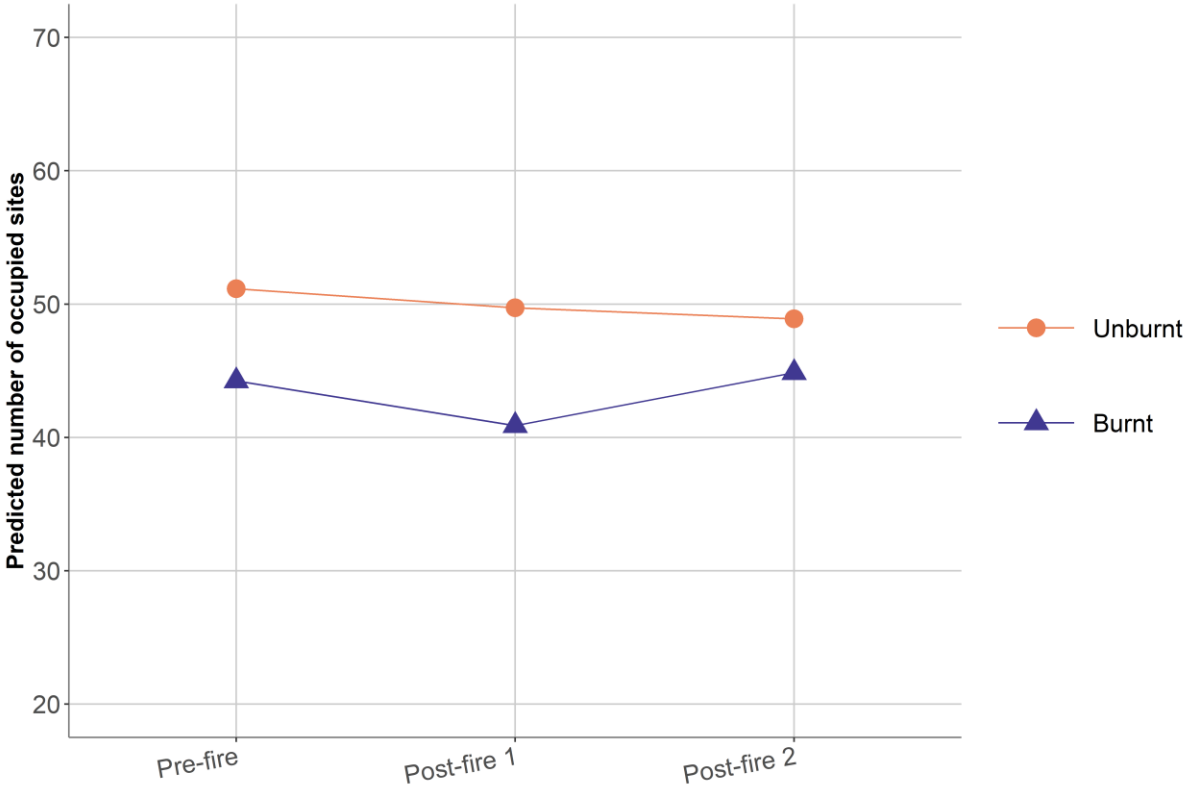

***Figure 4.4**: Predicted number of occupied unburnt and burnt sites, before the 2019/2020 fires, during post-fire sampling time period 1, and during post-fire sampling time period 2. Burnt sites = anywhere within a watershed is burnt.*

Using model 1, the β parameter estimates for the post-fire 1 and post-fire 2 variables were positive and did not overlap zero (post-fire 1 mean = 1.50, 95% credible interval (CI) = 0.61, 2.44; post-fire 2 mean = 1.40, 95% CI = 0.54, 2.46; Figure 4.5). This suggests that across all sites (burnt and unburnt), probability of platypus occupancy was higher in post-fire 1 and post-fire 2 than during pre-fire sampling. Interactions between burnt status and post-fire 1, and between burnt status and post-fire 2, were weak and uncertain (95% CIs overlapped zero). The effect of sample volume was also weak and uncertain.

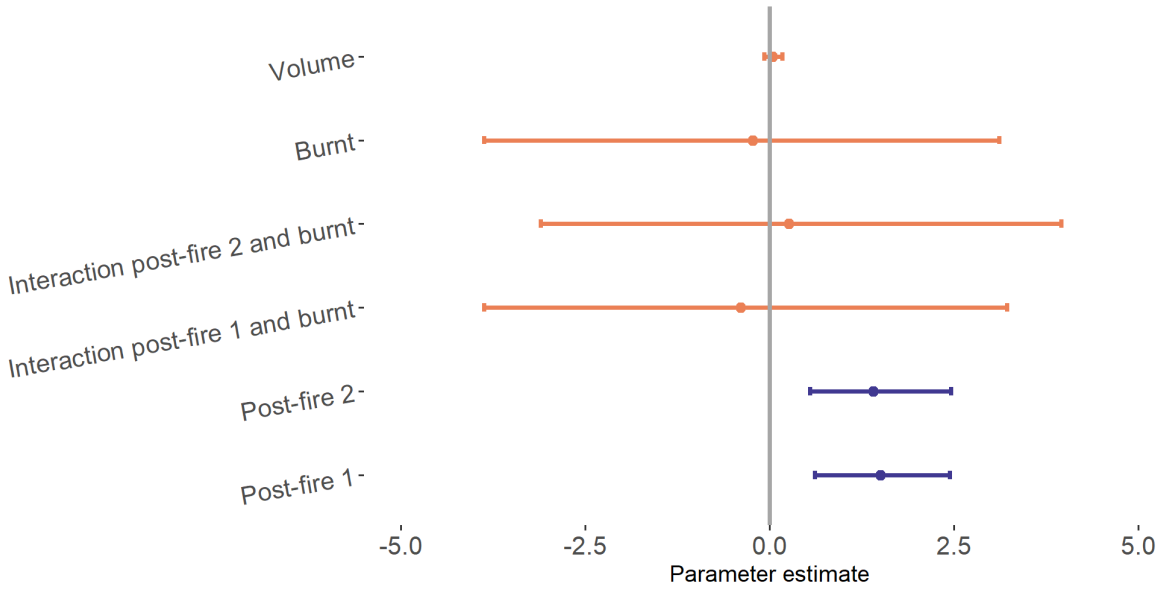

***Figure 4.5**: β parameter estimates displaying the effect of covariates on occupancy or availability (volume only) probabilities. Points represent the mean parameter estimate and lines are the 95% credible intervals. Dark purple lines do not overlap zero.*

We also used model 1 to explore the generality of our findings, classifying a site as burnt only if it fell directly within the mapped fire extent (as opposed to anywhere within the watershed being burnt, as above). The predicted number of occupied sites differed when using this classification of a burnt site (Figure 4.6). This classification of sites resulted in 79 sites being classed as unburnt and 43 as burnt. For burnt sites, the predicted number of occupied sites decreased from 38.1 in pre-fire to 31.7 I post-fire 1 and an increase to 32.8 in post-fire 2. For unburnt sites, the predicted number of occupied sites increased from 60.34 pre-fire to 61.1 in post-fire 1 and 64.2 in post-fire 2. Coefficient plots are presented in Appendix D, Figure D.5.

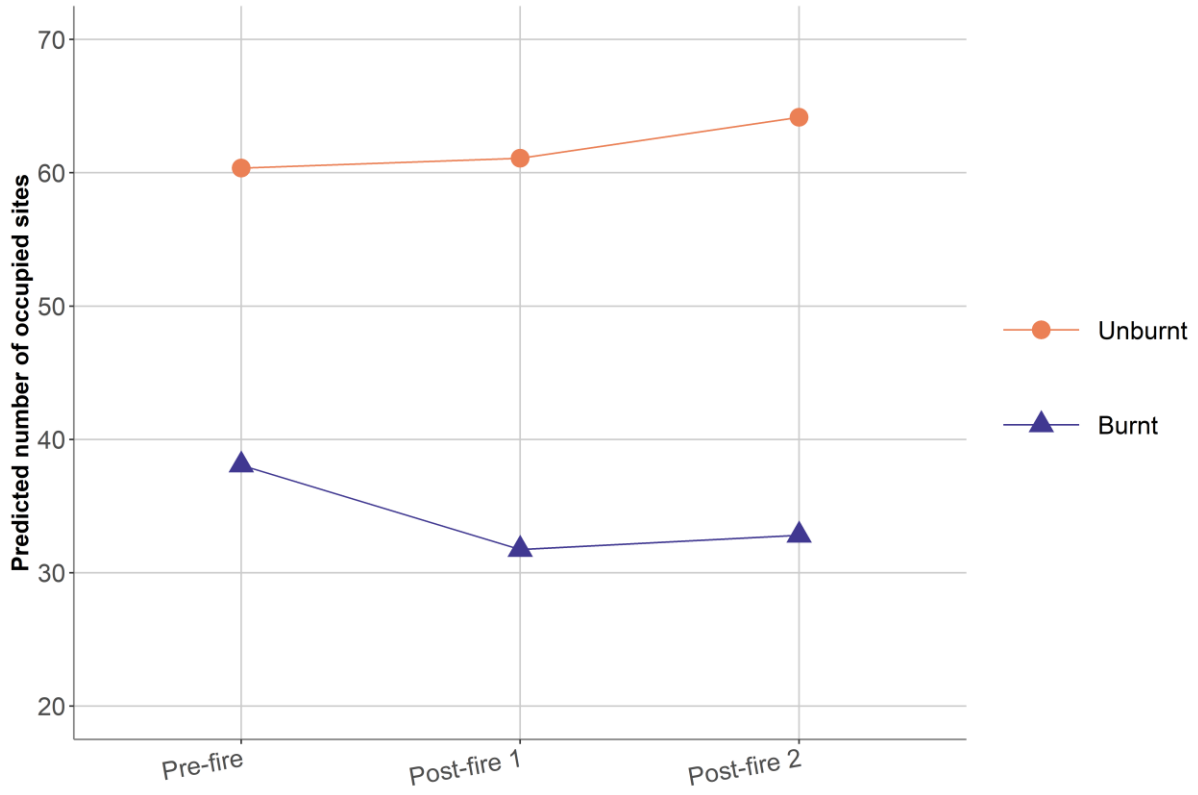

***Figure 4.6**: Predicted number of occupied unburnt and burnt sites, before the 2019/2020 fires, during post-fire sampling time period 1, and during post-fire sampling time period 2. Burnt sites = within the burnt extent.*

Model 2

The proportion of each watershed burnt by low severity fire or high severity fire varied from 0–1, with a mean of 0.15 for the former and 0.19 for the latter. Cumulative rainfall in a watershed during the 60 days post-fire varied from 67 mm to 50,236 mm with a mean of 2,652 mm. Using these variables in model 2, we found that the majority of β estimates were close to zero and/or had 95% CIs that overlapped zero (Figure 4.7). However, like model 1, estimates for post-fire 1 and post-fire 2 were positive and 95% CIs did not overlap zero (post-fire 1 mean = 1.38, 95% CI = 0.75, 2.10; post-fire 2 mean = 1.78, 95% CI = 1.06, 2.63). Interestingly, there was a strong negative interaction between rainfall post-fire and the proportion of a watershed burnt by high severity fire (mean = -2.30, 95% CIs = -4.28, -0.63). There was also a relatively strong (but more uncertain) interaction between rainfall post-fire and the proportion of a watershed burnt at low fire severity (mean = 1.67, 95% CIs = -0.50, 4.00).

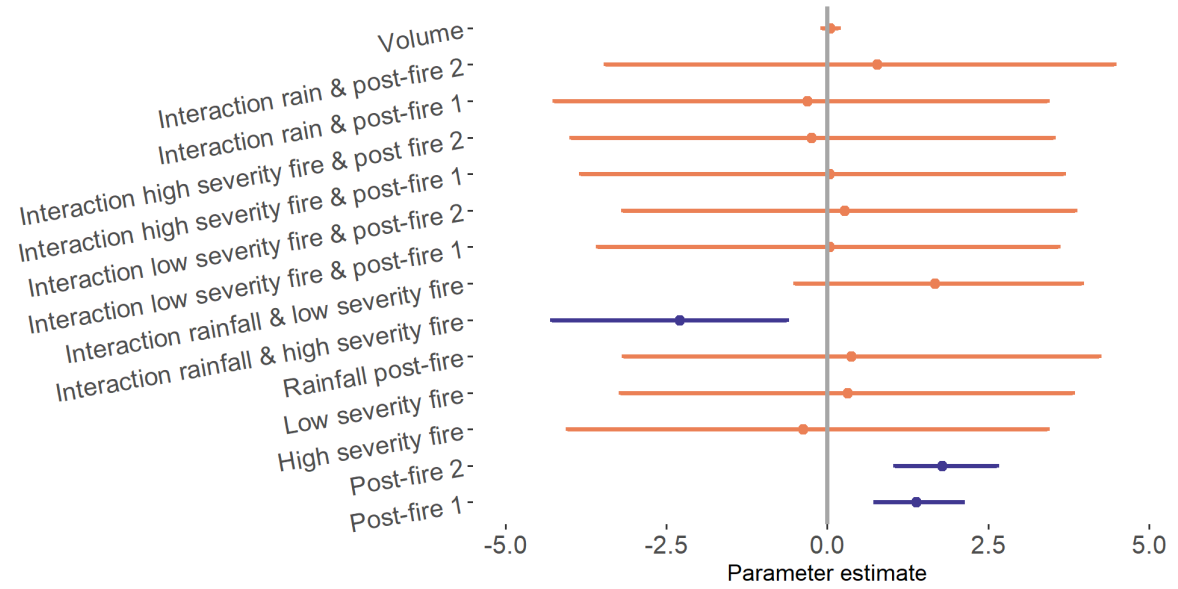

***Figure 4.7**: β parameter estimates for the effect of a covariate on occupancy, or availability (volume only) probabilities. Points are the parameter mean estimate and lines are the 95% credible intervals. Dark purple lines do not overlap zero, orange lines overlap zero*

Finally, we visualised how the interaction between rainfall post-fire and the proportion of a watershed burnt at high severity affected the probability of platypus occupancy (Figure 4.8). Moderate to high values of rainfall combined with moderate to high proportions of area burnt at high severity resulted in predictions of very low platypus occupancy probability. Conversely, sites in areas with moderate to high amounts of rainfall, but a low proportion burnt at high severity, had high probability of occupancy. A moderate occupancy probability was predicted where there were low levels of rainfall and a low to high proportion of a watershed burnt at high severity.

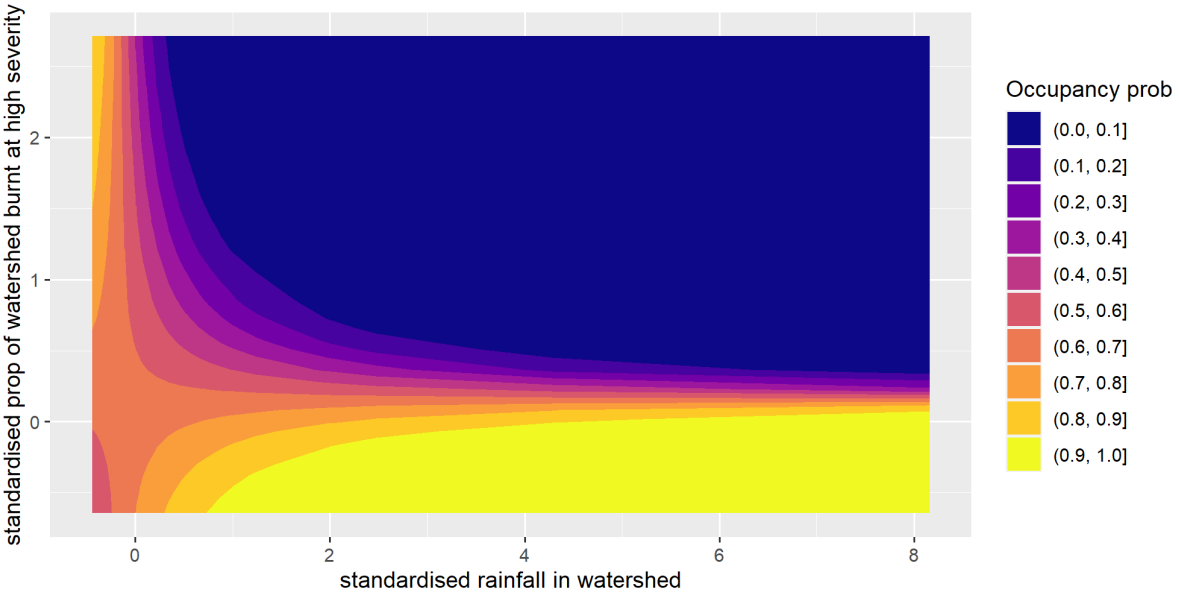

***Figure 4.8**: Contour plot of the relationship between the amount of rainfall in a watershed for the 60 days post - fire and the proportion of the watershed burnt at high severity and how this effects platypus occupancy probability.*

4.5 Discussion

Our results suggest that, in the case of the 2019/2020 Australian mega-fires, the presence or absence of fire did not affect platypus occurrence. However, we did detect an impact on platypus occurrence when including more proximate measures of fire impact. Specifically, we found that platypus occurrence depended on the interaction between fire severity and post-fire rainfall: watersheds with a high proportion of high severity fire and high rainfall had low probabilities of platypus occupancy. We also observed a decline in the predicted number of occupied sites when a site was classified as burnt at the watershed, or site level.

Regardless of whether a site was classified as burnt at the site or watershed scale, we observed similar patterns. We investigated these two resolutions because while a terrestrial site may be classified as burnt when the site itself experiences fire, an aquatic site may be impacted by fire in the surrounding watershed, due to downstream transport of debris, ash, and/or nutrients (Silva *et al.* 2020). A decline was observed in the predicted number of occupied sites for burnt site from pre-fire to post-fire 1. A larger increase or recovery was seen when a site was classified as burnt at the watershed scale, potentially demonstrating recovery at the landscape scale, but not at the site scale.

A second refined analysis, in which we included an interaction between the proportion of a watershed burnt at high severity and rainfall post-fire, led to more nuanced inferences regarding the impacts of the mega-fires on platypus occupancy. The nature of this interaction suggests that platypus occupancy was lowest at sites that were situated within a watershed that received high amounts of post-fire rainfall and that had a high proportion of their area burnt at high severity. High severity fire can kill vegetation, destabilise soil, and produce copious amounts of ash and debris, which, when followed by a high rainfall event, can be washed into waterways (Lyon and O'Connor 2008; Nyman *et al.* 2020; Neris *et al.* 2021). These events can produce sediment slugs and low oxygen concentrations, which can be fatal to fish and other fauna (Lyon and O'Connor 2008). While we do not have evidence that these events kill platypuses, there is evidence that they can negatively affect the abundance of their prey, such as aquatic macroinvertebrates (Verkaik *et al.* 2015). Direct mortality of platypuses seems unlikely, given that platypuses can be protected from fire by the burrows they use in consolidated earthen banks. These burrows are insulated from ambient air temperatures and can be extensive in size. For example Grant & Dawson, (1978) found that an artificially constructed burrow remained

between 14 and 18 °C when the outside air temperature was between -5 and 33 °C. Water temperatures can also increase during a fire, especially in small first or second order streams (Lyon and O'Connor 2008), but whether this would lead to platypus mortality is unknown. Perhaps the strongest evidence that our results reflect indirect effects of sedimentation and not mortality comes from the estimated interaction itself (Figure 4.8). In the absence of rain post-fire, probability of platypus occupancy showed little response to high severity fire. Regardless of the precise mechanism, our results suggest that rainfall following high severity fire is an important determinant of platypus distribution post-fire, as demonstrated by other studies of aquatic fauna.

There was also evidence, while not as strong, for a positive interaction between the proportion of a watershed burnt at low severity fire and rainfall post-fire. Low severity fire in combination with rainfall may not impact the occupancy of platypuses in the same way as the combination of high severity fire and rainfall, as environmental impacts may differ. Compared to high severity fires, lower severity fires are considered to have less of an impact on vegetation and soil erosion. Therefore, the hydrological impacts may be less. The positive impact on platypus occupancy probability may stem from increased rainfall providing more flows, which can be beneficial for platypuses (see Chapter 3).

We detected an effect of fire on platypus occurrence in this second analysis but failed to detect an effect of fire presence. The most plausible explanation is that because the fires were highly heterogeneous across the study region (44% was burnt at high severity, 47% was moderate to low severity, 9% was unburnt), the presence of fire was a coarse metric of potential impact when using model 1. This is particularly true at the watershed scale, as the presence of a single burnt pixel resulted in a watershed being classified as burnt. However, locally, there were very large patches of high severity fire, which could encompass an entire catchment (see (Collins *et al.* 2021)) or watershed. Our second analysis reflects this fact, incorporating a continuous measure of potential impact (proportion of watershed burnt) rather than a binary one (fire presence), and including fire severity (low vs high) and rainfall post-fire. As such, this refined analysis is a more robust test of ecological processes that have been postulated to impact aquatic fauna post-fire.

South-eastern Australia was experiencing the warmest and driest year on record prior to the 2019/2020 mega-fires (Abram *et al.* 2021), and there was a subsequent period of above average rainfall in most fire-affected catchments (Silva *et al.* 2020). These abnormal conditions may

help explain our observation that parameters post-fire 1 and post-fire 2 had a positive effect on platypus occupancy. First, drought conditions across much of the study area prior to the 2019/2020 mega-fires may have led to lower levels of platypus occupancy pre-fire. Droughts can reduce platypus prey abundance, reducing platypus foraging areas as well as habitat availability (Serena *et al.* 2014; Marchant and Grant 2015). Drying of waterways can also force platypuses to move overland, increasing predation risk (Bino *et al.* 2019). Furthermore, there may be synergistic effects from the combination of drought and fire. We did not explicitly investigate these effects, however, Bino *et al.* (2021) compared two rivers in New South Wales: one experienced wildfires during the 2019/2020 mega-fires, while both experienced extreme drought prior to the fires. They found that there were fewer platypuses at burnt sites, but that the drought could have impacted recruitment across both river catchments, as no juveniles were detected. Second, the post-fire increase in platypus occupancy we observed is likely due to drought-breaking rains that followed the 2019/2020 mega-fires (Silva *et al.* 2020). The habitat requirements of platypuses generally revolve around adequate flows, enabling foraging and movement (Milione and Harding 2009). Thus, the combination of pre-fire drought and post-fire rainfall is a plausible mechanism underpinning the observed temporal dynamics in platypus occupancy across the region.

One of the major strengths of our work is that we were able to conduct a Before-After Control-Impact (BACI) study design to investigate impacts of a large, natural disturbance. Indeed, pre-fire distribution data for many species was noted as a limiting factor in prioritising efforts after the 2019/2020 mega-fires (Legge *et al.* 2021; Southwell *et al.* 2021). More broadly, BACI study designs have been undertaken previously to investigate the effect of wildfire on species (e.g. Banks, Knight, McBurney, Blair, & Lindenmayer, (2011)), but to our knowledge, previous studies have not been undertaken across this broad of a spatial extent, nor for platypuses. BACI designs assume that the factors affecting the variables of interest are in a steady state between sampling periods, or across control/impact sites (Southwell *et al.* 2021), allowing for inferences to be made on the impact of a perturbation without interference. While these assumptions are difficult to prove, BACI designs remain a powerful tool to investigate the impact of disturbances, such as fire, on wildlife. The relatively large number of sites we were able to survey is another benefit of our study, as the statistical power to detect an impact increases with sample size (Wood 2021).

The primary reason that we were able to undertake a study of this size (both in terms of spatial extent and sample size) was our use of eDNA sampling. This technique provided an unprecedented ability to conduct both rapid and accurate assessments of platypus occupancy. Indeed, sites were able to be sampled in a matter of minutes, and previous work has shown that eDNA sampling is a highly sensitive technique for detecting platypuses, especially when compared to fyke netting (Lugg *et al.* 2017). Environmental DNA sampling is the only feasible method to extensively assess fire impacts on aquatic species over large areas, especially for widespread species, such as the platypus.

An additional benefit of using eDNA sampling is that it enabled us to leverage an extensive collection of pre-fire samples taken from across south-eastern Australia. Once eDNA has been extracted from samples, it can be used for other (unintended) purposes, as we have done here. For example, samples collected for a single species study, such as ours, could be used for eDNA metabarcoding, enabling post-fire assessments of entire taxonomic groups, such as vertebrates or fish. Alternatively, samples collected for one purpose could be used as baseline data for subsequent studies, such that occupancy data is available to understand impacts of future large fires or other disturbances, such as floods, land clearing, or climate change. This preparedness would allow for more targeted monitoring programs or emergency intervention actions implemented post-disturbance. While this is, to the best of our knowledge, the first study to use eDNA to assess the impacts of a major disturbance, we expect that the number of such studies will increase substantially over time, as eDNA sampling becomes more widely used.

With increasing evidence of climate change creating conditions that increases the duration, frequency, and size of wildfires (Abram *et al.* 2021), it is important to understand how biodiversity will be impacted. As we have demonstrated here, high severity fire, in combination with high rainfall events, could negatively impact platypus occupancy. While predictions of summer and autumn rainfall in south-eastern Australia under climate change are uncertain, extreme rainfall events are predicted to increase, whereby almost all climate models agree that the wettest day of the year will get wetter (Hope *et al.* 2018). An increase in the combination of large areas of high severity fire and high rainfall events could negatively impact not just platypuses, but other aquatic species, such as fish, crustaceans, and macroinvertebrates. Assessment of the impacts of fire and rainfall on a wider range of aquatic species is therefore a priority. Our study shows that eDNA sampling is a promising tool for this purpose.

Chapter 5
Using eDNA metabarcoding and multispecies occupancy models to investigate freshwater fish distributions over large spatial scales

5.1 Abstract

Freshwater systems are at risk globally and anthropogenic changes, such as river regulation and introduction of alien species, threaten many aquatic species. Traditional aquatic sampling methods are time-consuming and can fail to detect some species or life stages. Environmental DNA (eDNA) sampling can be conducted rapidly and at scale, providing an alternative approach for broad-scale monitoring of aquatic species. Here, we used eDNA metabarcoding and an environmentally stratified sampling design encompassing 237 sites across Victoria, Australia to understand responses of native and introduced fish species to environmental gradients. We used multi-species site occupancy-detection models to investigate species-level responses to environmental covariates, and to determine the extent to which the number of introduced species at a site impacted native species occupancy. We detected 49 fish and two lamprey taxa using a universal vertebrate primer pair. Distributions of native and introduced taxa responded similarly to water availability covariates (% of catchment covered by water, total runoff, and an index of zero-flow events). However, native and introduced taxa showed opposing responses to land use: occupancy probability of native taxa tended to increase with an increasing proportion of a contracted catchment covered by forest and agricultural land, whereas occupancy of introduced taxa tended to show the opposite trend. When restricted to common species detected more than 10 times, both native and introduced species predicted occupancy decreased with increased forest cover. The richness of introduced taxa at a site generally had a positive effect on native taxa occupancy, and the richness of native taxa generally had a positive effect on introduced taxa occupancy. This case study demonstrates that eDNA sampling can be used to conduct large-scale surveys of fish occupancy across environmental gradients. The methods used here can be implemented in other eDNA sampling programs to shed light on how the distributions of native and non-native taxa vary in response to global changes or disturbances.

5.2 Introduction

Freshwater habitats and their living inhabitants are imperilled globally (Dudgeon *et al.* 2006; Reid *et al.* 2019). Freshwater systems have a limited spatial extent and are subject to water extraction, regulation, and the continued impacts of a rapidly changing climate. As a result, freshwater systems are subject to both loss and fragmentation, which is a concern because there is often limited natural movement of taxa between catchments. Water quality is also declining in many regions due to anthropogenic changes (Lintermans *et al.* 2020). In addition, freshwater systems are increasingly occupied by invasive species (Lintermans *et al.* 2020). Australia hosts approximately 275 native freshwater fish species but very few of these are listed as threatened under national legislation. A recent study identified 22 species that are highly threatened, whereby over 90% of species were assessed to have a greater than 50% probability of extinction in the next 20 years (Lintermans *et al.* 2020). Invasive fish species are also considered a major threat to some threatened fish species, negatively impacting habitat quality and resulting in the homogenisation of communities across drainage basins (Olden *et al.* 2008). Invasive fish can also predate native fish, and are a particular threat for smaller fish, which in themselves are more threatened as a group (Lintermans *et al.* 2020).

Traditionally, fish populations are surveyed using time-consuming, labour-intensive, and invasive methods. For example, electrofishing or nets are often used to capture and morphologically identify species, which can potentially compromise animal health (Shaw *et al.* 2016). Some species may also be less likely to be detected using these methods, resulting in false negative detections. For example, the effectiveness of backpack electrofishing can vary among environments (Shepard *et al.* 2014) and some fish species are impossible to distinguish morphologically. These sampling techniques also require personnel experienced with the technique and with species identification, making broad-scale surveys extremely difficult, if not impossible, to undertake. In addition, traditional sampling methods pose health and safety concerns for those conducting the fish surveys —as surveyors typically need to enter the water body— as well as welfare concerns for the fish.

Environmental DNA (eDNA) sampling —the detection of intra- and extra- cellular DNA from an organism in an environmental sample— is increasingly being used for biodiversity or community-level surveys (Boivin-Delisle *et al.* 2021; West *et al.* 2021). This has been made possible by multi-species eDNA detection methods, such as eDNA metabarcoding.

Environmental DNA metabarcoding allows for an entire community to be targeted using a universal primer (e.g., targeting vertebrates or fish), enabling the collection of species-level presence/absence data across an entire group. Environmental DNA sampling has been shown to detect more species than electrofishing in a number of environments, including in freshwater streams of south-eastern Australia (McColl-Gausden *et al.* 2021) and South Australia (Shaw *et al.* 2016), and lakes in the United States (Sard *et al.* 2019). Environmental DNA sampling does not require capturing animals, reducing animal stress and impacts on individual health. Personnel also don't have to enter the water to collect eDNA samples, reducing health and safety concerns. Finally, eDNA samples can be taken with very little training. For example, citizen scientists can be sent kits to collect samples themselves (Biggs *et al.* 2015), standardising sample collection and increasing cost-efficiency. Some sites can also be sampled that are unsuitable for other methods such as electrofishing or trapping due to the environmental constraints on those methods. Thus, while animal capture will always be necessary for some studies (e.g., those on population health), eDNA metabarcoding provides a rapid, non-invasive, safe, and repeatable method to monitor the distributions of species in freshwater fish communities.

One of the greatest improvements in aquatic monitoring brought by eDNA sampling is that eDNA samples can be collected rapidly and at scale. Smaller sites can be sampled in a matter of minutes, enabling many sites to be visited in a short amount of time. Indeed, eDNA metabarcoding has recently been used to conduct surveys of aquatic communities across large spatial extents. For example, Pont *et al.* (2018) sampled fish along 524 km of river; West *et al.* (2021) surveyed bony fish, elasmobranchs and aquatic reptiles at 284 sites along 700 km of coastline in north west Australia; and Fraija-Fernández *et al.* (2020) collected marine samples for fish across an area of 120,000 km^2. The ability to conduct such surveys is important, as occupancy data across large spatial extents is needed to establish accurate estimates of species distributions, and to monitor for changes in those distributions over time (e.g., in response to disturbances or management actions). Given the rate and magnitude of anthropogenic influences on the environment, large scale and accurate survey methods are required to understand species' responses to environmental gradients, and to ensure anthropogenic impacts are documented and mitigated.

Site occupancy-detection models are an ideal way to analyse eDNA data and to understand species' responses to the environment. Hierarchical multi-species site occupancy-detection

models have been used to compare eDNA metabarcoding to electrofishing for freshwater fish (McColl-Gausden *et al.* 2021), and as a comparison to morphological identification of aquatic macroinvertebrates (Bush *et al.* 2020). These models have also been used to identify how detection errors can be accounted for when considering different experimental designs (Doi *et al.* 2019). Hierarchical multi-species site occupancy-detection models draw species-level parameters from a common statistical distribution. This allows data from a community to inform parameter estimates for individual species, including those with low occupancy and detection probabilities (Kéry and Royle 2016). Some fish species may be rarer than others in eDNA samples, and thus the ability to draw parameters from a community-level distribution is ideal for eDNA metabarcoding data. Effects of environmental covariates can also be included in hierarchical models at the species- and community-level, and these covariates can influence occupancy and/or detection. These models can also be used to inform management actions or interventions by providing data on factors which may influence a species' detectability or occupancy (Guillera-Arroita *et al.* 2010a; Shea *et al.* 2019).

In this study we demonstrate how eDNA metabarcoding can be used to study aquatic species' distributions rapidly and efficiently, and effects of environmental covariates on those distributions, using fish taxa across south-eastern Australia as a test case. We asked the following questions: a) how does the occupancy of native vs introduced fish species vary across land use gradients; and b) does the presence of introduced fish impact native fish occupancy, and vice versa?

5.3 Methods

Study area, eDNA sampling and site stratification

This study focused on data collected across the state of Victoria, Australia. Land cover in Victoria includes alpine regions, temperate forests, grasslands, agricultural industries, and urban areas. Contracted catchments (Commonwealth of Australia 2012) were chosen as sampling units as an attempt to account for the potential for movement of DNA within a river system (Pont *et al.* 2018). These catchments are derived from a nine second digital elevation model with a contracted node —a logical feature in the landscape such as the confluence of two major rivers or a significant monitoring location— at the outflow point (Commonwealth of Australia 2012). Only contracted catchments that were within a buffered 5 km area around major rivers (Commonwealth of Australia 2012) were considered. Environmental DNA

samples were collected between August 2018 and November 2019. Two independent water samples were taken at each site using two disposable syringes (60-mL Luer Lock sterile syringes; Hapool medical Technology, Shandong, China) and two detachable 0.22 µm filters (GP 22µm Filter Unit; Sterivex, EMD Millipore corporation, Billerica MA, United States). A maximum of 500 ml of water (collected using multiple draws from the 60-mL syringe) was pushed through each filter or until the filter was clogged. Filters were kept cool in a car fridge (approximately 3 ^{o}C) for up to three days before being transported to the laboratory and stored at -20 ^{o}C until DNA could be extracted.

A subset of sites from a larger single-species eDNA dataset (Chapter 3) was selected for use in this study. These original samples were stratified across seven land use categories and runoff categories. The seven land use categories were condensed from the Dynamic Land Cover dataset (Geoscience Australia, 2017): mines and urban areas, waterbodies, pasture and cropping (agriculture), grasslands and shrubs, and forested. Runoff (as a proxy for water availability) across each contracted catchment was calculated for 2018-2019 from the Australian Water Resources Assessment Landscape model (Frost *et al.* 2018) and an index of zero-flow annually was provided by the Victorian Department of Environment, Land, Water and Planning. A total of 179 sites from the larger dataset were selected. In addition, 58 sites from around the urban centre of Melbourne were also included. These Melbourne samples were collected in November and December 2020 and used the same field sampling methods as above. In total, 237 samples were included for this study across 115 contracted catchments (Figure 5.1).

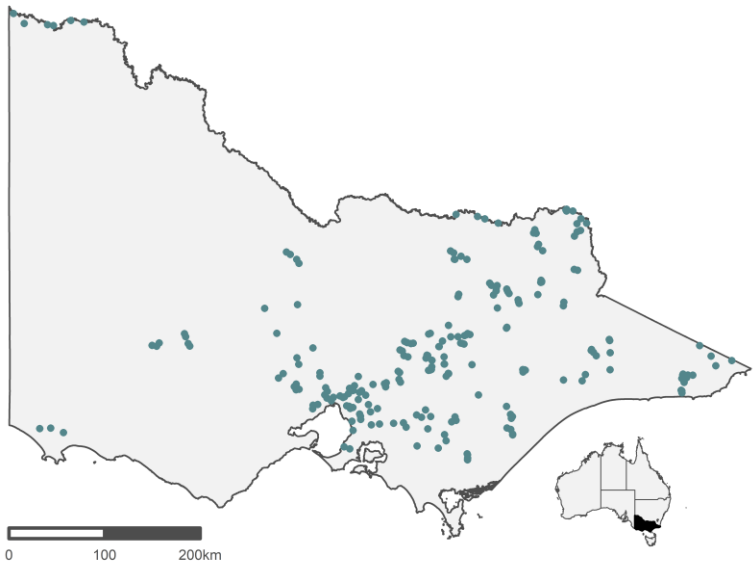

Figure 5.1*: Map of selected sites (blue dots) including samples from a larger-scale single-species dataset, and additional samples collected around Melbourne, Victoria, Australia.*

Environmental covariates

Seven land use categories and total runoff were calculated at the contracted catchment level, whereas the zero-flow index was calculated at the site level. In the zero-flow index dataset, a higher index value corresponded to less zero-flow days annually compared to other river systems in Victoria. Annual zero-flow data were in point form and therefore the nearest flow data point to the sampling location along the river network was obtained using QGIS (QGIS Geographic Information System 2021).

DNA extraction

For the samples collected for the large single-species eDNA dataset (Chapter 3), Qiagen DNeasy Blood & Tissue Kits (Spin-column protocol) were used to extract DNA from filters in a room dedicated to low-quantity DNA sources with qPCR setup undertaken in a laminar flow hood. A 540 µL of ATL buffer and 40 µL of proteinase K in each filter unit was used and then each filter was sealed and incubated for 3 hours at 56 °C with constant agitation. The lysis solution was transferred into new 2 mL tubes. The rest of the DNA extraction steps followed manufacturers protocol with some minor adjustments; in each sample, 500 µL AL buffer, 500 µL ethanol, and a final elution step of 100 µL AE buffer was used. A negative DNA extraction control (that replicated the process using a sterile Sterivex filter) was included in every batch of DNA extractions (~24 samples). There were 16 negative extraction controls. For the

additional samples collected around Melbourne, samples were extracted on the QIAcube HT, using a QIAamp 96 DNA QIAcube HT Kit. 540 µL of ATL buffer and 40 µL of proteinase K was added to each filter unit and then each filter sealed and incubated for 3 hours at 56 °C with constant agitation. After incubation, 200uL of this lysate was transferred to an S-Block for use on the QIAcube HT. Subsequent steps followed the manufacturer's instructions, excluding the use of TopElute (an optional step) and with samples eluted in 100uL of Buffer AE. One negative control was used per 96-well plate for a total of two negative controls.

Metabarcoding methods

In this study a universal primer targeting all vertebrates was used (Riaz *et al.* 2011; Shehzad *et al.* 2012). The assay marker region was located within the mitochondrial 12S ribosomal RNA (rRNA) gene and was ~100 bp. The two water sample replicates were analysed separately. To prepare metabarcoding libraries, a two-step PCR protocol was used, with primer sequences listed in Appendix E Table E.1. The first round PCR primers contained the marker-specific primer sequence, a (5′–3′) universal adaptor sequence and a 0- to 6-bp heterogeneity spacer (Fadrosh *et al.* 2014). Each reaction contained 2 µl eDNA, 1× PCR buffer, 3 mM MgCl2, 0.3 µM of each primer and 0.2 Units KAPA Plant DNA Polymerase (Kapa Biosystems). Reaction conditions were as follows; one cycle at 95 °C for 3 min, 40 cycles at 95°C for 20 s, 57°C for 15 s, 72 °C for 15 s; and one cycle at 72 °C for 1 min. For each sample, four PCR technical replicates were performed, one for each heterogeneity spacer combination with samples randomly allocated across PCR plates. 17 DNA extraction negative controls were used, and 96 PCR replicates were included and carried through to sequencing. Three eDNA samples were included from the Northern Territory (Australia) from a previous project where the fish community had already been determined using the same vertebrate markers to act as a known "mock community" and positive control.

The four technical replicates were pooled following the first-round PCR and purified with ExoSAP-IT PCR Product Cleanup Reagent (ThermoFisher). Second-round PCR primers contained (5′–3′) the Illumina p5 or p7 binding region, an 8-bp index sequence (Fadrosh *et al.* 2014) and a universal adaptor sequence. For each sample, unique forward and reverse index combinations were used. Reactions contained 2 µl purified PCR product, 0.5 µM of each primer and 1× Phusion Hot Start II High-Fidelity PCR Master Mix (Thermo Fisher Scientific). Reaction conditions were as follows; one cycle at 98 °C for 1 min; 15 cycles at 98 °C for 10 s, 72°C for 45 s; and one cycle at 72°C for 10 min. For the samples collected for the larger single-

species study (Chapter 3), sequencing was performed at the Monash Health Translation Precinct (MHTP) Medical Genomics Facility on an Illumina MiSeq platform (Illumina) using 150 bp PE chemistry. For the additional samples collected around Melbourne sequencing was performed on an iSeq 100 (Illumina) also using 150 bp PE chemistry.

Bioinformatics

Bioinformatic analyses were performed as in McColl-Gausden *et al.* (2021) using a custom analysis pipeline where a custom reference database for taxonomy assignment was built for vertebrate species known to occur in Victoria, Australia. To retrieve all 12S rDNA sequences present in GenBank (https://www.ncbi.nlm.nih.gov/genbank/), the R package RENTREZ v1.1.0 (Winter 2017) was used. The Genbank sequences were also supplemented by sequences from species of interest in our tissue collection, by Sanger sequencing the 12S region. An in-silico PCR undertaken to retrieve the appropriate gene region for reference species was performed with the primer sequences listed in Table S1 and the software packages OBITOOLS (Boyer *et al.* 2016) and ecoPCR (Ficetola *et al.* 2010).

After the merging of R1 and R2 paired-end sequences (using VSEARCH's fastq-mergepairs function), primer sequences were trimmed and sequencing reads dereplicated where those with an abundance of less than 10 were excluded from further analyses. Sequences were clustered into Operational Taxonomic Units (OTUs) using a pairwise identity of 100% (vsearch --id 1.0). By comparing against the custom vertebrate reference sequence database, taxonomic assignment was performed with VSEARCH (Rognes *et al.* 2016) whereby each OTU was assigned a taxonomic identity using a minimum bootstrap support of 95% (vsearch –sintax-cutoff 0.95). Filtering of samples was set at 0.1%.

Model description

The data was binary detection/non-detection observations of species *i* (1,2…n), at site *j* (1,2…n) in water sample *k* (1,2). A hierarchical multi-species site occupancy-detection model was used to estimate occupancy and detection probabilities for each species.

Equation 1 $$z_{ij} \sim Bernoulli(\psi_{ij})$$

Equation 2 $$logit(\psi_{ij}) = a_i + \beta_i * cov_j + \gamma_{catchment}$$

Equation 3 $$y_{ijk}|z_{ik} \sim Bernoulli(z_{ij}\omega_{ijk})$$

Equation 4 $$\psi_i \sim Normal(\mu_{\psi_{[g]}}, {\sigma^2}_{\psi_{[g]}})$$

Equation 5 $$\omega_i \sim Normal(\mu_{\omega_{[g]}}, {\sigma^2}_{\omega_{[g]}})$$

Equation 6 $$a_i \sim Normal(\mu_{a_{[g]}}, {\sigma^2}_{a_{[g]}})$$

Equation 7 $$\beta_i \sim Normal(\mu_{\beta_{g[i]}}, {\sigma^2}_{\beta_{g[i]}})$$

Equation 1 describes the occupancy process as the outcome of a Bernoulli trial. z_{ij} is a latent variable describing the presence or absence of taxa *i* at site *j* given the probability of occupancy for taxa *i* at site *j*. $z_{ij} = 1$ where the taxa is present and $z_{ij} = 0$ where the taxa is absent.

Equation 2 describes the taxa and site-level heterogeneity in occupancy (ψ_{ij}) and includes an overall intercept for each taxa (a_i) and effects of covariates ($\beta_i * cov_j$), where cov_j are the environmental covariates at each site and β_i is the posterior parameter estimate. This equation also incorporates a random effect for contracted catchment ($\gamma_{catchment}$).

Equation 3 describes the detection process, in which y_{ijk} denotes the detection ($y_{ijk} = 1$) or non-detection ($y_{ijk} = 0$) of taxa *i* at site *j* in replicate *k*. This is a function of the occurrence of a taxa (z_{ij}) and the probability of detection (ω_{ijk}).

Equations 4-7 describe the prior distributions from which taxa-level responses are drawn. These taxa-level responses are indexed by a grouping variable *g*. This grouping variable was used in the first analysis only. The first analysis used all available data, or a restricted dataset in which taxa were detected at least 10 times, and group was defined as "native taxa" or "introduced taxa". For the second analysis, data were restricted to only native taxa or only introduced taxa, and taxa richness of the opposite group was included as a covariate. For example, using the dataset where only native taxa were used, the number of introduced taxa present at a site was included as an additional covariate.

Taxa-specific prior distributions were specified on the logit scale, where ψ_{ij}, ω_{ijk}, a_i and β_i were normally distributed with mean μ and precision τ. Hyper-priors for μ_ψ, μ_ω, μ_α and μ_β were specified as uniformly distributed across the range of 0–1. Hyper-priors for σ^2 were specified as uniformly distributed across the range 0–5.

Models were fit in a Bayesian framework using R version 3.6.3 (R Development Core Team, 2020) and the software JAGS through rjags v.4-8. Three Markov chains were run for 300,000 iterations each. We discarded the first 100,000 iterations (burn-in), and the remaining samples were thinned by a factor of 100, resulting in 2000 samples per chain from the posterior distribution. Convergence was measured visually using traceplots from the jagsUI package v.1.5.0, which indicated that chains were well mixed and Gelman-Rubin statistic values $\hat{R}$ were below 1.1, indicating successful convergence of chains. Model code is provided in Appendix E.

5.4 Results

Three taxa were removed from the analysis (*Aldrichetta forsteri*, *Acanthopagrus butcheri* and *Tetractenos sp*.). These taxa were present in the mock community for the Melbourne sites and in study samples (but not in the negative controls) and are marine or estuarine taxa not known to occur in the areas where Melbourne samples were taken. This may represent contamination from the mock community in the processing of samples from around Melbourne only. *Macquaria colonorum* (which may share haplotypes with *Macquaria novemaculeata*) and *Gadopsis sp*. were included in the analysis even though they were present in the mock community, as they could plausibly occur in the regions sampled. Nonetheless, results for these taxa should be treated with caution.

Across 237 sites, we detected 49 fish and two lamprey taxa. Of these taxa, 16 were introduced, and 35 were native. The minimum number of taxa detected at a site was 0 and the maximum was 12, with a mean of 3.6. The mean number of sites at which a taxa was detected was 16.6 but 27 taxa were detected less than five times.

To investigate the impact of environmental covariates on the occupancy of native and introduced taxa, we predicted the mean occupancy of each group across the full range of possible values of each covariate. We used the full dataset (number of taxa = 51) (Figure 5.2), as well as a dataset restricted to taxa with at least 10 detections (number of taxa = 19) (Figure 5.3). For the full dataset, effects of two covariates on predicted mean occupancy differed between native and introduced taxa: the proportion of a contracted catchment covered by forest (Figure 5.2A), and the proportion covered by agriculture (Figure 5.2B). Predicted mean occupancy of native taxa increased with increasing forest or agricultural cover, whereas occupancy of introduced taxa showed the opposite pattern. The mean effect of the variable

describing the proportion of a contracted catchment covered by urban areas was relatively flat for both taxa groups but there was a high degree of interspecific variation (Figure 5.2C). The mean effect of the covariate describing the proportion of a contracted catchment covered by grasslands and shrubs was slightly negative for introduced taxa and relatively flat for native taxa (Figure 5.2D). Mean occupancy of native and introduced taxa increased slightly with an increase in the proportion of a catchment classified as water, although the slope of this relationship was marginally stronger for native taxa (Figure 5.2E). Finally, effects of the two flow-based covariates (zero-flow and runoff) were the most consistent across the native and introduced taxa groups, whereby there was a relatively flat relationship between each covariate and mean predicted occupancy (Figure 5.2F and Figure 5.2G). See Appendix E for coefficient plots for each environmental covariate and taxa (Figures E.1-E.7).

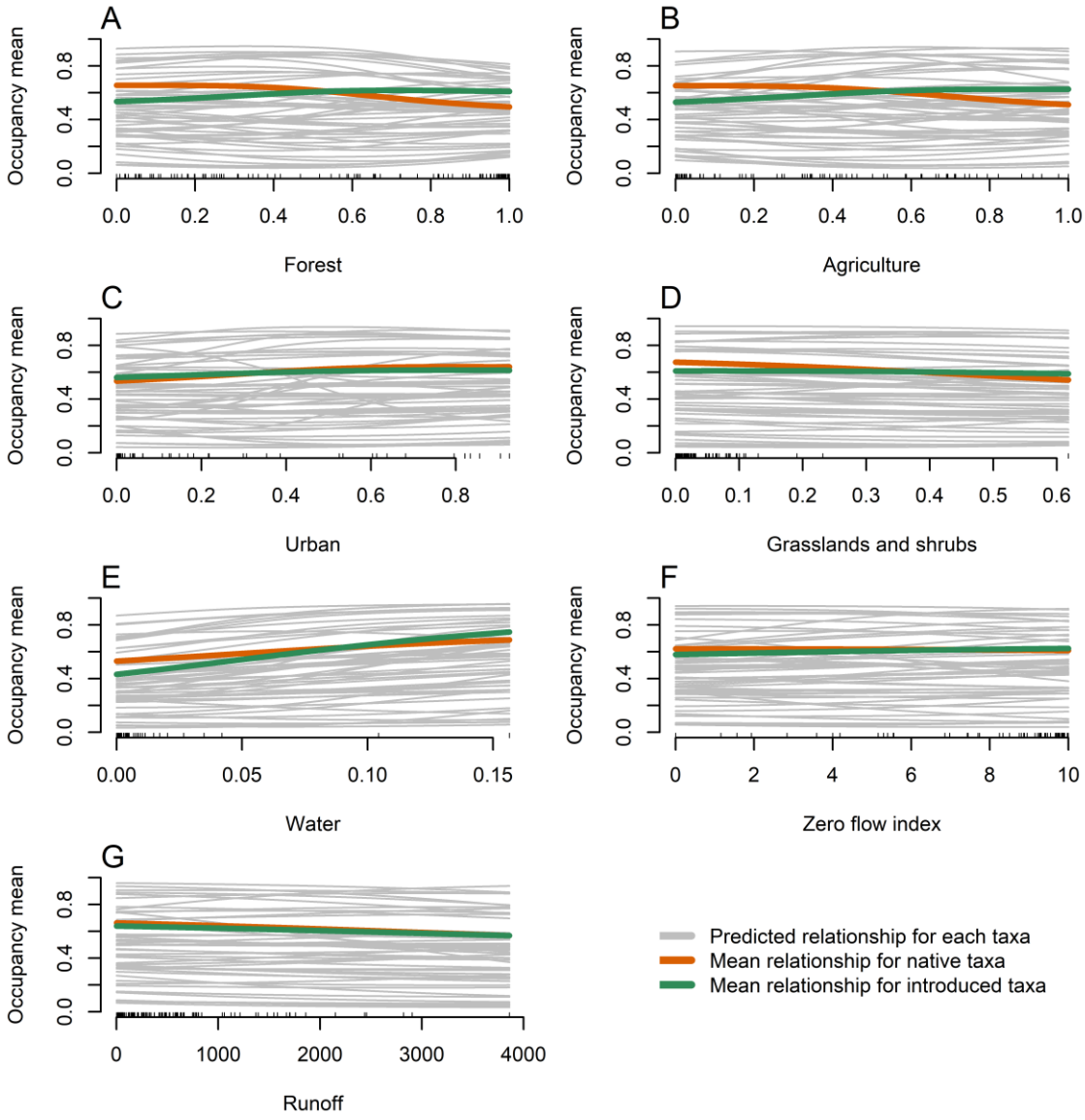

Figure 5.2*. Predicted relationships between occupancy probability for each taxa and environmental predictors using the full dataset (51 taxa: grey lines) The green and orange lines show the community posterior mean of these relationships for native and introduced species respectively.*

For the dataset restricted to taxa with at least 10 detections, general patterns were similar to the full dataset, except for responses to forest cover (Figure 5.3). Mean predicted occupancy for native and introduced taxa declined with an increase in the proportion of a contracted catchment covered by forest when using this smaller dataset (Figure 5.3A). The occupancy of both groups increased with an increase in agriculture, urban, water and the zero-flow index, but with the slopes for the groups differing across these covariates (Figure 5.3B, Figure 5.3C, Figure 5.3E and Figure 5.3F). Occupancy for both groups declined marginally with an increase in grasslands and shrubs, as well as with runoff (Figure 5.3D and Figure 5.3G).

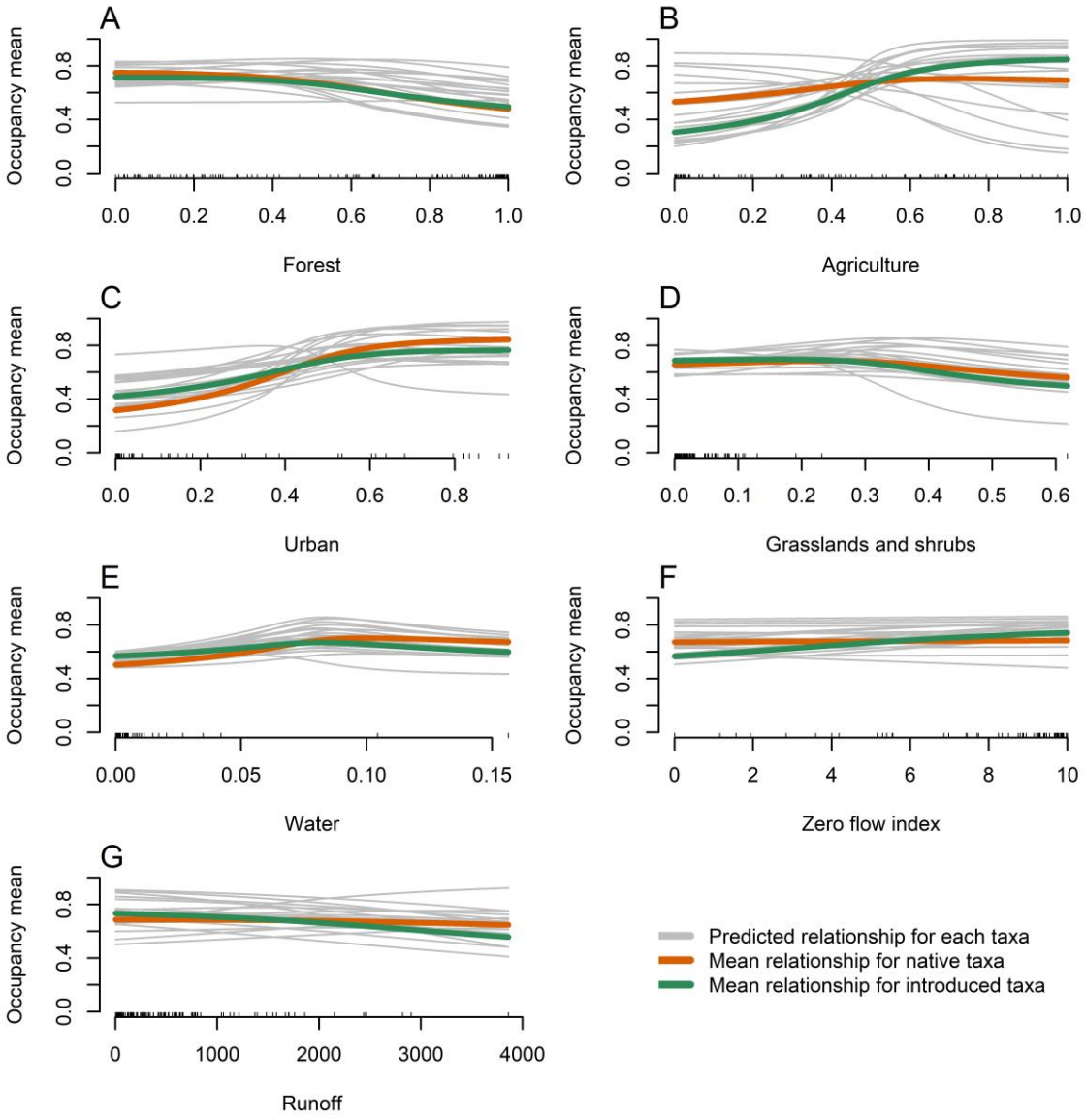

***Figure 5.3**: Predicted relationships between occupancy probability for each taxa and environmental predictors using the dataset restricted to taxa with more than 10 detections (19 taxa: grey lines) The green and orange lines show the community posterior mean of these relationships for native and introduced taxa, respectively.*

We used fitted models and the posterior predictive distribution of local species richness to predict richness for native and introduced taxa separately at the contracted catchment level across geographical space (Figure 5.4 and Figure 5.5). Species richness estimates were obtained from model estimates (sampled posterior distribution of presence absence (z_{ij}) for each species at each site; JAGS code provided in Appendix E) and standard deviation plots can also be found in Appendix E, Figures E.8 and E.9. Native and introduced richness were predicted to be higher around the urban centre of Melbourne. Introduced species richness was predicted to be lower around the cooler mountainous region of eastern Victoria and the Otway region to the south-west of Melbourne. When considering only common species detected 10 or more times, native species richness was lower in north-western Victoria, around the urban centre of Melbourne and in the more mountainous region to the north-east of Melbourne (Appendix E Figure E.10.). For more common introduced species, a similar pattern emerged where predicted richness was lower in the cooler, more mountainous regions of Victoria (Appendix E Figure E.11.).

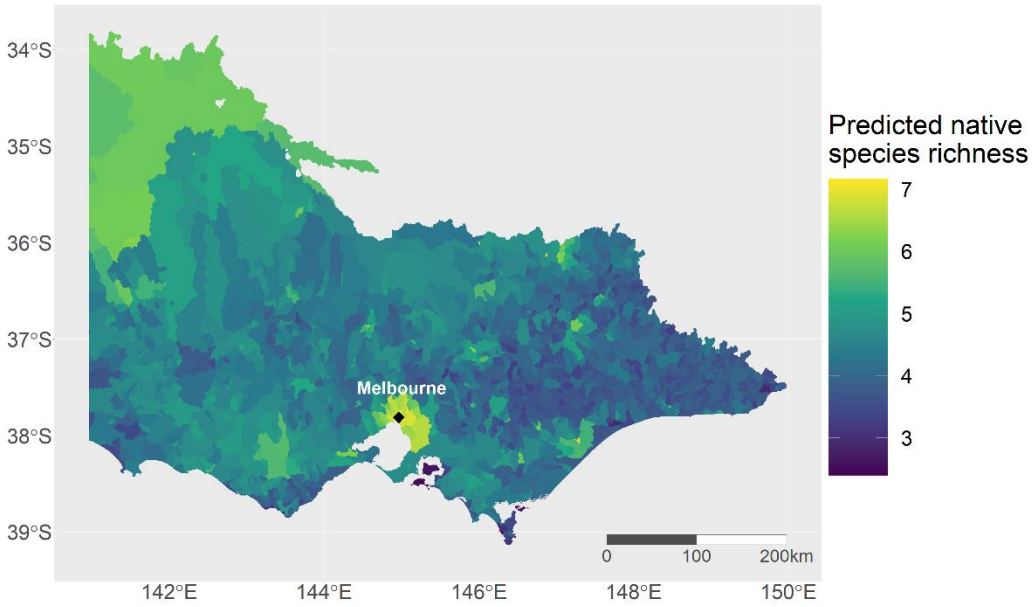

***Figure 5.4**: Predicted native fish species richness at the contracted catchment scale based on the community occupancy model presented above. The posterior mean estimate of species richness is used.*

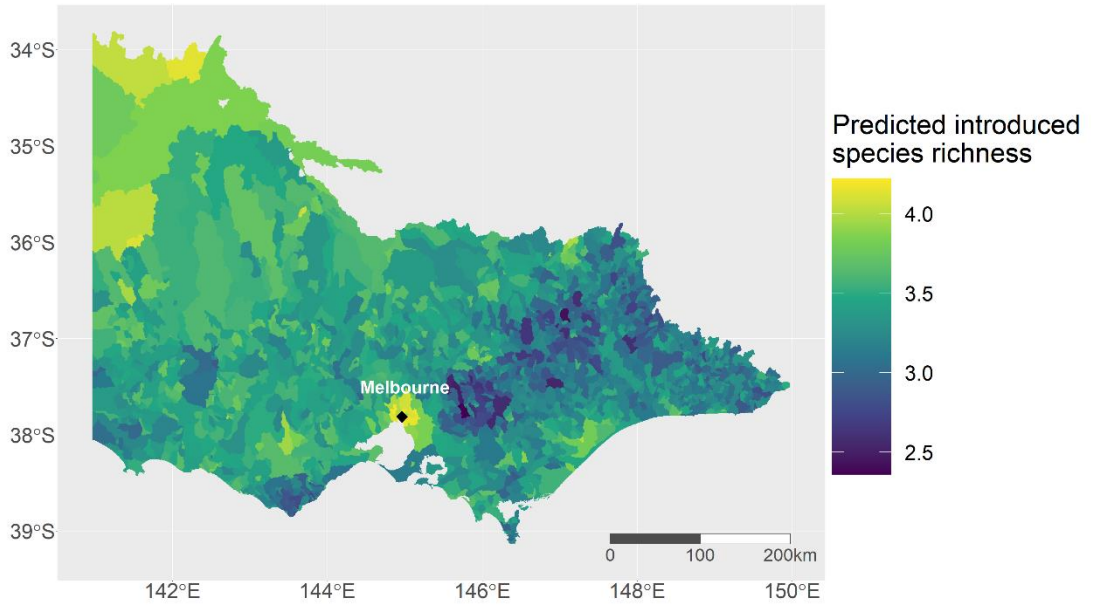

***Figure 5.5**: Predicted introduced fish species richness at the contracted catchment scale based on the community occupancy model presented above. The posterior mean estimate of species richness is used.*

In the second analysis, detections of only native taxa or only introduced taxa were considered as responses. In the case of the former, the number of introduced taxa at a site (introduced taxa richness) was used as an additional covariate to investigate the extent to which the presence of introduced taxa impacted the occupancy of native taxa. This analysis was then repeated for introduced taxa, whereby native taxa richness was used as a covariate. When the eDNA detection data was restricted to natives only, the occupancy probability of all 35 taxa increased with introduced taxa richness at a site (Figure 5.6). However, only 9/35 taxa —*Retropinna semoni, Philypnodon grandiceps, Percalates novemaculeata, Nannoperca sp., Macquaria sp., Maccullochella peelii, Galaxias sp., Gadopsis sp.* and *Anguilla australis*— had 95% credible intervals that did not overlap zero.

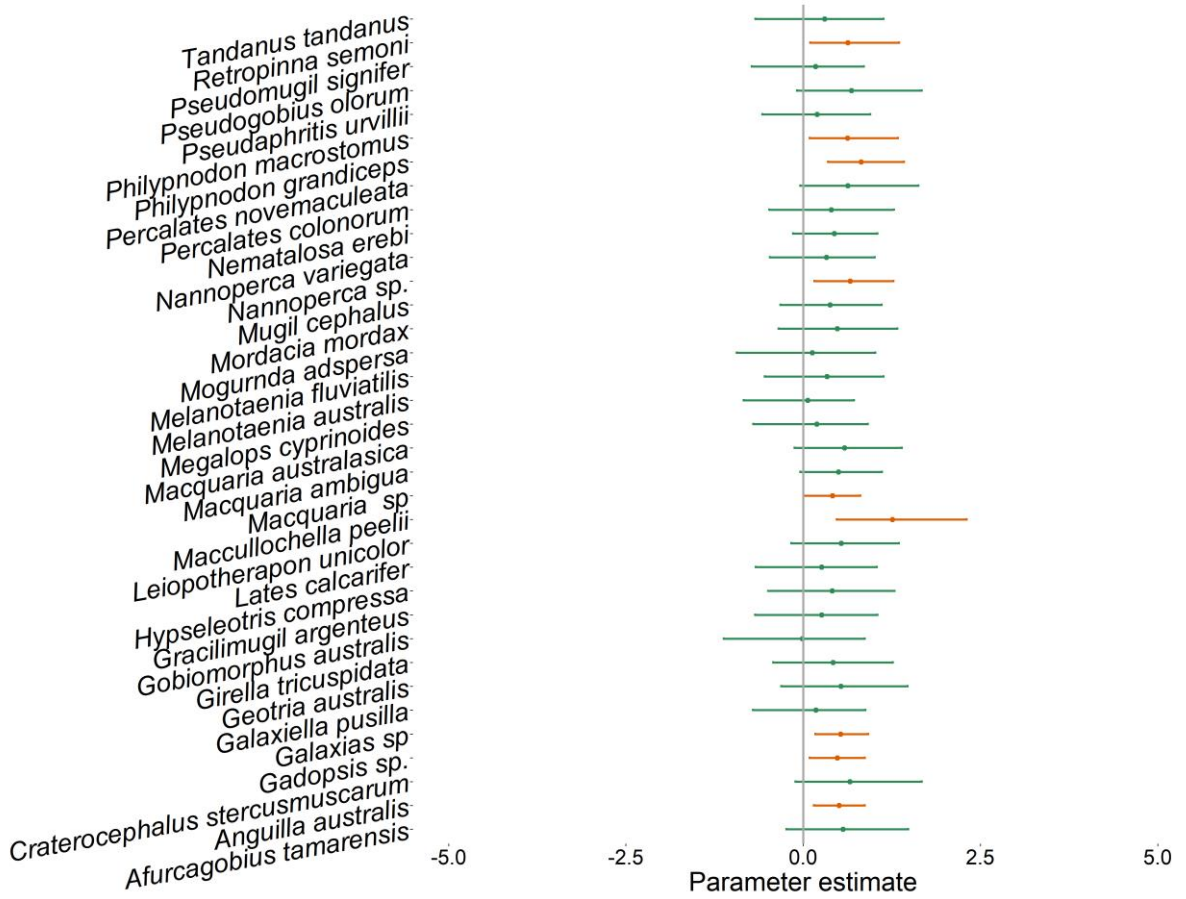

***Figure 5.6**: β coefficients describing effect of introduced taxa richness on native taxa occupancy. Points are the mean and lines are the 95% credible intervals. Orange lines do not overlap zero and green lines overlap zero.*

When data was restricted to introduced taxa only, we found a similar pattern: the occupancy probability of 15/16 introduced taxa increased with native taxa richness at a site. However, there was more uncertainty in this effect (as indicated by much wider 95% credible intervals) than was the case for native species occupancy. Nonetheless, the occupancy probability of 10/16 introduced taxa —*Tinca tinca, Phoxinus phoxinus, Oncorhynchus tshawytscha, Oncorhynchus sp., Misgurnus anguillicaudatus, Maulisia mauli, Gambusia holbrooki, Cyprinus carpio, Carassius auratus and Anguilla japonica*— increased with introduced taxa richness at a site (mean β estimates were positive and 95% credible intervals did not overlap zero: Figure 5.7). The occupancy of one species —*Salmo trutta*— displayed a negative response to native taxa richness. Coefficient plots using data where species were detected more than 10 times showed similar patterns and are presented in Appendix E Figures E.12 and E.13.

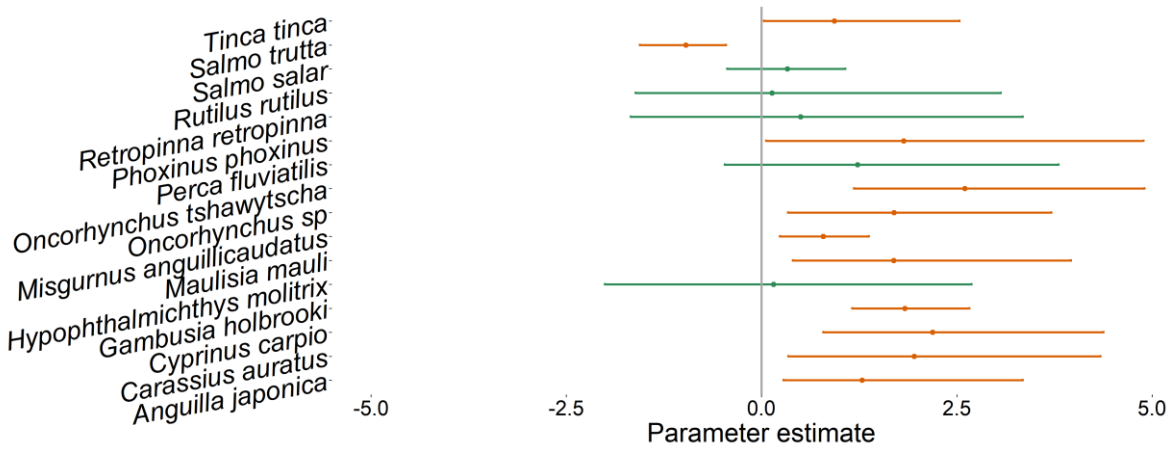

Figure 5.7*: β coefficients describing effect of native taxa richness on introduced taxa occupancy. Points are the mean and lines are the 95% credible intervals. Orange lines do not overlap zero and green lines overlap zero.*

5.5 Discussion

The distribution and composition of aquatic communities are important facets of conservation management. We demonstrate that eDNA sampling is a valid alternative to traditional sampling methods for studying aquatic communities. Environmental DNA sampling enables rapid and efficient collection of data on species occupancy across large areas for the purposes of determining species distributions, community composition, and the effects of environmental gradients on these patterns. Below, we discuss the ecological patterns uncovered by our case study, before using our case study to demonstrate the benefits of eDNA sampling for identifying fish communities and the potential for large scale, routine monitoring programs.

Case study: eDNA sampling for fish in Victoria, Australia

Large and extensive datasets on species detections across broad spatial extents are invaluable for estimating species' contemporary distributions and their responses to environmental gradients. Such data is also needed to study how species' distributions change over time and in response to environmental triggers. Fish, in particular, are sensitive to environmental changes (e.g., in flow regimes or river condition), and thus are useful indicators of river health (Evans *et al.* 2016). Here we collected baseline data on the distributions of 51 fish and lamprey taxa from 217 sites across Victoria, Australia, encompassing an area (227,444 km^2) larger than the size of Great Britain. Importantly, sites were stratified by land use and flow regimes to ensure

representation across environmental gradients. This stratified sampling design enabled us to model correlations between the occupancy of individual fish taxa and surrounding land use, and to investigate whether these correlations differed between native and introduced taxa.

Our analyses suggest that for some native and introduced taxa, water availability metrics (the proportion of a contracted catchment covered by water, zero-flow index, and runoff) were important predictors of occupancy. However, using the full dataset of all taxa, we found that both groups had relatively flat responses to increases in the zero-flow index and runoff at the community level (Figure 5.2). This result was unexpected, as all fish species, whether they are native or introduced, require suitable water availability or appropriate flow regimes (Morrongiello *et al.* 2011). This result may be due to a lack of representation for some taxa across the studied environmental gradients, with 27 taxa being detected less than five times. Indeed, when we used the same model on the more common species (those that were detected at least 10 times; 19 taxa), we found an increase in the community-level occupancy of native and introduced taxa with an increase in the zero-flow index (less zero-flow days) although there were negligible changes in predicted responses to runoff (Figure 5.3). Interestingly, the predicted increase in occupancy probability with less zero-flow days was more pronounced for native taxa, highlighting the impact that flow regime change, drought and water extractions could have on native species, especially under future climatic conditions (Morrongiello *et al.* 2011). Predictions of water availability in the future are uncertain but some predictions indicate a substantial decrease in spring and winter rain in eastern Victoria, western Victoria and the Murray basin (Hope *et al.* 2018).

More predictably, occupancy of both native and introduced taxa increased with an increase in the proportion of a contracted catchment covered by water, with native taxa responding slightly more positively compared to introduced taxa. This also held true for the restricted dataset containing only common species. The spatial distribution of native taxa also seemed to be influenced strongly by water availability (Figure 5.4). Native species richness was predicted to be higher in areas with higher water availability, such as the urban centre of Melbourne; the lakes 130 km south-west of Melbourne; the Gippsland Lakes in eastern Victoria, 200 km east of Melbourne; and along the Murray River, with its billabongs and wetlands, 450 km north-west of Melbourne. These areas could provide a wide variety of habitats, enabling multiple species to co-exist (Sutela *et al.* 2020).

Responses to the proportion of a contracted catchment covered by forest varied between the full dataset (Figure 5.2) and the restricted dataset including only common taxa (Figure 5.3). For the full dataset, even within the native and introduced groups, there was a range of positive and negative responses to the proportion of forest cover. However, changes in the predicted community-mean occupancy probability for both groups were relatively linear, with the community-mean response for native fish showing a slight increase in occupancy probability with an increase in forest cover, with introduced taxa showing the opposite pattern (Figure 5.2). However, when restricted to only common taxa, a slightly different pattern emerged (Figure 5.3), whereby predicted community-mean occupancy of native and introduced taxa declined with increasing forest cover.

This pattern of decreasing occupancy probability with forest cover can also be seen in the spatial predictions of taxa richness, especially predictions for introduced taxa. Relatively low taxa richness was predicted in the cooler, more forested areas of the mountainous region 100 km north-east of Melbourne, and in the Otway region 120 km to the south-west (Figure 5.4 and Figure 5.5). Less species are generally expected to occupy headwater streams due to conditions at these sites such as fast running, shallow and narrow streams (Sutela *et al.* 2020). Another explanation for the negative response of taxa to forest cover is the presence of self-sustaining populations of introduced brown trout (*Salmo trutta*) in high elevation or cooler areas (Jarvis *et al.* 2019). Brown trout were introduced into Australia in 1864 by acclimatization societies and have been shown to outcompete or prey upon other native and introduced species (Jarvis *et al.* 2019). These interactions could result in lower species richness in forested areas, as brown trout, being a large-bodied fish, dominate these environments (Cadwallader 1996).

Another line of evidence for the effect that brown trout have on the distributions of other species is our finding that native taxa richness had a significant negative impact on brown trout. This impact does not indicate causation; we cannot infer that native taxa negatively impact the occupancy of brown trout. Rather, this finding likely indicates that brown trout are less likely to occupy sites where there is high native richness. Brown trout are dominant in many cool water areas of the state, and therefore are more likely to be found at sites with lower species richness, as trout reduce species richness by outcompeting and predating on other fish (Jarvis *et al.* 2019). Brown trout, along with other introduced species —rainbow trout (*Oncorhynchus mykiss*), chinook salmon (*Oncorhynchus tshawytscha*) and brook trout (*Salvelinus fontinalis*)— and native species —Murray cod (*Maccullochella peelii*), golden perch

(*Macquaria ambigua*), Australian bass (*Macquaria novemaculeata*), silver perch (*Bidyanus bidyanus*), estuary perch (*Macquaria colonorum*) and Macquarie perch (*Macquaria australasica)*— are also stocked in large numbers throughout Victoria (Jarvis *et al.* 2019).

The occupancy probabilities of both native and introduced taxa increased or remained stable with an increase in the proportion of a contracted catchment covered by a form of agriculture, with a more pronounced positive effect seen for native taxa in both datasets. This result was unexpected, as intensive agriculture is generally reported to have negative impacts on native species due to alterations in flow regimes and native habitat (Koehn *et al.* 2020). However, south-eastern Australia continues to host a large number of native fish species, despite extensive land use changes, including agriculture, urbanisation, and forestry (Morrongiello *et al.* 2011). Dividing agriculture into the different forms that the agricultural industry takes would allow for a more nuanced perspective of the impact that agriculture has on fish occupancy. For example, modelling the proportion of a contracted catchment that is covered by pasture and cropping separately.

Somewhat surprisingly, the proportion of urban environment in a contracted catchment had a positive effect on both native and introduced species in both datasets. When restricted to more common species with more than 10 detections, the positive response was even more pronounced. This suggests these more common species may be generalists, and able to thrive in more altered landscapes (Marques *et al.* 2021). Supporting this observation, species richness for native and introduced species was predicted to be high around the urban centre of Melbourne (Figure 5.4 and Figure 5.5).

We did not find any significant negative impact of introduced fish richness on native fish occupancy, counter to our expectations (Figure 5.6). Previous research suggests that some native Australian species are negatively impacted by invasive fish through processes such as competition and predation. For example, the endangered species *Galaxiella pusilla* is negatively impacted by at least three invasive fish species —*Perca fluviatilis*, *Cyprinus carpio* and *Gambusia holbrooki*— through predation, habitat degradation, and competition, respectively (Coleman *et al.* 2019). In contrast, we found many positive correlations between the occupancy of native fish and the richness of introduced taxa at a site. Likewise, the occupancy of introduced fish was often positively correlated with native fish richness (Figure 5.7). This could be because site occupancy, for many taxa, is predominately influenced by the environment or microhabitats, rather than other fish species. That is, the richness of taxa at a

site may be a proxy for the abiotic suitability of a site (Tingley *et al.* 2011; Souza *et al.* 2011). For example, sites with many different microhabitats, which may promote site occupancy of some species, also tend to host a large array of different fish species. A future consideration, which may help tease apart these species-specific interactions, is the use of Joint Species Distribution Models, JSDMs (Wilkinson *et al.* 2021). JSDMs model the response of each species to the environment, and residual correlations between species can be used to infer species interactions or missing environmental variables (Zurell *et al.* 2018).

eDNA sampling for species composition across large spatial scales

Our case study demonstrates that eDNA metabarcoding provides an efficient means to standardise routine monitoring programs. Large and extensive datasets on fish detections can take years to collate with traditional sampling methods, as sites cannot be sampled frequently and extensively without a large amount of effort (Zajicek and Wolter 2018). What's more, some sites are inaccessible with traditional sampling methods because of equipment requirements. In contrast, eDNA sampling imposes far fewer sampling restrictions and can be deployed rapidly, providing a more comprehensive picture of community composition in different environments in a shorter period of time.

Our study illustrates one of the major benefits of using eDNA metabarcoding rather than single-species eDNA detection methods: a wide range of species can be identified from water samples —e.g., mammals (Lyet *et al.* 2021) or fish (West *et al.* 2021)— using a single consistent method. This makes large-scale surveys, such as those conducted here, not only possible but also highly efficient and systematic.

Environmental DNA sampling is particularly useful for the detection of rare, cryptic, or low abundance species, or individuals at different life stages that may not be routinely detected during surveys (Deiner *et al.* 2017). Indeed, many of the species we detected were rare, with only 19/52 species being detected 10 or more times. Because of potentially low detection probabilities of some species using traditional sampling methods, indicator species are sometimes used to represent the health of a system (Ji *et al.* 2013). Using a method such as eDNA sampling combined with metabarcoding, in which the probability of detecting all or most species at a site is much higher, reduces the need for reliance on indicator species, as the entire community can be captured in one or more water samples. As such, eDNA

metabarcoding can be used to routinely monitor species composition at a set of sites to assess changes over time, including the detection/non-detection of introduced or threatened species. When used in this way, routine eDNA metabarcoding could be used as an early warning signal for potentially undesirable changes in community composition, prompting the deployment of other sampling methods, such as trapping or electrofishing, to verify results and assess the extent of change. For example, four species that are listed as threatened or endangered in Victoria were detected in this study (*Galaxiella pusilla*, *Nannoperca variegata*, *Maccullochella peelii* and *Macquaria australasica*). Environmental DNA detections of these species in previously unknown locations may trigger more extensive traditional surveys to locate new populations so that conservation actions can be implemented, or so that more detailed data can be collected (e.g., abundance, health status, age class, breeding status, genetic samples). Similarly, eDNA metabarcoding could be used to monitor for the encroachment of invasive species, or to evaluate the effectiveness or longevity of a translocation (Hempel *et al.* 2020). Thus, eDNA metabarcoding shows great promise for the routine monitoring of aquatic fauna.

5.6 Conclusions

We demonstrate that eDNA sampling can be used to investigate broad-scale patterns of occupancy in fish communities using a systematic study design and multi-species site occupancy-detection models. We also demonstrate that data obtained from eDNA sampling can be used to investigate how different groups of taxa —such as native or introduced species— respond to environmental covariates or the species richness of other groups. These results support the use of eDNA sampling for monitoring species distributions in Australia as well as globally.

Chapter 6
General discussion

6.1 Overview

In this thesis, I examined the use of eDNA sampling, combined with site occupancy-detection models, for detecting aquatic vertebrates using single- and multi-species eDNA detection methods. My aims were to i) compare the sensitivity of eDNA metabarcoding and a single-species detection method (qPCR); ii) implement a systematic study design to demonstrate how broad-scale eDNA sampling can be used to investigate the impact of land use and site level factors on platypus occupancy; iii) determine whether platypus occupancy was impacted by the 2019/2020 mega-fires in south-eastern Australia using eDNA sampling and a Before-After Control-Impact (BACI) design; and iv) demonstrate how broad-scale occupancy data from eDNA sampling can be collected for multiple fish species, revealing occupancy patterns and the impact of different land uses on native and introduced fish. To achieve these aims I conducted extensive fieldwork across south-eastern Australia and combined eDNA sampling data with hierarchical site occupancy-detection models to estimate occupancy, availability, and detection probabilities. In this chapter, I outline key findings and management implications from my four empirical chapters. I will also discuss the limitations of this work and outline my recommendations for future research directions.

6.2 Key findings

Environmental DNA method selection

Chapter 2 aimed to provide a robust and uniform approach to compare the sensitivity of single-species and multi-species eDNA detection methods. The relative sensitivity of these methods is of great importance when considering which method to choose for any given project. While there have been some comparisons of these two approaches in the literature (Harper *et al.* 2018; Bylemans *et al.* 2019; Wood *et al.* 2019; Schenekar *et al.* 2020), all have been analysed using different field and laboratory methods and different statistical analyses. This chapter advances the field by providing a consistent framework with which to compare single and multi-species eDNA detection methods between different datasets (numbers of sites, samples, replicates), different study systems and different target species. I found that, generally, qPCR was more sensitive than metabarcoding at detecting target species. Importantly, however, different methodological considerations affected the extent of the difference between the two approaches. For example, thresholds placed on data using either eDNA detection method,

altered its sensitivity for detecting target species. Therefore, I recommend emphasis be placed on the aims of the study before choosing a method, as well as careful consideration as to what methodological decisions will be made prior to selecting a detection method. For example, while metabarcoding does provide data on many species, sensitivity to detect individual species is reduced compared to qPCR. In situations where false absences would be detrimental to a project, a targeted species detection method, such as qPCR, should be chosen to reduce the probability of false negatives. A combination of both methods could also be considered. For example, Pukk *et al.* (2021) used eDNA metabarcoding to determine fish diversity and aquatic invasive species (AIS) prevalence in lakes. The maximum number of fish species detected at a site was 42, and the maximum number of AIS was three. In this circumstance, where there is a small number of invasive species, a hybrid approach could be used (employing both metabarcoding and a single-species detection method) to ensure a high probability of detection for invasive species at each site. This could ensure AIS occupancy estimates are not compromised by a less sensitive method. This approach would result in more accurate detection data for AIS and therefore potentially better-informed management actions. Hybrid studies may become more widespread as primers and probes for different species are developed and can be used more widely.

Environmental DNA sampling at broad spatial scales

Chapter 3 aimed to demonstrate that eDNA data can be collected at a broad spatial scale to investigate correlates of species occupancy. I focused on platypuses for this study. I used a stratified sampling design to ensure contracted catchments (unit of replication for stratification) were selected across different land uses. This study, to the best of my knowledge, is the largest scale eDNA study using a systematic study design. I collected eDNA samples across south-eastern Australia between August 2018 and March 2021, focusing on the peak platypus activity times of spring and autumn. The fieldwork was completed despite significant delays associated with catastrophic bushfires in south-eastern Australia, and then by the COVID-19 pandemic, whereby regional or interstate travel became impossible. Using hierarchical site occupancy-detection models, I investigated whether chosen environmental covariates affected the probability that platypuses would occupy a site. I found that land use types —such as agriculture, and grasslands and shrubs— resulted in a lower probability that platypuses would occupy a site. In contrast, the presence of appropriate burrowing banks, increased runoff, and fewer zero-flow days resulted in a higher occupancy probability. My results were consistent

with the existing body of knowledge on platypus habitat preferences (Grant and Temple-Smith 1998; Serena *et al.* 1998; Grant and Temple-Smith 2003; Bino *et al.* 2020) and therefore demonstrate that valid data can be collected using eDNA detection methods over broad spatial scales.

Chapter 5 extends the work of chapter 3, considering the broad-scale detection of fish across the state of Victoria using eDNA metabarcoding. It again demonstrates that eDNA sampling can be efficiently deployed across vast areas, not only for single species (chapter 3) but for entire biological communities. I used this data to consider how native and introduced fish respond to environmental factors. I found that both native and introduced fish species responded similarly to water availability and flow-based metrics. However, occupancy of native fish was higher at sites with a higher proportion of forest and agriculture in the surrounding contracted catchment, compared to introduced fish when using the full dataset. When only considering common species detected more than 10 times, native and introduced species occupancy probability fell with an increase in forest cover. This could be due to the presence of the cooler water species, brown trout (*Salmo trutta*), or the physical conditions present in these cooler and higher elevation regions. This chapter again demonstrates the validity of data collected using eDNA sampling with a systematic study design. We can use this method and subsequent data to help us understand patterns in the occupancy of multiple species and to map species occurrence across vast areas.

Measuring change in occupancy after a disturbance using eDNA sampling

Chapter 4 demonstrates the advantages of collecting eDNA detection data over different time periods and following extreme environmental disturbances. I collected extensive pre-fire data on platypus occupancy using eDNA sampling within and surrounding areas which were subsequently burnt in the 2019/2020 mega-fires. We then returned to 118 sites in both 2020 and 2021, resulting in sites sampled across three time points. I adapted hierarchical site occupancy-detection models to incorporate a Before-After Control-Impact (BACI) design to investigate if platypus occupancy changed in the short and medium term after the wildfires. An advantage of the BACI design was that factors other than the impact of the disturbance were controlled between sampling time points and between burnt/unburnt sites (Southwell *et al.* 2021). I found that platypus occupancy did not vary significantly between control and impact sites, indicating that platypus occupancy was not generally affected by the presence of fire.

This is a positive outcome for platypuses, a species now classed as *vulnerable* in the state of Victoria under the Flora and Fauna Guarantee Act 1998. I did, however, find a significant impact on platypus occupancy when considering the interaction between the proportion of a watershed burnt at high severity and the amount of rainfall in a watershed post-fire. The occupancy probability of platypuses was lower at sites where the local watershed both contained a large proportion of vegetation burnt at high severity and experienced large amounts of post-fire rainfall. High severity fire in combination with high rainfall has already been found to be a significant threat to aquatic organisms such as fish and macroinvertebrates due to decreases in water quality and sediment flows (Shakesby *et al.* 2007; Lyon and O'Connor 2008; Verkaik *et al.* 2014). This finding demonstrates that semi-aquatic species such as platypuses can also be impacted by extreme wildfire events.

Management implications for platypuses

In this thesis, I have shown that eDNA sampling is a very useful tool to detect platypuses and to efficiently investigate correlates of their occupancy over large spatial extents. Platypuses are a widespread species, previously assumed to be common. Their distribution spans eastern Australia from Tasmania in the south to northern Queensland in the north (Woinarski and Burbidge 2016). They are now classed as *vulnerable* in the state of Victoria due to mounting evidence of their decline in some areas due to urban growth and habitat destruction, as well as drought and altered flow regimes (Woinarski and Burbidge 2016; Scientific Advisory Committee 2020). Platypuses are a difficult species to detect. They are cryptic and predominately nocturnal, so incidental observations, while useful, do not give the full picture of platypus distribution (Lugg *et al.* 2017). Trapping methods are generally used to gain population-level data on the health and status of platypuses (Bino *et al.* 2019; Hawke *et al.* 2021), however this method cannot be deployed on a distribution-wide scale for this species in a timely and cost effective manner. I have demonstrated in chapters 3 and 4 that eDNA sampling can be used as a broad-scale sampling tool to monitor the occupancy of this species across a large proportion of their widespread distribution. Environmental DNA sampling can identify locations where trapping for population monitoring will be useful and can provide baseline data on occupancy that is likely to be critical for monitoring distributions over time. This baseline data will be especially important for a cryptic species such as platypuses, where distribution data is patchy at best and lacking in many areas, especially in difficult to access and remote regions.

As I demonstrated in chapter 3, eDNA sampling can also be used to efficiently investigate correlates of a species distribution. A key finding was that platypuses were less likely to occupy sites with predominately agricultural and grassland land uses and more likely to occupy sites with suitable banks for burrowing, more runoff, and less zero-flow days. Reductions in winter and autumn rainfall are predicted as the climate in south-eastern Australia changes (Hope *et al.* 2018), potentially resulting is less runoff and more zero-flow days. This could have negative consequences for the distribution of platypuses as they are more likely to occupy areas with more (or more consistent) rainfall, leading to increased runoff and less zero-flow days.

My results in chapter 4 show that platypus occupancy could be negatively impacted by the combination of high severity fire and high rainfall post-fire. There is a trend of larger mega-fires occurring globally where greater extents of forested areas are being burnt at high severity (Collins *et al.* 2021). High model agreement suggests that south-eastern Australia will receive more extreme rain events (Hope *et al.* 2018). This combination of extreme rainfall and high severity fires could impact platypuses and other aquatic species significantly and thus should be considered in the conservation and management decisions of this group in the future.

Despite being an iconic species once considered widespread and common, I have demonstrated that, in the future, they could be at risk of decline. The future implications of climate change need to be considered in any management, and ongoing monitoring is essential for providing the data necessary to document any decline. Environmental DNA sampling would be the optimal tool for such a monitoring program and the data accumulated during this PhD can provide a baseline for monitoring change.

6.3 Key advances

This thesis advances knowledge of eDNA sampling and the inferences that can be made from this emerging technology. It was especially important to demonstrate the use of eDNA sampling across vast areas of Australia (chapters 3 and 5), as this method has not yet been routinely applied across such large scales to detect rare, cryptic, or invasive species (though it's use is increasing rapidly; Smart *et al.* 2015; Bylemans *et al.* 2016; Hinlo *et al.* 2017; Piggott 2017; Tingley *et al.* 2021; West *et al.* 2021; McColl-Gausden *et al.* 2021). These results and applications are globally relevant, demonstrating the huge spatial scales over which high quality data can be collected, and the conclusions that can be drawn from such data. For example, chapters 3 and 5 considered platypus and fish distributions, and without even sighting

a platypus or a fish during fieldwork, I was able to investigate how different land uses, flow, and runoff, and even species richness, affected a species' occupancy. Chapter 5 also demonstrated that threatened species can be detected using metabarcoding —I detected four species of threatened fish— without *a priori* knowledge of their occupancy. This data provides a baseline for future monitoring upon which expanded eDNA surveys could be used to efficiently map their distribution. Chapter 4 was born out of circumstance, but without the extensive eDNA sampling I had conducted prior to the 2019/2020 mega-fires, I would not have been able to undertake the subsequent analysis and determine the impact of these fires on platypus occupancy. Pre-fire data across some very remote parts of south-eastern Australia were invaluable for not only considering the impacts of the large-scale fires on platypuses, but also impacts on other vertebrates using the multi-species detection method metabarcoding. The use of site occupancy-detection models across all chapters enabled me to incorporate uncertainty at the inherently hierarchal levels of eDNA sampling, the site, sample, and replicate levels. I was able to adapt these models to answer my research aims, demonstrating the promise of this very flexible modelling framework.

6.4 Limitations

There are several limitations in this work that should be addressed to assist interpretation. An evolving area of research is the transport or movement of eDNA in lotic systems. Estimates of eDNA transport vary considerably depending on the size and substrate of the waterway (Fremier *et al.* 2019). Environmental DNA can be degraded to the point where it cannot be detected, as well as be adsorbed to the substrate (Nukazawa *et al.* 2018). I assumed in each of my chapters that DNA detected at a site was from an individual or individuals occupying that site (or in close proximity). However, due to eDNA transport, individuals could be located upstream of a site. I attempted to account for transport with covariates in chapters 3, 4 and 5 by considering factors that might affect occupancy at the contracted catchment or watershed level. For example, the proportion of a contracted catchment covered by a certain land use, as opposed to the land use at the sampling site. By considering these catchment level factors, the effect eDNA transport has on any inferences is greatly reduced. For the site-level covariates (chapters 3 and 5) my assumption of site occupancy may not be correct in all circumstances, but I believe it is justified given current understanding of eDNA transport, the size of the rivers sampled (generally smaller rivers and creeks), and the distance between sampling points

(generally more than 7km). Improvements in our understanding of eDNA transport could considerably improve the accuracy of this work.

Another limitation of this research, which I discussed in chapter 3, pertains to the distribution of site level variables. While the contracted catchments selected for this study were stratified across different vegetation states, I could not ensure the appropriate stratification for site level variables, as these could not be determined prior to the measurements being taken in the field. This resulted in some site level variables being under-represented in the dataset, potentially resulting in more uncertainty around those variables. For example, the "excellent" category for burrowing banks was only represented by 45/504 sites. Additional targeted sampling could add more sites into these under-represented categories and thus potentially reduce the uncertainty around the parameter estimates presented in chapter 3.

In chapter 3, I collected ordinal data on the suitability of a site for platypuses based on previous research of platypus habitat preferences (e.g., good, fair, or excellent habitat). I deemed ordinal data as appropriate at the time of the study design due to the need for multiple people to take eDNA samples for the project, including the potential addition of citizen science data. I predicted that we would be more precise in our estimates of the status of a site if the site characteristics were placed into categories, as opposed to continuous values. However, there were limitations in this use of ordinal data to represent habitat conditions at a site. For example, ordinal data results in reduced statistical power compared to continuous variables, as data is lost within categories. This may have impacted the ability to determine the relationship between site characteristics and platypus occupancy accurately.

I developed different hierarchical site occupancy-detection models for all chapters, based on those presented in Schmidt *et al.* (2013) and *Applied hierarchical modelling in ecology* (Kéry and Royle 2016). Due to the novel nature of these models and time constraints, I was not able to conduct thorough sensitivity analyses on these models for all chapters. A traditional sensitivity analysis, varying the probability of occupancy and the parameters of interest, would be ideal. I was able to conduct a thorough sensitivity analysis for chapter 3, varying occupancy and β estimates (see Appendix C), however, more parameters could be investigated. Ideally, such sensitivity analyses could be expanded to include the models used in chapters 4 and 5.

Finally, in chapter 4, the power to detect a change in occupancy (Guillera-Arroita and Lahoz-Monfort 2012) was an important component which was lacking. A power analysis, considering

the hierarchical BACI design we developed, was beyond the scope of the chapter and would take more time to develop. It would, however, be a worthwhile pursuit to consider how many sites need to be sampled across such a large spatial scale to detect different levels of change in occupancy or distribution, and with different initial occupancy probabilities.

6.5 Future research directions

My thesis has explored several components relevant to the use of eDNA sampling, however, there are many future priorities that I have identified for further research.

First, I have collected eDNA data for platypuses covering approximately 37% of their predicted distribution and identified some covariates which affect the species' occupancy. This extensive dataset could be used to construct a species distribution model (SDM) to help predict where platypuses are most likely to occur. A SDM could help inform future fine-scale trapping surveys or identify areas of important platypus habitat for future interventions (Coleman and Bragg 2021).

Second, predicting how platypus distribution may shift over time with land use, climate, and fire regime changes will be an important step to ensuring their conservation. Metapopulation models, such as those used in Bino *et al.* (2020), have been used to predict platypus population declines with drought and fragmentation under climate change. Updated distribution data using eDNA sampling could be used in combination with trapping and population data to develop spatially explicit population viability analyses (PVAs), such as those using *steps* (Visintin *et al.* 2020) or RAMAS GIS as in Bino *et al.* (2020). These packages and programs allow for the integration of static or temporally variable spatial layers to define change in the landscape. It would therefore be interesting to incorporate additional data not previously considered about land use change, as well as fire and high rainfall events which I found to impact the occupancy of platypuses. PVAs could help demonstrate how a population may change over time under different management or climate scenarios (Bino *et al.* 2020), providing evidence for species classifications and for informing conservation planning.

Third, multi-species eDNA detection methods and the samples I collected could be used to investigate how other species were impacted by the 2019/2020 mega-fires. It would be interesting to investigate if fire presence, or other more local impacts such as the interaction of fire severity and rainfall post-fire, impacted other aquatic species occupancy strongly. These

types of findings could provide further evidence that aquatic species are not always buffered from the effects of fire, and that if these combinations of high fire severity and high rainfall post-fire keep occurring, many aquatic species may be more threatened than previously thought. This baseline data can also help managers plan or prioritise areas for post-fire assessments in the future.

Fourth, as highlighted in the limitations of my work, eDNA transport is an area of active research. Developing methods to predict eDNA transport is complicated, due to the complex nature of eDNA detection in a waterway. Environmental DNA needs to be treated as a "sticky" particle that can adsorb to the substrate and desorb back into the water column. Environmental DNA can also be degraded in water, depending on the temperature and amount of UV light. Progress is being made in this area (Fremier *et al.* 2019; Nevers *et al.* 2020) but a thorough understanding of eDNA movement will ensure conclusions drawn from eDNA sampling are as accurate as possible. An example of how eDNA transport could be measured in the systems I worked in could be by conducting an experiment where a known concentration of a species eDNA (from a zoo or sanctuary for example) is poured at a steady state into a waterway where that species is not present. The detection of that DNA over time and at different distances from the source could help build a model of DNA movement in that system.

Fifth, chapter 5 used multi-species detection methods and hierarchical site occupancy-detection models to investigate how native and introduced species responded to environmental factors. An extension of this work would be to use joint species distribution models (JSDMs) to investigate the interactions between individual species. These species interactions will be more helpful for management interventions. For example, if an introduced species is negatively interacting with a native threatened species, the introduced species can be targeted or monitored. JSDMs are also being developed which could account for imperfect detection (Hogg *et al.* 2021) enabling multi-level data like eDNA data to be incorporated.

Lastly, eDNA detection methods are improving in their accuracy and the costs of using them are reducing (Hunter *et al.* 2019; Loeza-Quintana *et al.* 2020; Burian *et al.* 2021). These developments could allow the use of multiple methods in a single study to help confirm any findings robustly, as well as enable sampling across larger spatial scales. For example, metabarcoding using different target regions, in combination with targeted single-species detection methods for certain priority species, could become a part of routine monitoring programs. Biobanking could play a role for the centralised, long term storage of samples and

data, enabling long-term datasets to be established, and newer technologies to be used on older samples (Jarman *et al.* 2018).

6.6 Conclusions

This thesis presents novel methods to analyse eDNA detection data and demonstrates the vast array of questions which can be explored using eDNA sampling and site occupancy-detection models. I have presented a consistent model with which to compare the sensitivity of eDNA metabarcoding and a single-species detection method, highlighting the need for careful consideration of the detection method based on the aims of the study. I have demonstrated the broad-scale sampling that can be achieved using both single- and multi-species eDNA detection methods, highlighting the benefits of a systematic study design in drawing conclusions about the factors affecting species occupancy. I have also shown that platypus occupancy was not affected significantly by the 2019/2020 mega-fires in the areas we sampled but have added to the growing evidence that aquatic species can be negatively impacted by the combination of high severity fire and large rainfall events post-fire. This thesis highlights the utility and benefits of using eDNA detection methods to monitor aquatic biodiversity, providing important baseline data from which to predict the effects of global changes.

References

Abram, N. J., Henley, B. J., Sen Gupta, A., Lippmann, T. J. R., Clarke, H., Dowdy, A. J., Sharples, J. J., Nolan, R. H., Zhang, T., Wooster, M. J., Wurtzel, J. B., Meissner, K. J., Pitman, A. J., Ukkola, A. M., Murphy, B. P., Tapper, N. J., and Boer, M. M. (2021). Connections of climate change and variability to large and extreme forest fires in southeast Australia. *Communications Earth & Environment* **2**. doi:https://doi.org/10.1038/s43247-020-00065-8

Andersen, K., Bird, K. L., Rasmussen, M., Haile, J., Breuning-madsen, H., Kjær, K. H., Orlando, L., Gilbert, M. T. P., and Willerslev, E. (2012). Meta-barcoding of ‘ dirt ’ DNA from soil reflects vertebrate biodiversity. *Molecular Ecology* **21**, 1966–1979. doi:10.1111/j.1365-294X.2011.05261.x

Banks, S. C., Knight, E. J., McBurney, L., Blair, D., and Lindenmayer, D. B. (2011). The effects of wildfire on mortality and resources for an arboreal marsupial: Resilience to fire events but susceptibility to fire regime change. *PLoS ONE* **6**. doi:10.1371/journal.pone.0022952

Benyon, R. G., and Lane, P. N. J. (2013). Ground and satellite-based assessments of wet eucalypt forest survival and regeneration for predicting long-term hydrological responses to a large wildfire. *Forest Ecology and Management* **294**, 197–207. doi:10.1016/j.foreco.2012.04.003

Biggs, J., Ewald, N., Valentini, A., Gaboriaud, C., Dejean, T., Griffiths, R. A., Foster, J., Wilkinson, J. W., Arnell, A., Brotherton, P., Williams, P., and Dunn, F. (2015). Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt (Triturus cristatus). *Biological Conservation* **183**, 19–28. doi:10.1016/j.biocon.2014.11.029

Bino, G., Hawke, T., and Kingsford, R. T. (2021). Synergistic effects of a severe drought and fire on platypuses. *Science of the Total Environment* **777**, 146137. doi:10.1016/j.scitotenv.2021.146137

Bino, G., Kingsford, R. T., Archer, M., Connolly, J. H., Day, J., Dias, K., Goldney, D., Gongora, J., Grant, T., Griffiths, J., Hawke, T., Klamt, M., Lunney, D., Mijangos, L., Munks, S., Sherwin, W., Serena, M., Temple-Smith, P., Thomas, J., Williams, G., and Whittington, C. (2019). The platypus: Evolutionary history, biology, and an uncertain

future. *Journal of Mammalogy* **100**, 308–327. doi:10.1093/jmammal/gyz058

Bino, G., Kingsford, R. T., and Wintle, B. A. (2020). A stitch in time – Synergistic impacts to platypus metapopulation extinction risk. *Biological Conservation* **242**, 108399. doi:10.1016/j.biocon.2019.108399

Blackman, R. C., Ling, K. K. S., Harper, L. R., Shum, P., Hänfling, B., and Lawson-Handley, L. (2020). Targeted and passive environmental DNA approaches outperform established methods for detection of quagga mussels, Dreissena rostriformis bugensis in flowing water. *Ecology and Evolution* **10**, 13248–13259. doi:10.1002/ece3.6921

Blake, D., Nyman, P., Nice, H., D'souza, F. M. L., Kavazos, C. R. J., and Horwitz, P. (2020). Assessment of post-wildfire erosion risk and effects on water quality in south-western Australia. *International Journal of Wildland Fire* **29**, 240–257. doi:10.1071/WF18123

Boer, M. M., Resco de Dios, V., and Bradstock, R. A. (2020). Unprecedented burn area of Australian mega forest fires. *Nature Climate Change* **10**, 170. doi:10.1038/s41558-020-0710-7

Bohmann, K., Chua, P., Holman, L. E., and Lynggaard, C. (2021). DNAqua-Net conference unites participants from around the world with the quest to standardize and implement DNA-based aquatic biomonitoring. *Environmental DNA*, 1–5. doi:10.1002/edn3.207

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., Yu, D. W., and de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution* **29**, 358–367. doi:10.1016/j.tree.2014.04.003

Boivin-Delisle, D., Laporte, M., Burton, F., Dion, R., Normandeau, E., and Bernatchez, L. (2021). Using environmental DNA for biomonitoring of freshwater fish communities: Comparison with established gillnet surveys in a boreal hydroelectric impoundment. *Environmental DNA* **3**, 105–120. doi:10.1002/edn3.135

Bowman, D., Williamson, G., Yebra, M., Lizundia-Loiola, J., Pettinari, M. L., Shah, S., Bradstock, R., and Chuvieco, E. (2020). Wildfires: Australia needs national monitoring agency. *Nature* **584**, 188–191. doi:10.1038/d41586-020-02306-4

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., and Coissac, E. (2016). obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology*

Resources **16**, 176–182. doi:10.1111/1755-0998.12428

Bureau of Meteorology (2021). Climate Statistics.

Burian, A., Mauvisseau, Q., Bulling, M., Domisch, S., Qian, S., and Sweet, M. (2021). Improving the reliability of eDNA data interpretation. *Molecular Ecology Resources* **21**, 1422–1433. doi:10.1111/1755-0998.13367

Bush, A., Monk, W. A., Compson, Z. G., Peters, D. L., Porter, T. M., Shokralla, S., Wright, M. T. G., Hajibabaei, M., and Baird, D. J. (2020). DNA metabarcoding reveals metacommunity dynamics in a threatened boreal wetland wilderness. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 8539–8545. doi:10.1073/pnas.1918741117

Buxton, A. S., Groombridge, J. J., and Griffiths, R. A. (2018). Seasonal variation in environmental DNA detection in sediment and water samples. *PLoS ONE* **13**, 1–14. doi:10.1371/journal.pone.0191737

Bylemans, J., Furlan, E. M., Hardy, C. M., McGuffie, P., Lintermans, M., and Gleeson, D. M. (2017). An environmental DNA-based method for monitoring spawning activity: a case study, using the endangered Macquarie perch (Macquaria australasica). *Methods in Ecology and Evolution* **8**, 646–655. doi:10.1111/2041-210X.12709

Bylemans, J., Furlan, E. M., Pearce, L., Daly, T., and Gleeson, D. M. (2016). Improving the containment of a freshwater invader using environmental DNA (eDNA) based monitoring. *Biological Invasions*, 1–9. doi:10.1007/s10530-016-1203-5

Bylemans, J., Gleeson, D. M., Hardy, C. M., and Furlan, E. (2018). Toward an ecoregion scale evaluation of eDNA metabarcoding primers : A case study for the freshwater fish biodiversity of the Murray – Darling Basin (Australia). *Ecology and Evolution*, 8697–8712. doi:10.1002/ece3.4387

Bylemans, J., Hardy, C. M., Gleeson, D. M., Duncan, R. P., and Furlan, E. M. (2019). A performance evaluation of targeted eDNA and eDNA metabarcoding analyses for freshwater fishes. *Environmental DNA* **1**, 402–414. doi:10.1002/edn3.41

Cadwallader, P. L. (1996). Overview of the impacts of introduced salmonids on Australian native fauna. Canberra.

Chia, E. K., Bassett, M., Leonard, S. W. J., Holland, G. J., Ritchie, E. G., Clarke, M. F., and

Bennett, A. F. (2016). Effects of the fire regime on mammal occurrence after wildfire: Site effects vs landscape context in fire-prone forests. *Forest Ecology and Management* **363**, 130–139. doi:10.1016/j.foreco.2015.12.008

Christie, A. P., Amano, T., Martin, P. A., Shackelford, G. E., Simmons, B. I., and Sutherland, W. J. (2019). Simple study designs in ecology produce inaccurate estimates of biodiversity responses. *Journal of Applied Ecology* **56**, 2742–2754. doi:10.1111/1365-2664.13499

Clark, K., Karsch-mizrachi, I., Lipman, D. J., Ostell, J., and Sayers, E. W. (2016). GenBank. *Nucleic Acids Research* **44**, 67–72. doi:10.1093/nar/gkv1276

Clarke, P. J., Lawes, M. J., Murphy, B. P., Russell-Smith, J., Nano, C. E. M., Bradstock, R., Enright, N. J., Fontaine, J. B., Gosper, C. R., Radford, I., Midgley, J. J., and Gunton, R. M. (2015). A synthesis of postfire recovery traits of woody plants in Australian ecosystems. *Science of the Total Environment* **534**, 31–42. doi:10.1016/j.scitotenv.2015.04.002

Closek, C. J., Santora, J. A., Starks, H. A., Schroeder, I. D., Andruszkiewicz, E. A., Sakuma, K. M., Bograd, S. J., Hazen, E. L., Field, J. C., and Boehm, A. B. (2019). Marine Vertebrate Biodiversity and Distribution Within the Central California Current Using Environmental DNA (eDNA) Metabarcoding and Ecosystem Surveys. *Frontiers in Marine Science* **6**, 1–17. doi:10.3389/fmars.2019.00732

Coissac, E., Riaz, T., and Puillandre, N. (2012). Bioinformatic challenges for DNA metabarcoding of plants and animals. *Moecular Ecology* **21**, 1834–1847. doi:10.1111/j.1365-294X.2012.05550.x

Coleman, M. A., and Bragg, J. G. (2021). A decision framework for evidence-based climate adaptation interventions. *Global Change Biology* **27**, 472–474. doi:10.1111/gcb.15429

Coleman, R., Raadik, T., and Freeman, R. (2019). Galaxiella pusilla. The IUCN Red List of Threatened Species 2019. doi:https://dx.doi.org/10.2305/IUCN.UK.2019-3.RLTS.T8820A123377818.en.

Collins, L. (2020). Eucalypt forests dominated by epicormic resprouters are resilient to repeated canopy fires. *Journal of Ecology* **108**, 310–324. doi:10.1111/1365-2745.13227

Collins, L., Bradstock, R. A., Clarke, H., Clarke, M. F., Nolan, R. H., and Penman, T. D.

(2021). The 2019/2020 mega-fires exposed Australian ecosystems to an unprecedented extent of high-severity fire. *Environmental Research Letters* **16**. doi:10.1088/1748-9326/abeb9e

Collins, L., Griffioen, P., Newell, G., and Mellor, A. (2018). The utility of Random Forests for wildfire severity mapping. *Remote Sensing of Environment* **216**, 374–384. doi:10.1016/j.rse.2018.07.005

Collins, L., McCarthy, G., Mellor, A., Newell, G., and Smith, L. (2020). Training data requirements for fire severity mapping using Landsat imagery and random forest. *Remote Sensing of Environment* **245**, 111839. doi:10.1016/j.rse.2020.111839

Commonwealth of Australia (2012). Bureau of Meteorology.

Cowart, D. A., Matabos, M., Brandt, M. I., Marticorena, J., and Sarrazin, J. (2020). Exploring Environmental DNA (eDNA) to Assess Biodiversity of Hard Substratum Faunal Communities on the Lucky Strike Vent Field (Mid-Atlantic Ridge) and Investigate Recolonization Dynamics After an Induced Disturbance. *Frontiers in Marine Science* **6**, 1–21. doi:10.3389/fmars.2019.00783

Cristescu, R. H., Foley, E., Markula, A., Jackson, G., Jones, D., and Frère, C. (2015). Accuracy and efficiency of detection dogs: a powerful new tool for koala conservation and management. *Scientific Reports* **5**. doi:10.1038/srep08349

Crookes, S., Heer, T., Castañeda, R. A., Mandrak, N. E., Heath, D. D., Weyl, O. L. ., MacIsaac, H. J., and Foxcroft, L. C. (2020). NeoBiota Monitoring the silver carp invasion in Africa : a case study using environmental DNA (eDNA) in dangerous watersheds. *NeoBiota* **47**, 31–47. doi:10.3897/neobiota.56.47475

Dalén, L., Ötherström, A., Eijer, T., and Hapiro, B. (2007). Recovery of DNA from Footprints in the Snow. *Canadian Field-Naturalist* **121**, 321–324. doi:http://dx.doi.org/10.22621/cfn.v121i3.482

Darling, J. A., and Mahon, A. R. (2011). From molecules to management: Adopting DNA-based methods for monitoring biological invasions in aquatic environments. *Environmental Research* **111**, 978–988. doi:10.1016/j.envres.2011.02.001

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E., and Bernatchez, L.

(2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology* **26**, 5872–5895. doi:10.1111/mec.14350

Dejean, T., Valentini, A., Miquel, C., Taberlet, P., Bellemain, E., and Miaud, C. (2012). Improved detection of an alien invasive species through environmental DNA barcoding: The example of the American bullfrog Lithobates catesbeianus. *Journal of Applied Ecology* **49**, 953–959. doi:10.1111/j.1365-2664.2012.02171.x

Doi, H., Fukaya, K., and Miya, M. (2019). Evaluation of detection probabilities at the water-filtering and initial PCR steps in environmental DNA metabarcoding using a multispecies site occupancy model. *Scientific Reports* **9**. doi:10.1038/s41598-019-40233-1

Doi, H., Takahara, T., Minamoto, T., Matsuhashi, S., Uchii, K., and Yamanaka, H. (2015). Droplet digital polymerase chain reaction (PCR) outperforms real-time PCR in the detection of environmental DNA from an invasive fish species. *Environmental science & technology* **49**, 5601–8. doi:10.1021/acs.est.5b00253

Dorazio, R. M., and Erickson, R. A. (2018). EDNAOCCUPANCY : An R package for multiscale occupancy modelling of environmental DNA data. *Molecular Ecology* **18**, 368–380. doi:10.1111/1755-0998.12735

Dragovich, D., and Morris, R. (2002). Fire intensity, slopewash and bio-transfer of sediment in Eucalypt Forest, Australia. *Earth Surface Processes and Landforms* **27**, 1309–1319. doi:10.1002/esp.413

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z. I., Knowler, D. J., Lévêque, C., Naiman, R. J., Prieur-Richard, A. H., Soto, D., Stiassny, M. L. J., and Sullivan, C. A. (2006). Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews of the Cambridge Philosophical Society* **81**, 163–182. doi:10.1017/S1464793105006950

Elbrecht, V., and Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. *PLoS ONE* **10**, 1–16. doi:10.1371/journal.pone.0130324

Elbrecht, V., and Leese, F. (2017). Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment. *Frontiers in Environmental*

Science **5**, 1–11. doi:10.3389/fenvs.2017.00011

Ellem, B. A., Bryant, A., and O'Connor, A. (1998). Statistical Modelling of Platypus, Ornithorhynchus anatinus, Habitat Preferences Using Generalised Linear Models. *Australian Mammalogy* **20**, 281. doi:10.1071/am98281

Esling, P., Lejzerowicz, F., and Pawlowski, J. (2015). Accurate multiplexing and filtering for high-throughput amplicon-sequencing. *Nucleic Acids Research* **43**, 2513–2524. doi:10.1093/nar/gkv107

Evans, N. T., and Lamberti, G. A. (2017). Freshwater fisheries assessment using environmental DNA: A primer on the method, its potential, and shortcomings as a conservation tool. *Fisheries Research* **197**, 60–66. doi:10.1016/j.fishres.2017.09.013

Evans, N. T., Olds, B. P., Renshaw, M. A., Turner, C. R., Li, Y., Jerde, C. L., Mahon, A. R., Pfrender, M. E., Lamberti, G. A., and Lodge, D. M. (2016). Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Molecular Ecology Resources* **16**, 29–41. doi:10.1111/1755-0998.12433

Evans, N. T., Shirey, P. D., Wieringa, J. G., Mahon, A. R., and Lamberti, G. A. (2017). Comparative Cost and Effort of Fish Distribution Detection via Environmental DNA Analysis and Electrofishing. *Fisheries* **42**, 90–99. doi:10.1080/03632415.2017.1276329

Fadrosh, D. W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R. M., and Ravel, J. (2014). An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* **2**, 6. doi:10.1186/2049-2618-2-6

Fediajevaite, J., Priestley, V., Arnold, R., and Savolainen, V. (2021). Meta-analysis shows that environmental DNA outperforms traditional surveys, but warrants better reporting standards. *Ecology and Evolution* **11**, 4803–4815. doi:10.1002/ece3.7382

Ficetola, G. ., Taberlet, P., and Coissac, E. (2016). How to limit false positives in environmental DNA and metabarcoding? *Molecular Ecology Resources* **16**, 604–607. doi:10.1111/1755-0998.12508

Ficetola, G., Coissac, E., Zundel, S., Riaz, T., Taberlet, P., and Pompanon, F. (2010). An In silico approach for the evaluation of DNA barcodes. *BMC genomics* **11**. doi:https://doi.org/10.1186/1471-2164-11-434

Filkov, A. I., Ngo, T., Matthews, S., Telfer, S., and Penman, T. D. (2020). Impact of Australia's catastrophic 2019/20 bushfire season on communities and environment. Retrospective analysis and current trends. *Journal of Safety Science and Resilience* **1**, 44–56. doi:10.1016/j.jnlssr.2020.06.009

Fonseca, V. G. (2018). "Pitfalls in relative abundance estimation using eDNA metabarcoding". *Molecular Ecology Resources* **18**, 923–926. doi:10.1111/1755-0998.12902

Fraija-Fernández, N., Bouquieaux, M. C., Rey, A., Mendibil, I., Cotano, U., Irigoien, X., Santos, M., and Rodríguez-Ezpeleta, N. (2020). Marine water environmental DNA metabarcoding provides a comprehensive fish diversity assessment and reveals spatial patterns in a large oceanic area. *Ecology and Evolution* **10**, 7560–7584. doi:10.1002/ece3.6482

Fremier, A. K., Strickler, K. M., Parzych, J., Powers, S., and Goldberg, C. S. (2019). Stream Transport and Retention of Environmental DNA Pulse Releases in Relation to Hydrogeomorphic Scaling Factors. *Environmental Science & Technology* **53**, 6640–6649. doi:10.1021/acs.est.8b06829

Frost, A. J., Ramchurn, A., and Smith, A. (2018). The Australian Landscape Water Balance model (AWRA-L v6). Technical Description of the Australian Water Resources Assessment Landscape model version 6.

Fukaya, K., Kondo, N. I., Matsuzaki, S. ichiro S., and Kadoya, T. (2022). Multispecies site occupancy modelling and study design for spatially replicated environmental DNA metabarcoding. *Methods in Ecology and Evolution* **13**, 183–193. doi:10.1111/2041-210X.13732

Fukumoto, S., Ushimaru, A., and Minamoto, T. (2015). A basin-scale application of environmental DNA assessment for rare endemic species and closely related exotic species in rivers: A case study of giant salamanders in Japan. *Journal of Applied Ecology* **52**, 358–365. doi:10.1111/1365-2664.12392

Furlan, E. M., Gleeson, D., Hardy, C. M., and Duncan, R. P. (2016). A framework for estimating the sensitivity of eDNA surveys. *Molecular Ecology Resources* **16**, 641–654. doi:10.1111/1755-0998.12483

Geoscience, A. (2017). Dynamic Land Cover Dataset. , 1–15.

Giglio, L., Justice, C., Boschetti, L., and Roy, D. (2015). MCD64A1 MODIS/Terra+Aqua Burned Area Monthly L3 Global 500m SIN Grid V006 [Data set]. *NASA EOSDIS Land Processes DAAC*. doi:https://doi.org/10.5067/MODIS/MCD64A1.006

Goldberg, C. S., Pilliod, D. S., Arkle, R. S., and Waits, L. P. (2011). Molecular detection of vertebrates in stream water: A demonstration using rocky mountain tailed frogs and Idaho giant salamanders. *PLoS ONE* **6**. doi:10.1371/journal.pone.0022746

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., Spear, S. F., McKee, A., Oyler-McCance, S. J., Cornman, R. S., Laramie, M. B., Mahon, A. R., Lance, R. F., Pilliod, D. S., Strickler, K. M., Waits, L. P., Fremier, A. K., Takahara, T., Herder, J. E., and Taberlet, P. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*. doi:10.1111/2041-210X.12595

Goodenough, A. E., Carpenter, W. S., Mactavish, L., Mactavish, D., Theron, C., and Hart, A. G. (2017). Empirically testing the effectiveness of thermal imaging as a tool for identification of large mammals in the African bushveldt. *African Journal of Ecology* **56**, 51–62. doi:10.1111/aje.12416

Grant, T. ., and Temple-Smith, P. . (2003). Conservation of the platypus, Ornithorhynchus anatinus : Threats and challenges. *Aquatic Ecosystem Health and Management* **6**, 5–18. doi:10.1080/14634980390151538

Grant, T. R., and Dawson, T. (1978). Temperature regulation in the platypus, Ornithorhynchus anatinus: Production and loss of Metabolic heat in air and water. *Physiological Zoology* **51**.

Grant, T. R., and Temple-Smith, P. D. (1998). Field biology of the platypus (Ornithorhynchus anatinus): Historical and current perspectives. *Philosophical Transactions of the Royal Society B: Biological Sciences* **353**, 1081–1091. doi:10.1098/rstb.1998.0267

Griffiths, J., Kelly, T., and Weeks, A. (2013). Net-avoidance behaviour in platypuses. *Australian Mammalogy* **35**, 245–247. doi:10.1071/AM12051

Griffiths, J., van Rooyen, A., and Weeks, A. R. (2016). Platypus distribution and relative

abundance in the MacKenzie River and Upper Wimerra 2016 (Report to Wimmera Catchment Management Authority). Parkville, VIC.

Griffiths, J., and Weeks, A. (2014). Platypus distribution and abundance in the MacKenzie River: impact of fire. Report to Wimmera Catchment Management Authority. Parkville.

Griffiths, J., and Weeks, A. R. (2018). Investigating the current distribution of platypuses, rakali, and blackfish in the upper Wimmera region (Report to Project Platypus). Parkville, VIC.

Griffiths, J., and Weeks, A. R. (2013). Platypus distribution and abundance in the upper Wimmera (Report for Project Platypus). Parkville, VIC.

Guillera-Arroita, G., and Lahoz-Monfort, J. J. (2012). Designing studies to detect differences in species occupancy: Power analysis under imperfect detection. *Methods in Ecology and Evolution* **3**, 860–869. doi:10.1111/j.2041-210X.2012.00225.x

Guillera-Arroita, G., Lahoz-Monfort, J. J., Milner-Gulland, E. J., Young, R. P., and Nicholson, E. (2010a). Using occupancy as a state variable for monitoring the Critically Endangered Alaotran gentle lemur Hapalemur alaotrensis. *Endangered Species Research* **11**, 157–166. doi:10.3354/esr00274

Guillera-Arroita, G., Lahoz-Monfort, J. J., van Rooyen, A. R., Weeks, A. R., and Tingley, R. (2017). Dealing with false-positive and false-negative errors about species occurrence at multiple levels. *Methods in Ecology and Evolution* **8**, 1081–1091. doi:10.1111/2041-210X.12743

Guillera-Arroita, G., Ridout, M. S., and Morgan, B. J. T. (2010b). Design of occupancy studies with imperfect detection. *Methods in Ecology and Evolution* **1**, 131–139. doi:10.1111/j.2041-210x.2010.00017.x

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., Mcdonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M. R., Possingham, H. P., and Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters* **16**, 1424–1435. doi:10.1111/ele.12189

Gust, N., and Handasyde, K. (1995). Seasonal-Variation in the ranging behavior of the

platypus (ornithorhynchus-anatinus) on the Goulburn River, Victoria. *Australian Journal of Zoology* **43**, 193–208. doi:10.1071/ZO9950193

Hammill, K. A., and Bradstock, R. A. (2006). Remote sensing of fire severity in the Blue Mountains: Influence of vegetation type and inferring fire intensity. *International Journal of Wildland Fire* **15**, 213–226. doi:10.1071/WF05051

Harper, L. R., Handley, L. L., Hahn, C., Boonham, N., Rees, H. C., Gough, K. C., Lewis, E., Adams, I. P., Brotherton, P., Phillips, S., and Haenfling, B. (2018). Needle in a haystack? A comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (Triturus cristatus). *Ecology and Evolution* **8**, 6330–6341. doi:https://doi.org/10.1002/ece3.4013

Hawke, T., Bino, G., and Kingsford, R. T. (2019). Short Communication A silent demise : Historical insights into population changes of the iconic platypus (Ornithorhynchus anatinus). *Global Ecology and Conservation* **20**, e00720. doi:10.1016/j.gecco.2019.e00720

Hawke, T., Bino, G., Kingsford, R. T., Iervasi, D., Iervasi, K., and Taylor, M. D. (2020). Fine-scale movements and interactions of platypuses , and the impact of an environmental flushing flow. *Freshwater Biology*, 1–12. doi:10.1111/fwb.13628

Hawke, T., Bino, G., Kingsford, R. T., Iervasi, D., Iervasi, K., and Taylor, M. D. (2021). Long - term movements and activity patterns of platypus on regulated rivers. *Scientific Reports*, 1–11. doi:10.1038/s41598-021-81142-6

Hempel, C. A., Peinert, B., Beermann, A. J., Elbrecht, V., Macher, J. N., Macher, T. H., Jacobs, G., and Leese, F. (2020). Using Environmental DNA to Monitor the Reintroduction Success of the Rhine Sculpin (Cottus rhenanus) in a Restored Stream. *Frontiers in Ecology and Evolution* **8**. doi:10.3389/fevo.2020.00081

Hinlo, R., Furlan, E., Suitor, L., and Gleeson, D. (2017). Environmental DNA monitoring and management of invasive fish: comparison of eDNA and fyke netting. *Management of Biological Invasions* **8**, 89–100. doi:10.3391/mbi.2017.8.1.09

Hodgson, J. C., Baylis, S. M., Mott, R., Herrod, A., and Clarke, R. H. (2016). Precision wildlife monitoring using unmanned aerial vehicles. *Scientific Reports* **6**, 22574. doi:10.1038/srep22574

Hogg, S. E., Wang, Y., and Stone, L. (2021). Effectiveness of joint species distribution models in the presence of imperfect detection. *Methods in Ecology and Evolution* **12**, 1458–1474. doi:10.1111/2041-210X.13614

Hope, P., Hernaman, V., Heady, C., Pearce, K., and Bluhm, S. (2018). Our changing climate: Southern Australia rainfall: long-term trends and future projections. doi:10.1016/0168-1923(90)90138-V

Hunter, M. E., Ferrante, J. A., Meigs-Friend, G., and Ulmer, A. (2019). Improving eDNA yield and inhibitor reduction through increased water volumes and multi-filter isolation techniques. *Scientific Reports* **9**, 1–9. doi:10.1038/s41598-019-40977-w

Hunter, M. E., Meigs-Friend, G., Ferrante, J. A., Kamla, A. T., Dorazio, R. M., Diagne, L. K., Luna, F., Lanyon, J. M., and Reid, J. P. (2018). Surveys of environmental DNA (eDNA): a new approach to estimate occurrence in Vulnerable manatee populations. *Endangered Species Research* **35**, 101–111. doi:10.3354/esr00880

Hunter, M. E., Oyler-McCance, S. J., Dorazio, R. M., Fike, J. A., Smith, B. J., Hunter, C. T., Reed, R. N., and Hart, K. M. (2015). Environmental DNA (eDNA) sampling improves occurrence and detection estimates of invasive Burmese pythons. *PLoS ONE* **10**, e0121655. doi:10.1371/journal.pone.0121655

Jackson, W., Argent, R., Bax, N., Bui, E., Clark, G., Coleman, S., Cresswell, I., Emmerson, K., Evans, K., Hibberd, M., Johnston, E., Keywood, M., Klekociuk, A., Mackay, R., Metcalfe, D., Murphy, H., Rankin, A., Smith, D., and Wienecke, B. (2016). Overview: Land-use change, and habitat fragmentation and degradation threaten ecosystems and resilience. In 'Australia state of the environment'. (Australian Government Department of the Environment and Energy: Canberra.) doi:10.4226/94/58b65510c633b

Jarman, S. N., Berry, O., and Bunce, M. (2018). The value of environmental DNA biobanking for long-term biomonitoring. *Nature Ecology & Evolution* **2**, 1192–1193. doi:10.1038/s41559-018-0614-3

Jarvis, M., Closs, G., Dedual, M., Dorsey, L., Fulton, W., Gabrielsson, R., Lintermans, M., and Trotter, M. (2019). The Introduced Trout and Char of Australia and New Zealand. In 'Trout and Char of the World'. (Eds J. . Kershner, J. E. Williams, R. . Gresswell, and J. Lóbon-Cerviá.) pp. 573–604. (American Fisheries Society: Bethesda, Maryland, USA.)

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P. M., Woodcock, P., Edwards, F. A., Larsen, T. H., Hsu, W. W., Benedick, S., Hamer, K. C., Wilcove, D. S., Bruce, C., Wang, X., Levi, T., Lott, M., Emerson, B. C., and Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters* **16**, 1245–1257. doi:10.1111/ele.12162

Jolly, C. J., Dickman, C. R., Doherty, T. S., Eeden, L. M., Geary, W. L., Legge, S. M., Woinarski, J. C. Z., and Nimmo, D. G. (2022). Animal mortality during fire. *Global Change Biology*, 1–13. doi:10.1111/gcb.16044

Keeley, J. E. (2009). Fire intensity, fire severity and burn severity: A brief review and suggested usage. *International Journal of Wildland Fire* **18**, 116–126. doi:10.1071/WF07049

Kelly, R. P., Shelton, A. O., and Gallego, R. (2019). Understanding PCR Processes to Draw Meaningful Conclusions from Environmental DNA Studies. *Scientific Reports* **9**, 1–14. doi:10.1038/s41598-019-48546-x

Kéry, M., and Royle, A. J. (2016). 'Applied Hierachical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS'. (Elsevier: London, UK.)

Klamt, M., Thompson, R., and Davis, J. (2011). Early response of the platypus to climate warming. *Global Change Biology* **17**, 3011–3018. doi:10.1111/j.1365-2486.2011.02472.x

Klobucar, S. L., Rodgers, T. W., and Budy, P. (2017). At the forefront: evidence of the applicability of using environmental DNA to quantify the abundance of fish populations in natural lentic waters with additional sampling considerations. *Canadian Journal of Fisheries and Aquatic Sciences* **74**, 2030–2034. doi:10.1139/cjfas-2017-0114

Klymus, K. E., Richter, C. A., Thompson, N., and Hinck, J. E. (2017). Metabarcoding of environmental DNA samples to explore the use of uranium mine containment ponds as a water source for wildlife. *Diversity* **9**, 1–18. doi:10.3390/d9040054

Koehn, J. D., Balcombe, S. R., Baumgartner, L. J., Bice, C. M., Burndred, K., Ellis, I., Koster, W. M., Lintermans, M., Pearce, L., Sharpe, C., Stuart, I., and Todd, C. R. (2020). What is needed to restore native fishes in Australia's Murray-Darling Basin? *Marine and Freshwater Research* **71**, 1464–1468. doi:10.1071/MF20248

Lacoursière-Roussel, A., Dubois, Y., Normandeau, E., Bernatchez, L., and Adamowicz, S. (2016). Improving herpetological surveys in eastern North America using the environmental DNA method. *Genome* **59**, 991–1007. doi:10.1139/gen-2015-0218

Lahoz-Monfort, J. J., Guillera-Arroita, G., and Tingley, R. (2016). Statistical approaches to account for false-positive errors in environmental DNA samples. *Molecular Ecology Resources* **16**, 673–685. doi:10.1111/1755-0998.12486

Lahoz-Monfort, J. J., Guillera-Arroita, G., and Wintle, B. A. (2014). Imperfect detection impacts the performance of species distribution models. *Global Ecology and Biogeography* **23**, 504–515. doi:10.1111/geb.12138

Lahoz-Monfort, J. J., and Tingley, R. (2018). The technology revolution: improving species detection and monitoring using new tools and statistical methods. In 'Monitoring threatened species and ecological communities'. (Eds S. Legge, D. B. Lindenmayer, N. Robinson, B. C. Scheele, D. Southwell, and B. A. Wintle.) (CSIRO PUBLISHING.)

Langlois, V. S., Allison, M. J., Bergman, L. C., To, T. A., and Helbing, C. C. (2021). The need for robust qPCR-based eDNA detection assays in environmental monitoring and species inventories. *Environmental DNA* **3**, 519–527. doi:10.1002/edn3.164

Laramie, M. B., Pilliod, D. S., Goldberg, C. S., and Strickler, K. M. (2015). Environmental DNA sampling protocol - filtering water to capture DNA from aquatic organisms. *U.S Geological Survey Techniques and Methods* **Book 2**, 15 p. doi:10.3133/TM2A13

Lee, P. L. M. (2017). DNA amplification in the field: move over PCR, here comes LAMP. *Molecular ecology resources* **17**, 138–141. doi:10.1111/1755-0998.12548

Legge, S., Woinarski, J. C. Z., Scheele, B. C., Garnett, S. T., Lintermans, M., Nimmo, D. G., Whiterod, N. S., Southwell, D. M., Ehmke, G., Buchan, A., Gray, J., Metcalfe, D. J., Page, M., Rumpff, L., Leeuwen, S., Williams, D., Ahyong, S. T., Chapple, D. G., Cowan, M., Hossain, M. A., Kennard, M., Macdonald, S., Moore, H., Marsh, J., McCormack, R. B., Michael, D., Mitchell, N., Newell, D., Raadik, T. A., and Tingley, R. (2021). Rapid assessment of the biodiversity impacts of the 2019–2020 Australian megafires to guide urgent management intervention and recovery and lessons for other regions. *Diversity and Distributions* **00**, 1–21. doi:10.1111/ddi.13428

Legge, S., Woinarski, J., Garnett, S., Nimmo, D., Scheele, B., Lintermans, M., Mitchell, N.,

and Ferris, J. (2020). Rapid analysis of impacts of the 2019-20 fires on animal species, and prioritisation of species for management response – preliminary report. Available at: https://www.environment.gov.au/biodiversity/bushfire-recovery/research-and-resources

Leontidou, K., Verseni, C., de Groeve, J., Cristofolini, F., Vokou, D., and Cristofori, A. (2018). Taxonomic identification of airborne pollen from complex environmental samples by DNA metabarcoding: a methodological study for optimizing protocols. *Aerobiologia* **34**, 63–74. doi:10.1101/099481

Lesslie, R., Thackway, R., and Smith, J. (2010). 'A national-level Vegetation Assets, States and Transitions (VAST) dataset for Australia (version 2.0)'. (Bureau of Rural Sciences: Canberra.)

Lindsay, J. (2014). The Whitebox Geospatial Analysis Tools project and open-access GIS. In 'Proceedings of the GIS Research UK 22nd Annual Conference'. (The University of Glasgow.) doi:10.13140/RG.2.1.1010.8962

Lintermans, M., Geyle, H. M., Beatty, S., Brown, C., Ebner, B. C., Freeman, R., Hammer, M. P., Humphreys, W. F., Kennard, M. J., Kern, P., Martin, K., Morgan, D. L., Raadik, T. A., Unmack, P. J., Wager, R., Woinarski, J. C. Z., and Garnett, S. T. (2020). Big trouble for little fish: Identifying Australian freshwater fishes in imminent risk of extinction. *Pacific Conservation Biology* **26**, 365–377. doi:10.1071/PC19053

Loeza-Quintana, T., Abbott, C. L., Heath, D. D., Bernatchez, L., and Hanner, R. H. (2020). Pathway to Increase Standards and Competency of eDNA Surveys (PISCeS)—Advancing collaboration and standardization efforts in the field of eDNA. *Environmental DNA* **2**, 255–260. doi:10.1002/edn3.112

Loeza-Quintana, T., Abbott, C. L., Heath, D. D., Bernatchez, L., and Hanner, R. H. (2020). Pathway to Increase Standards and Competency of eDNA Surveys (PISCeS)—Advancing collaboration and standardization efforts in the field of eDNA. *Environmental DNA* **2**, 255–260. doi:10.1002/edn3.112

Lugg, W. H., Griffiths, J., van Rooyen, A. R., Weeks, A. R., and Tingley, R. (2017). Optimal survey designs for environmental DNA sampling. *Methods in Ecology and Evolution* **9**, 1049–1059. doi:10.1111/2041-210X.12951

Lyet, A., Pellissier, L., Valentini, A., Dejean, T., Hehmeyer, A., and Naidoo, R. (2021).

eDNA sampled from stream networks correlates with camera trap detection rates of terrestrial mammals. *Scientific Reports* **11**, 1–14. doi:10.1038/s41598-021-90598-5

Lyon, J. P., and O'Connor, J. P. (2008). Smoke on the water: Can riverine fish populations recover following a catastrophic fire-related sediment slug? *Austral Ecology* **33**, 794–806. doi:10.1111/j.1442-9993.2008.01851.x

Mächler, E., Deiner, K., Spahn, F., and Altermatt, F. (2016). Fishing in the Water: Effect of Sampled Water Volume on Environmental DNA-Based Detection of Macroinvertebrates. *Environmental Science and Technology* **50**, 305–312. doi:10.1021/acs.est.5b04188

Marchant, R., and Grant, T. R. (2015). The productivity of the macroinvertebrate prey of the platypus in the upper Shoalhaven River, New South Wales. *Marine and Freshwater Research* **66**, 1128–1137. doi:10.1071/MF14301

Marques, P., Zandonà, E., Mazzoni, R., and El-Sabaawi, R. (2021). Individual variation in feeding morphology, not diet, can facilitate the success of generalist species in urban ecosystems. *Ecology and Evolution*, 18342–18356. doi:10.1002/ece3.8425

Mcclenaghan, B., Compson, Z. G., and Hajibabaei, M. (2020). Validating metabarcoding-based biodiversity assessments with multi-species occupancy models : A case study using coastal marine. *PLoS ONE* **15**, 1–17. doi:10.1371/journal.pone.0224119

McColl-Gausden, E. F., Weeks, A. R., Coleman, R. A., Robinson, K. L., Song, S., Raadik, T. A., and Tingley, R. (2021). Multispecies models reveal that eDNA metabarcoding is more sensitive than backpack electrofishing for conducting fish surveys in freshwater streams. *Molecular Ecology* **30**, 3111–3126. doi:10.1111/mec.15644

McColl-Gausden, E. F., Weeks, A. R., and Tingley, R. (2020). A field ecologist's guide to environmental DNA sampling in freshwater environments. *Australian Zoologist* **40**, 641–651. doi:10.7882/AZ.2019.025

Mcdevitt, A. D., Benvenuto, C., Guimarães, N., Samuel, S., Sparnenn, A. O., Mariani, S., Wangensteen, O. S., and Coscia, I. (2019). Environmental DNA metabarcoding as an effective and rapid tool for fish monitoring in canals. *Journal of Fish Biology* **95**, 679–682. doi:10.1111/jfb.14053

Mcelroy, M. E., Dressler, T. L., Titcomb, G. C., Wilson, E. A., Deiner, K., Dudley, T. L.,

Eliason, E. J., Evans, N. T., Gaines, S. D., Lafferty, K. D., Lamberti, G. A., Li, Y., Lodge, D. M., Love, M. S., Mahon, A. R., Pfrender, M. E., Renshaw, M. A., Selkoe, K. A., and Allan, E. A. (2020). Calibrating Environmental DNA Metabarcoding to Conventional Surveys for Measuring Fish Species Richness. *frontiers in ecology and Evolution* **8**, 1–12. doi:10.3389/fevo.2020.00276

McLachlan-Troup, T. A., Dickman, C. R., and Grant, T. R. (2010). Diet and dietary selectivity of the platypus in relation to season, sex and macroinvertebrate assemblages. *Journal of Zoology* **280**, 237–246. doi:10.1111/j.1469-7998.2009.00645.x

Metcalfe, D., and Bui, E. (2016). Land: Regional and landscape-scale pressures: Land clearing. In 'Australia state of the environment 2016'. (Australian Government Department of the Environment and Energy: Canberra.) doi:10.4226/94/58b6585f94911

Meyer, R. S., Curd, E. E., Schweizer, T., Gold, Z., Ramos, D. R., Shirazi, S., Kandlikar, G., Kwan, W. Y., Lin, M., Freise, A., Moberg-Parker, J., Ramos, M. M., Shapiro, B., Sexton, J. P., Pipes, L., Vedrenne, A. G., Mejia, M. P., Aronson, E. L., Moore, T., Nielsen, R., Lewin, H., Barber, P., Wall, J., Kraft, N., and Wayne, R. K. (2019). The California environmental DNA "CALeDNA" program. *bioRxiv*, 1–33. doi:10.1101/503383

Milione, M., and Harding, E. (2009). Habitat use by platypus (Ornithorhynchus anatinus) in a modified Australian Wet Tropics catchment, north-eastern Queensland. *Australian Mammalogy* **31**, 35–46. doi:10.1071/AM08110

Morrongiello, J. R., Beatty, S. J., Bennett, J. C., Crook, D. A., Ikedife, D. N. E. N., Kennard, M. J., Kerezsy, A., Lintermans, M., McNeil, D. G., Pusey, B. J., and Rayner, T. (2011). Climate change and its implications for Australia's freshwater fish. *Marine and Freshwater Research* **62**, 1082–1098. doi:10.1071/MF10308

Neris, J., Santin, C., Lew, R., Robichaud, P. R., Elliot, W. J., Lewis, S. A., Sheridan, G., Rohlfs, A. M., Ollivier, Q., Oliveira, L., and Doerr, S. H. (2021). Designing tools to predict and mitigate impacts on water quality following the Australian 2019/2020 wildfires: Insights from Sydney's largest water supply catchment. *Integrated Environmental Assessment and Management* **17**, 1151–1161. doi:10.1002/ieam.4406

Nevers, M. B., Przybyla-Kelly, K., Shively, D., Morris, C. C., Dickey, J., and Byappanahalli, M. N. (2020). Influence of sediment and stream transport on detecting a source of

environmental DNA. *PLoS ONE* **15**, 1–21. doi:10.1371/journal.pone.0244086

Nguyen, B., Id, T., Tanase, M. A., Bennett, L. T., and Aponte, C. (2020). High-severity wildfires in temperate Australian forests have increased in extent and aggregation in recent decades. *Plos One* **15**, e0242484. doi:10.1371/journal.pone.0242484

Nukazawa, K., Hamasuna, Y., and Suzuki, Y. (2018). Simulating the Advection and Degradation of the Environmental DNA of Common Carp along a River. *Environmental Science & Technology* **52**, 10562–10570. doi:10.1021/acs.est.8b02293

Nyman, P., Box, W. A. C., Stout, J. C., Sheridan, G. J., Keesstra, S. D., Lane, P. N. J., and Langhans, C. (2020). Debris-flow-dominated sediment transport through a channel network after wildfire. *Earth Surface Processes and Landforms* **45**, 1155–1167. doi:10.1002/esp.4785

Olden, J. D., Kennard, M. J., and Pusey, B. J. (2008). Species invasions and the changing biogeography of Australian freshwater fishes. *Global Ecology and Biogeography* **17**, 25–37. doi:10.1111/j.1466-8238.2007.00340.x

Olds, B. P., Jerde, C. L., Renshaw, M. A., Li, Y., Evans, N. T., Turner, C. R., Deiner, K., Mahon, A. R., Brueseke, M. A., Shirey, P. D., Pfrender, M. E., Lodge, D. M., and Lamberti, G. A. (2016). Estimating species richness using environmental DNA. *Ecology and Evolution* **6**, 4214–4226. doi:10.1002/ece3.2186

Peixoto, S., Chaves, C., Velo-Antón, G., Beja, P., and Egeter, B. (2021). Species detection from aquatic eDNA: Assessing the importance of capture methods. *Environmental DNA* **3**, 435–448. doi:10.1002/edn3.130

Piggott, M. P. (2017). An environmental DNA assay for detecting Macquarie perch, Macquaria australasica. *Conservation Genetics Resources* **9**, 257–259. doi:10.1007/s12686-016-0666-0

Piggott, M. P., Banks, S. C., Broadhurst, B. T., Fulton, C. J., and Lintermans, M. (2021). Comparison of traditional and environmental DNA survey methods for detecting rare and abundant freshwater fish. *Aquatic Conservation: Marine and Freshwater Ecosystems* **31**, 173–184. doi:10.1002/aqc.3474

Piñol, J., Senar, M. A., and Symondson, W. O. C. (2019). The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can

be quantitative. *Molecular Ecology* **28**, 407–419. doi:10.1111/mec.14776

Plante, F., Bourgault, P., Dubois, Y., and Bernatchez, L. (2021). Environmental DNA as a detection and quantitative tool for stream-dwelling salamanders: A comparison with the traditional active search method. *Environmental DNA* **3**, 1128–1141. doi:10.1002/edn3.233

Plough, L. V, Ogburn, M. B., Fitzgerald, C. L., Geranio, R., Marafino, G. A., and Richie, K. D. (2018). Environmental DNA analysis of river herring in Chesapeake Bay : A powerful tool for monitoring threatened keystone species. *Plos One*, 1–25. doi:10.5061/dryad.n60t8b3

Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., Roset, N., Schabuss, M., Zornig, H., and Dejean, T. (2018). Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Scientific Reports* **8**, 10361. doi:10.1038/s41598-018-28424-8

Pukk, L., Kanefsky, J., Heathman, A. L., Weise, E. M., Nathan, L. R., Herbst, S. J., Sard, N. M., Scribner, K. T., and Robinson, J. D. (2021). eDNA metabarcoding in lakes to quantify influences of landscape features and human activity on aquatic invasive species prevalence and fish community diversity. *Diversity and Distributions* **27**, 2016–2031. doi:10.1111/ddi.13370

QGIS Geographic Information System (2021). QGIS.org.

R Development Core Team (2022). R: A Language and Environment for Statistical Computing.

Ratcliffe, F. C., Uren Webster, T. M., Garcia de Leaniz, C., and Consuegra, S. (2021). A drop in the ocean: Monitoring fish communities in spawning areas using environmental DNA. *Environmental DNA* **3**, 43–54. doi:10.1002/edn3.87

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J., Kidd, K. A., MacCormack, T. J., Olden, J. D., Ormerod, S. J., Smol, J. P., Taylor, W. W., Tockner, K., Vermaire, J. C., Dudgeon, D., and Cooke, S. J. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews* **94**, 849–873. doi:10.1111/brv.12480

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., and Coissac, E. (2011).

EcoPrimers: Inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research* **39**, 1–11. doi:10.1093/nar/gkr732

Rice, C. J., Larson, E. R., and Taylor, C. A. (2018). Environmental DNA detects a rare large river crayfish but with little relation to local abundance. *Freshwater Biology* **63**, 443–455. doi:10.1111/fwb.13081

Robson, B. J., Chester, E. T., Matthews, T. G., and Johnston, K. (2018). Post-wildfire recovery of invertebrate diversity in drought-affected headwater streams. *Aquatic Sciences* **80**, 1–15. doi:10.1007/s00027-018-0570-7

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH : a versatile open source tool for metagenomics. *PeerJ* **4**. doi:10.7717/peerj.2584

Rohweder, D. A., and Baverstock, P. R. (1999). Distribution of platypus Ornithorhynchus anatinus in the Richmond river catchment, northern New South Wales. *Australian Zoologist* **31**, 30–37. doi:10.7882/AZ.1999.005

Sard, N. M., Herbst, S. J., Nathan, L., Uhrig, G., Kanefsky, J., Robinson, J. D., and Scribner, K. T. (2019). Comparison of fish detections, community diversity, and relative abundance using environmental DNA metabarcoding and traditional gears. *Environmental DNA* **1**, 368–384. doi:https://doi.org/10.1002/edn3.38

Schenekar, T., Schletterer, M., Lecaudey, L. A., and Weiss, S. J. (2020). Reference databases, primer choice, and assay sensitivity for environmental metabarcoding: Lessons learnt from a re-evaluation of an eDNA fish assessment in the Volga headwaters. *River Research and Applications*, 1004–1013. doi:10.1002/rra.3610

Schmelzle, M. C., and Kinziger, A. P. (2016). Using occupancy modelling to compare environmental DNA to traditional field methods for regional-scale monitoring of an endangered aquatic species. *Molecular Ecology Resources* **16**, 895–908. doi:10.1111/1755-0998.12501

Schmidt, B. R., Kéry, M., Ursenbacher, S., Hyman, O. J., and Collins, J. P. (2013). Site occupancy models in the analysis of environmental DNA presence/absence surveys: A case study of an emerging amphibian pathogen. *Methods in Ecology and Evolution* **4**, 646–653. doi:10.1111/2041-210X.12052

Scientific Advisory Committee (2020). FLORA AND FAUNA GUARANTEE -

SCIENTIFIC ADVISORY COMMITTEE PRELIMINARY RECOMMENDATION ON A NOMINATION FOR LISTING Ornithorhynchus anatinus.

Serena, M. (1994). Use of time and space by platypus (Ornithorhynchus anatinus: Monotremata) along a Victorian stream. *Journal of Zoology* **232**, 117–131. doi:10.1111/j.1469-7998.1994.tb01563.x

Serena, M., Thomas, J. L., Williams, G. A., and Officer, R. C. E. (1998). Use of stream and river habitats by the platypus, Ornithorhynchus anatinus, in an urban fringe environment. *Australian Journal of Zoology* **46**, 267–282. doi:10.1071/ZO98034

Serena, M., and Williams, G. A. (2004). Progress report: An experimental reintroduction of platypus (Ornithorhynchus anatinus) to Cardinia Creek April-June 2004. Whittlesea.

Serena, M., Williams, G. A., Weeks, A. R., and Griffiths, J. (2014). Variation in platypus (Ornithorhynchus anatinus) life-history attributes and population trajectories in urban streams. *Australian Journal of Zoology* **62**, 223–234. doi:10.1071/ZO13079

Serena, M., Worley, M., Swinnerton, M., and Williams, G. A. (2001). Effect of food availability and habitat on the distribution of platypus (ornithorhynchus anatinus) foraging activity. *Australian Journal of Zoology* **49**, 263–277. doi:10.1071/ZO00089

Shakesby, R. A., Wallbrink, P. J., Doerr, S. H., English, P. M., Chafer, C. J., Humphreys, G. S., Blake, W. H., and Tomkins, K. M. (2007). Distinctiveness of wildfire effects on soil erosion in south-east Australian eucalypt forests assessed in a global context. *Forest Ecology and Management* **238**, 347–364. doi:10.1016/j.foreco.2006.10.029

Shaw, J. L. A., Clarke, L. J., Wedderburn, S. D., Barnes, T. C., Weyrich, L. S., and Cooper, A. (2016). Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biological Conservation* **197**, 131–138. doi:10.1016/j.biocon.2016.03.010

Shaw, J. L. A., Weyrich, L., and Cooper, A. (2017). Using environmental (e)DNA sequencing for aquatic biodiversity surveys: a beginner's guide. *Marine and Freshwater Research* **68**, 20–33. doi:http://dx.doi.org/10.1071/MF15361

Shea, C. P., Eaton, M. J., and MacKenzie, D. I. (2019). Implementation of an occupancy-based monitoring protocol for a widespread and cryptic species, the New England cottontail (Sylvilagus transitionalis). *Wildlife Research* **46**, 222–235.

doi:10.1071/WR18058

Shehzad, W., Mccarthy, T. M., Pompanon, F., Purevjav, L., Riaz, T., and Taberlet, P. (2012). Prey Preference of Snow Leopard (Panthera uncia) in. *PLoS ONE* **7**, e32104. doi:10.5061/dryad.bj376f61

Shepard, B. B., Nelson, L. M., Taper, M. L., and Zale, A. V. (2014). Factors Influencing Successful Eradication of Nonnative Brook Trout from Four Small Rocky Mountain Streams Using Electrofishing. *North American Journal of Fisheries Management* **34**, 988–997. doi:10.1080/02755947.2014.942042

Shogren, A. J., Tank, J. L., Andruszkiewicz, E., Olds, B., Mahon, A. R., Jerde, C. L., and Bolster, D. (2017). Controls on eDNA movement in streams : Transport, Retention, and Resuspension. *Scientific Reports* **7**, 5065. doi:10.1038/s41598-017-05223-1

Shogren, A. J., Tank, J. L., and Egan, S. P. (2019). Riverine distribution of mussel environmental DNA reflects a balance among density , transport , and removal processes. , 1–13. doi:10.1111/fwb.13319

Sigsgaard, E. E., Nielsen, I. B., Bach, S. S., Lorenzen, E. D., Robinson, D. P., Knudsen, S. W., Pedersen, M. W., Jaidah, M. Al, Orlando, L., Willerslev, E., Møller, P. R., and Thomsen, P. F. (2016). Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. *Nature Ecology & Evolution* **1**, 0004. doi:10.1038/s41559-016-0004

Silva, L. G. M., Doyle, K. E., Duffy, D., Humphries, P., Horta, A., and Baumgartner, L. J. (2020). Mortality events resulting from Australia's catastrophic fires threaten aquatic biota. *Global Change Biology* **26**, 5345–5350. doi:10.1111/gcb.15282

Simpfendorfer, C., Kyne, P., Noble, T., Goldsbury, J., Basiita, R., Lindsay, R., Shields, A., Perry, C., and Jerry, D. (2016). Environmental DNA detects Critically Endangered largetooth sawfish in the wild. *Endangered Species Research* **30**, 109–116. doi:10.3354/esr00731

Smart, A. S., Tingley, R., Weeks, A. R., Van Rooyen, A. R., and McCarthy, M. A. (2015). Environmental DNA sampling is more sensitive than a traditional survey technique for detecting an aquatic invader. *Ecological Applications* **25**, 1944–1952. doi:10.1890/14-1751.1

Smucker, K. M., Hutto, R. L., and Steele, B. M. (2005). Changes in bird abundance after wildfire: Importance of fire severity and time since fire. *Ecological Applications* **15**, 1535–1549. doi:10.1890/04-1353

Southwell, D., Legge, S., Woinarski, J., Lindenmayer, D., Lavery, T., and Wintle, B. (2021). Design considerations for rapid biodiversity reconnaissance surveys and long-term monitoring to assess the impact of wildfire. *Diversity and Distributions* **00**, 1–12. doi:10.1111/ddi.13427

Souza, L., Bunn, W. A., Simberloff, D., Lawton, R. M., and Sanders, N. J. (2011). Biotic and abiotic influences on native and exotic richness relationship across spatial scales: Favourable environments for native species are highly invasible. *Functional Ecology* **25**, 1106–1112. doi:10.1111/j.1365-2435.2011.01857.x

Spear, M. J., Embke, H. S., Krysan, P. J., and Vander Zanden, M. J. (2021). Application of eDNA as a tool for assessing fish population abundance. *Environmental DNA* **3**, 83–91. doi:10.1002/edn3.94

Spens, J., Evans, A. R., Halfmaerten, D., Knudsen, S. W., Sengupta, M. E., Mak, S. S. T., Sigsgaard, E. E., and Hellström, M. (2017). Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol : advantage of enclosed filter. *Methods in Ecology and Evolution* **8**, 635–645. doi:10.1111/2041-210X.12683

Stratton, C., Sepulveda, A. J., Hoegh, A., and Stratton, C. (2020). msocc : Fit and analyse computationally efficient multi-scale occupancy models in r. *Methods in Ecology and Evolution* **11**, 1113–1120. doi:10.1111/2041-210X.13442

Strickland, G. J., and Roberts, J. H. (2019). Utility of eDNA and occupancy models for monitoring an endangered fish across diverse riverine habitats. *Hydrobiologia* **826**, 129–144. doi:10.1007/s10750-018-3723-8

Sutela, T., Vehanen, T., and Jounela, P. (2020). Longitudinal patterns of fish assemblages in European boreal streams. *Hydrobiologia* **847**, 3277–3290. doi:10.1007/s10750-020-04330-x

Taberlet, P., Coissac, E., Hajibabaei, M., and Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology* **21**, 1789–1793. doi:10.1111/j.1365-294X.2012.05542.x

Tiago, P., Ceia-Hasse, A., Marques, T. A., Capinha, C., and Pereira, H. M. (2017). Spatial distribution of citizen science casuistic observations for different taxonomic groups. *Scientific Reports* **7**, 1–9. doi:10.1038/s41598-017-13130-8

Tillotson, M. D., Kelly, R. P., Duda, J. J., Hoy, M., Kralj, J., and Quinn, T. P. (2018). Concentrations of environmental DNA (eDNA) reflect spawning salmon abundance at fine spatial and temporal scales. *Biological Conservation* **220**, 1–11. doi:10.1016/j.biocon.2018.01.030

Tingley, R., Coleman, R., Gecse, N., van Rooyen, A., and Weeks, A. (2021). Accounting for false positive detections in occupancy studies based on environmental DNA: A case study of a threatened freshwater fish (Galaxiella pusilla). *Environmental DNA* **3**, 388–397. doi:10.1002/edn3.124

Tingley, R., Greenlees, M., Oertel, S., van Rooyen, A. ., and Weeks, A. . (2019). Environmental DNA sampling as a surveillance tool for cane toad Rhinella marina introductions on offshore islands. *Biological Invasions* **21**, 1–6. doi:10.1007/s10530-018-1810-4

Tingley, R., Phillips, B. L., and Shine, R. (2011). Establishment Success of Introduced Amphibians Increases in the Presence of Congeneric Species. *American Naturalist* **177**, 382–388. doi:10.1086/658342

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G. H., Geniez, P., Pont, D., Argillier, C., Baudoin, J. M., Peroux, T., Crivelli, A. J., Olivier, A., Acqueberge, M., Le Brun, M., Møller, P. R., Willerslev, E., and Dejean, T. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology* **25**, 929–942. doi:10.1111/mec.13428

Verkaik, I., Prat, N., Rieradevall, M., Reich, P., and Lake, P. S. (2014). Effects of bushfire on macroinvertebrate communities in south-east Australian streams affected by a megadrought. *Marine and Freshwater Research* **65**, 359–369. doi:10.1071/MF13039

Verkaik, I., Vila-Escalé, M., Rieradevall, M., Baxter, C. V., Lake, P. S., Minshall, G. W., Reich, P., and Prat, N. (2015). Stream macroinvertebrate community responses to fire: Are they the same in different fire-prone biogeographic regions? *Freshwater Science* **34**, 1527–1541. doi:10.1086/683370

Vieira, N. K. M., Clements, W. H., Guevara, L. S., and Jacobs, B. F. (2004). Resistance and resilience of stream insect communities to repeated hydrologic disturbances after a wildfire. *Freshwater Biology* **49**, 1243–1259. doi:10.1111/j.1365-2427.2004.01261.x

Visintin, C., Briscoe, N. J., Woolley, S. N. C., Lentini, P. E., Tingley, R., Wintle, B. A., and Golding, N. (2020). steps: Software for spatially and temporally explicit population simulations. *Methods in Ecology and Evolution* **11**, 596–603. doi:10.1111/2041-210X.13354

Ward, M., Tulloch, A. I. T., Radford, J. Q., Williams, B. A., Reside, A. E., Macdonald, S. L., Mayfield, H. J., Maron, M., Possingham, H. P., Vine, S. J., O'Connor, J. L., Massingham, E. J., Greenville, A. C., Woinarski, J. C. Z., Garnett, S. T., Lintermans, M., Scheele, B. C., Carwardine, J., Nimmo, D. G., Lindenmayer, D. B., Kooyman, R. M., Simmonds, J. S., Sonter, L. J., and Watson, J. E. M. (2020). Impact of 2019–2020 mega-fires on Australian fauna habitat. *Nature Ecology and Evolution* **4**, 1321–1326. doi:10.1038/s41559-020-1251-1

Warton, D. I., Stoklosa, J., Guillera-Arroita, G., MacKenzie, D. I., and Welsh, A. H. (2017). Graphical diagnostics for occupancy models with imperfect detection. *Methods in Ecology and Evolution* **8**, 408–419. doi:10.1111/2041-210X.12761

West, K., Travers, M. J., Stat, M., Harvey, E. S., Richards, Z. T., DiBattista, J. D., Newman, S. J., Harry, A., Skepper, C. L., Heydenrych, M., and Bunce, M. (2021). Large-scale eDNA metabarcoding survey reveals marine biogeographic break and transitions over tropical north-western Australia. *Diversity and Distributions* **00**, 1–16. doi:10.1111/ddi.13228

Wilcox, T. M., McKelvey, K. S., Young, M. K., Sepulveda, A. J., Shepard, B. B., Jane, S. F., Whiteley, A. R., Lowe, W. H., and Schwartz, M. K. (2016). Understanding environmental DNA detection probabilities: A case study using a stream-dwelling char Salvelinus fontinalis. *Biological Conservation* **194**, 209–216. doi:10.1016/j.biocon.2015.12.023

Wilkinson, D. P., Golding, N., Guillera-Arroita, G., Tingley, R., and McCarthy, M. A. (2021). Defining and evaluating predictions of joint species distribution models. *Methods in Ecology and Evolution* **12**, 394–404. doi:10.1111/2041-210X.13518

Wilkinson, S., Wallbrink, P., Shakesby, R., Blake, W., and Doerr, S. (2007). Impacts on

water quality by sediments and nutrients released during extreme bushfires: Summary of findings.

Williams, B. K., and Brown, E. D. (2019). Sampling and analysis frameworks for inference in ecology. *Methods in Ecology and Evolution* **10**, 1832–1842. doi:10.1111/2041-210X.13279

Willoughby, J. R., Wijayawardena, B. K., Sundaram, M., Swihart, R. K., and DeWoody, J. A. (2016). The importance of including imperfect detection models in eDNA experimental design. *Molecular Ecology Resources* **16**, 837–844. doi:10.1111/1755-0998.12531

Winter, D. J. (2017). rentrez : An R package for the NCBI eUtils API. *The R Journal* **9**, 520–526.

Woinarski, J., and Burbidge, A. A. (2016). Ornithorhynchus anatinus. *The IUCN Red List of Threatened Species*. doi:http://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T40488A21964009.en

Wood, C. M. (2021). Optimizing landscape-scale monitoring programmes to detect the effects of megafires. *Diversity and Distributions* **00**, 1–14. doi:10.1111/ddi.13308

Wood, S. A., Pochon, X., Laroche, O., Ammon, U. Von, Adamson, J., and Zaiko, A. (2019). A comparison of droplet digital polymerase chain reaction (PCR), quantitative PCR and metabarcoding for species - specific detection in environmental DNA. *Molecular Ecology Resources*, 1407–1419. doi:10.1111/1755-0998.13055

Woon, P. M. . (1995). Habitat preferences of platypus. University of Sydney.

Yates, M. C., Fraser, D. J., and Derry, A. M. (2019). Meta - analysis supports further refinement of eDNA for monitoring aquatic species - specific abundance in nature. *Environmental DNA*, 5–13. doi:10.1002/edn3.7

Zajicek, P., and Wolter, C. (2018). The gain of additional sampling methods for the fish-based assessment of large rivers. *Fisheries Research* **197**, 15–24. doi:10.1016/j.fishres.2017.09.018

Zurell, D., Pollock, L. J., and Thuiller, W. (2018). Do joint species distribution models reliably detect interspecific interactions from co-occurrence data in homogenous environments? *Ecography* **41**, 1812–1819. doi:10.1111/ecog.03315

Appendix A
Published in Australian Zoologist

A field ecologist's guide to environmental DNA sampling in freshwater environments

Emily F. McColl-Gausden[1], Andrew R. Weeks[1,2] and Reid Tingley[1]

1. School of BioSciences, The University of Melbourne, Parkville, Victoria 3010, Australia

2. cesar Pty Ltd, 293 Royal Pde, Parkville, VIC 3052, Australia

Corresponding author: mccolle@student.unimelb.edu.au

Keywords: detection, eDNA, freshwater, monitoring, occupancy

Abstract

Environmental DNA, or eDNA —DNA shed from organisms and extracted from environmental samples— is an emerging survey technique that has the potential to transform biodiversity monitoring in freshwater ecosystems. We provide a brief overview of the primary methodological aspects of eDNA sampling which ecologists should consider before taking environmental samples in the field. We outline five key methodological considerations: (i) targeting single-species vs multiple-species; (ii) where and when to sample; (iii) how much water to collect; (iv) how many samples to take; and (v) recognising potential sources of false positives. The need to account for false negatives and false positives in eDNA surveys, and the power of site occupancy-detection models in accounting for imperfect detection, is also discussed.

Introduction

Understanding spatial changes in biodiversity patterns is important for informing conservation efforts, but landscape-level data on species distributions can be logistically and financially difficult to collect. Environmental DNA (eDNA) sampling —the detection of extra-organismal DNA sourced from environmental samples without any obvious biological source material being present (Taberlet *et al.* 2012)— is an emerging survey technique that could, for many species, facilitate efficient and cost-effective collection of landscape-level data. DNA can be shed into the environment from various sources, including skin cells, mucous, faeces, or even individual hairs. Previous studies have shown samples of water, snow, soil, and air can be used to determine species presence-absence at a site (Dalén *et al.* 2007; Andersen *et al.* 2012; Lugg *et al.* 2017; Leontidou *et al.* 2018). In this brief review, we focus on the detection of eDNA presence-absence in freshwater ecosystems, which are globally imperilled and particularly amenable to eDNA sampling. Indeed, freshwater ecology is an area in which significant technical and practical progress has been made in the field of eDNA sampling (Cristescu and Hebert 2018), although similar progress is also starting to occur in marine environments (Foote *et al.* 2012; Thomsen *et al.* 2012; Kelly *et al.* 2014).

The first step in assessing the suitability of any emerging technology, such as eDNA sampling, is to compare the results it produces against traditional sampling methods (Lahoz-Monfort and Tingley 2018). As with other emerging technologies in ecology, such as drones (Hodgson *et al.* 2016), thermal cameras (Goodenough *et al.* 2017), and detector dogs (Cristescu *et al.* 2015),

eDNA sampling in aquatic environments has been compared to traditional survey methods, such as trapping (Shaw *et al.* 2016), aural surveys (Valentini *et al.* 2016), and electrofishing (Evans *et al.* 2017). In many aquatic systems tested to date, eDNA sampling has been capable of detecting species or communities more effectively than traditional methods (see Lugg *et al.* 2017). Environmental DNA sampling lends itself to baseline data collection on species distributions, as well as routine monitoring programs that aim to track changes in species distributions over space and time. If more detailed data on population or individual health (e.g., reproductive output, juvenile recruitment, sex, genetic variation, abundance) are required, eDNA can highlight sampling locations for more traditional, time-intensive capture methods. However, eDNA methods also show some promise in this regard (providing population- and individual-based information), particularly as technologies improve (Sigsgaard *et al.* 2016; Bylemans *et al.* 2017).

Environmental DNA sampling has many elements that make it a promising survey method. It allows for landscape-level data to be collected cost-efficiently for many species without invasive sampling (e.g., capturing), or even sighting animals. Thus, eDNA sampling largely eliminates concerns around animal ethics. An additional benefit of eDNA sampling is that it reduces health and safety concerns for researchers and employers. For example, eDNA sampling eliminates high-risk activities, such as entering waterways to check fyke nets for fish and platypuses (Serena 1994) or conducting backpack or boat electrofishing surveys for fish (Wilcox *et al.* 2016). This technique can potentially detect cryptic or rare species that evade capture with other methods, and detect species at life stages that were previously difficult to detect or distinguish between (Dejean *et al.* 2012).

Despite these benefits, using eDNA sampling in biodiversity surveys is currently more complex than using most traditional survey methods. This is largely because eDNA sampling is a relatively new monitoring technique and standardised approaches have not been developed (see Goldberg *et al.* 2016). Thus, our objective in this paper is to provide a brief overview of the primary methodological aspects of eDNA sampling that ecologists who are interested in using eDNA should consider before taking environmental samples. Because our work is aimed at practitioners collaborating with a fully-equipped genetics laboratory, we do not consider other important, but more technical considerations, such as DNA extraction method, gene region, primer selection, and marker specificity (see Freeland 2017). Nor do we touch on quality control issues, such as field and laboratory negative controls, mock community positive

controls, index switching (exclusive to metabarcoding and pooling samples), OTU thresholds, or read depth (Deiner *et al.* 2017). For excellent reviews incorporating such considerations see Bohmann *et al.* (2014), Cristescu and Hebert (2018), Deiner *et al.* (2017), and Rees *et al.* (2014).

Our review is structured around five key methodological considerations: (i) targeting single species vs multiple species or communities; (ii) where and when to sample; (iii) how much water; (iv) how many samples; and (v) mitigating false positive detections (Figure A.1). We conclude with a discussion of the need to consider imperfect detection in eDNA surveys, and briefly highlight the power of species occupancy detection models (SODM) in this regard.

Figure A.1 Key methodological considerations for eDNA sampling in the field, with notes on relevant methodological decisions.

Targeting single species vs multiple species

Environmental DNA detection methods can be divided into two broad categories: single- and multi-species methods. Single-species methods focus on eDNA quantification from a single target species using probes that are specific to the species or population(s) of interest. Most single-species studies use real-time quantitative polymerase chain reaction (qPCR), although traditional PCR (Goldberg *et al.* 2011) and Droplet Digital PCR (ddPCR) (Doi *et al.* 2015) have also been used. Single-species detection methods have been used broadly to successfully detect rare (Laramie *et al.* 2015; Schmelzle and Kinziger 2016; Simpfendorfer *et al.* 2016) and invasive (Hunter *et al.* 2015; Hinlo *et al.* 2017; Klymus *et al.* 2017; Tingley *et al.* 2018) species in aquatic environments. Multi-species detection methods —also known as metabarcoding— take a broader, community-focused approach. High-throughput next generation sequencing (NGS) technologies for DNA sequencing enable all species from one or more target groups (e.g., fish, amphibians, or decapods) to be identified (Taberlet *et al.* 2012).

Whether to adopt a single- or multi-species eDNA approach depends primarily on the specific aims of the study. Clearly, a single-species approach is not appropriate when the scientific or management objective pertains to community-level patterns. But given the rich amount of data afforded by metabarcoding, why would one ever adopt a single-species approach, even if a study intends to focus solely on a single species? The answer to this question lies primarily in the relative complexity, sensitivity, and cost-efficiency of each eDNA approach.

Single-species approaches require species-specific primers/probes for a target DNA region, whereas metabarcoding requires primers designed to bind to conserved DNA regions across the target group. The ability to distinguish species within the target group then relies on a reference library of known sequences matched to species, with which the sequenced eDNA data can be compared. Publicly available sequence databases, such as GenBank (Clark *et al.* 2016), can be used to generate reference libraries at a broad level, but often native species are poorly represented in these databases. Therefore, a local reference library, consisting of genetic sequences of each species from the study area (as in Valentini *et al.* 2016), is recommended; this will also help account for any intraspecific variation within the gene region being sequenced. Once this library has been established, it can be used by future studies; however, its initial development can be expensive and time-consuming relative to the resources needed to design a species-specific primer. Different target gene regions will also provide different levels of species resolution, and this needs to be considered when undertaking metabarcoding

surveys, as some gene regions may not differentiate closely related species (Bylemans *et al.* 2018). Primer bias can also result in some species not being detected in a sample when they may be present in the environment, as some species may amplify more readily than others (Elbrecht and Leese 2017).

Metabarcoding requires bioinformatic pipelines to demultiplex samples —assigning sequences to samples— and assign genetic sequences (or haplotypes) to species (Coissac *et al.* 2012). Once again, these pipelines can be reused once developed, but the bioinformatic tools needed for metabarcoding are much more labour-intensive and require a different level of expertise compared to a single-species approach, which is much more routine in a molecular laboratory.

Few studies have directly compared the sensitivity of single- vs multi-species eDNA assays for individual species, although Harper *et al.* (2018) demonstrated that qPCR resulted in greater detectability for Great Crested Newts (*Triturus cristatus*) relative to metabarcoding. Nonetheless, this increased sensitivity needs to be evaluated against a study's aims and the nature of the study system. In some cases, accepting a slightly lower detection probability for a target species may be worthwhile if additional data on biotic interactions (e.g., presence of prey, competitors, predators, disease) would change scientific inferences or management decisions. In other cases, such as determining whether an endangered species is present at a site for an environmental impact assessment, a potentially less-sensitive metabarcoding assay may not be the optimal choice. There may be situations where combining approaches (metabarcoding and a single-species assay) is also appealing, particularly where sensitivity for the target species is paramount in the survey design, but there is also a need for the additional data on community composition. The advantage here is that sampling and DNA extraction have already been undertaken for one method; there is therefore a significant reduction in cost for the second method.

Single-species eDNA assays using qPCR or ddPCR have revealed relationships between DNA copy number and species abundance within a sample, indicating that the method can be used to estimate density at a site (Klobucar *et al.* 2017; Tillotson *et al.* 2018). This provides additional information beyond just presence-absence and is therefore an additional consideration when designing an eDNA study. While, in theory, multi-species eDNA approaches can also estimate a measure of DNA copy number within a sample, this relationship is much more complex in metabarcoding due to the methodology employed, and therefore generally considered a relatively weak indicator of species abundance at a site (Fonseca 2018).

The final consideration as to whether to adopt a single- or multi-species approach —cost-efficiency— has not, to the best of our knowledge, been explicitly studied. Metabarcoding is considerably more expensive than single species PCR-based approaches when only considering a single target species, but it remains to be seen at what point (e.g. the number of species) metabarcoding is more cost-efficient. Clearly this will depend on the number of species present in a system and the availability of single-species eDNA assays (Shaw *et al.* 2017).

For the remainder of our review, we focus primarily on considerations for single-species eDNA studies, as this approach has been developed to a much greater extent than metabarcoding. However, many of our recommendations pertain to general sampling design and thus are also relevant for eDNA metabarcoding.

When and where to sample?

As with other survey methods, the timing and location of eDNA sampling needs to be carefully considered. The optimal timing of eDNA sampling will depend on a species' phenology, as well as the likelihood of DNA transportation and retention. When the primary objective is to maximise the probability of detecting a single species, sampling at a time of year when the target species is most active (e.g., the breeding season), or is in highest abundance (Buxton *et al.* 2017), may maximise the amount of DNA present in the system. This assumes, of course, that this information is known for the target species, and that DNA shedding rates are equivalent across different life stages of the target species.

In lotic (flowing) systems, the transportation and retention of eDNA is also important. Environmental DNA can be transported downstream to an unoccupied site or diluted to undetectable levels at the point of origin (Pont *et al.* 2018). Estimates of eDNA transportation differ greatly between studies. Balasingham *et al.* (2017) found residual eDNA at detectable levels only 960 m downstream, whereas Deiner and Altermatt (2014), in a faster flowing but similar-sized stream, detected a species' eDNA 9.1 km downstream. The retention of eDNA in a system also needs to be considered, as eDNA can interact with environmental or biotic elements in the system that cause it to be retained, removed, or released (Shogren *et al.* 2017). Shogren *et al.* (2017) considered some of these complex issues by investigating the transport, retention and resuspension of eDNA empirically using controlled experimental streams. They

suggest that the complexity of eDNA detection in lotic systems should be considered but requires further progress to be made towards predictive modelling of eDNA transport.

Where to take a sample within a site is also important. The optimal sampling location within a site will depend on a species' habitat preferences and the likelihood of DNA retention (Buxton *et al.* 2018). Targeting preferred habitat could increase the likelihood of detecting the target species but certain habitats will retain DNA for longer periods of time. For example, the DNA of a species that occupies vegetation around the stream edge may be retained at a site for a longer period compared to a species that prefers the water column in the middle of the stream, simply because of the water's physical movement.

Sampling from the water column compared to the sediment could yield different concentrations of eDNA and represent different timescales of species presence. eDNA in sediments is more concentrated than eDNA from the water column, due to eDNA settling through the water column (Turner *et al.* 2015). The timescales represented by sediment samples could be much longer compared to those of the water column (Turner *et al.* 2015) due to the exponential decay of eDNA; a higher starting concentration results in a longer decay time (Thomsen *et al.* 2012; Barnes *et al.* 2014).

Environmental conditions can also affect eDNA detectability via DNA degradation. Strickler *et al.* (2015), for example, concluded that habitats with less solar radiation and more alkaline water retained eDNA for longer periods. However, there can also be a trade-off; environments with higher levels of solar radiation that are warmer can increase DNA shedding rates from a target species, offsetting higher degradation rates (Robson *et al.* 2016). Other biotic factors, such as the presence of biofilm, have been shown to increase the degradation rate of eDNA in flowing systems (Shogren *et al.* 2018).

How much water?

Filtration and precipitation methods have been used to capture eDNA and each method can process different volumes of water (Li *et al.* 2018). Precipitation involves adding sodium acetate and ethanol to water samples (as in Dejean *et al.* (2011) and Harper *et al.* (2018)). However, only a small subsample (typically 15 ml, but up to 30 ml (Eichmiller *et al.* 2016; Li *et al.* 2018)) can be used, which could impact the amount of eDNA that is recovered (Eichmiller *et al.* 2016). Filtration methods, in contrast, can process larger water volumes (Li *et al.* 2018).

Below we focus on filtration, as it is the most commonly used method (Smart *et al.* 2015; Shaw *et al.* 2016; Hinlo *et al.* 2017; Lugg *et al.* 2017; Tingley *et al.* 2018). Filtration can involve on-site filtration by hand (Lugg *et al.* 2017), using a peristaltic pump (Goldberg *et al.* 2011), or the collection of site water in sterile water bottles, which are later filtered in the laboratory using a vacuum pump (Piaggio *et al.* 2014; Smart *et al.* 2015).

The eDNA capture method selected should optimise eDNA preservation until DNA extraction can take place. In the case of filtering methods, there are various types of filters available, each made of different materials, and of open or closed varieties. Closed filters, which consist of a filter that is enclosed in a casing, have been used more recently in eDNA studies (Lugg *et al.* 2017; Spens *et al.* 2017). The advantage of using a closed filtering system is increased DNA preservation (reduced degradation) and potentially lower contamination risk (Spens *et al.* 2017). More traditional, open filter units require the filter to be handled, as well as the use of a filter funnel and vacuum pump, thus increasing contamination risk. The most efficient filtering membrane used in closed filters is polyethersulfone (PES), whereas open filters contain mixed cellulose ester membranes, polycarbonate track-etched filters, or glass fibre filters (Spens *et al.* 2017). Some membrane types can affect the ability to filter water regardless of pore size. For example, smaller volumes can be filtered with polyvinylidene difluoride (PVDF) Sterivex 0.45 µm filters than with PES Sterivex 0.22 µm filters, due to the hydrophobic nature and protein retention of the former. There are various preservation methods for filters (in the field or laboratory), including drying, freezing, ethanol, or the use of a buffer, such as Longmire's solution (Goldberg *et al.* 2016). The preservation method selected may depend on availability of field resources; for example, the availability of a freezer at field sites. PCR can now also be conducted in the field using portable qPCR machines, such as BioMeme® (Philadelphia, PA, USA), meaning the presence/absence of a target species can be confirmed within approximately one hour of sampling.

The pore size of the filter used to capture eDNA and the amount of water sampled can impact the amount and quality of DNA extracted from environmental samples. The source of DNA, for example extracellular or cellular DNA, could also dictate the most appropriate pore size (Taberlet *et al.* 2012). Some studies have found that smaller pore sizes can retain greater quantities and smaller particles of eDNA (Shaw *et al.* 2017). Turner *et al.* (2014) found that common carp (*Cyprinus carpio*) eDNA (an invasive species in Australia) ranged in size from > 180 µm to < 0.2 µm, and recommended using a 0.2 µm pore size to capture these smaller

particles. Eichmiller *et al.* (2016) and Liang and Keeley (2013) both found that smaller pore sizes (such as those around 0.2-0.6 µm) enabled the extraction of more eDNA. However, Li *et al.* (2018) found that pore size did not affect eDNA yield or species detectability, although they only tested filter pore sizes between 0.45 µm and 1.2 µm.

Given the emerging consensus in the literature that smaller pore sizes do not hinder, but in some cases, improve eDNA quantification, why might one consider a filter with a larger pore size? There is an inherent trade-off between filter pore size and the volume of water that can be passed through the filter (Mächler *et al.* 2016; Minamoto *et al.* 2016). Smaller pores get clogged more easily, and thus limit the amount of water that can be filtered for a given sample (Li *et al.* 2018). Filtration of a larger water volume has, for some species, shown to increase detection rate (Mächler *et al.* 2016). Presumably, filtering larger volumes of water increases the chance of capturing a species' DNA from a site; filtering small amounts of water risks missing the species' DNA entirely. However, to the best of our knowledge, this has not been rigorously tested across a range of species and therefore more research is required to improve our understanding of how water volume interacts with filter pore size for detectability of different species.

In practice, the choice of pore size will depend on the conditions of the study area. For example, in more turbid environments, such as farm dams or heavily disturbed sites, larger pore sizes may be required, as they will clog less often, enabling a larger volume of water to be processed. In clearer, faster-flowing streams, it may be preferential to use a smaller pore size to ensure smaller eDNA particles are captured. Collecting a larger quantity of water is also likely to become more important as water body size increases, although collecting additional water samples can help alleviate this issue.

How many samples?

The optimal number of samples to take at each site depends on the detectability of the target species. A species' detectability using eDNA depends on various ecological factors, including abundance (more individuals = more eDNA), life-history (fully aquatic or not), and habitat (fast-flowing river vs lentic pond). Detection probability also depends on the efficiency of the capture and DNA extraction methods, any sample interference (inhibition), and primer sensitivity (see Goldberg *et al.* 2016). The number of samples taken should reflect this detectability, as accurately estimating site occupancy of a species is generally the main

objective of any biodiversity survey. Detectability can be estimated with a pilot study at field sites where a species is simultaneously observed, or via mesocosm or laboratory trials. In the absence of resources for such studies, detectability can be estimated using taxonomically- or ecologically-similar species, although this approach has not, to the best of our knowledge, been validated for eDNA sampling.

The reason that replicate samples are recommended is that there is likely to be heterogeneity between samples (Schmidt *et al.* 2013; Furlan *et al.* 2016; Shogren *et al.* 2017), due to the nature of random sampling, and the potential for uneven distribution of eDNA molecules in the environment (Hunter *et al.* 2015). This heterogeneity is introduced by the water sampling process; for any given water sample, there is an associated 'availability' probability (Schmidt *et al.* 2013). That is, there is a probability that the water sample in question actually captures a species' DNA from the site when present. Note that this replication is distinct from replication at the qPCR level (technical replication), in which multiple qPCR assays are run on each water sample (see ***Accounting for imperfect detection: SODM for eDNA data***).

Greater numbers of water samples may be required at sites that present difficult filtration conditions (e.g. high turbidity or PCR inhibition). Under such conditions, increasing the number of samples collected could increase the probability of capturing the eDNA of the target species in the sampled water. Mächler *et al.* (2016) detected a positive relationship between the volume of water sampled and detection rate for one macroinvertebrate species but found no relationship for two others. They speculated that this relationship could exist for the other two species, but that the smallest volume of water tested did not reach the lower limit of detection for those species, and thus no relationship was observed. A review undertaken by Willoughby *et al.* (2016) found that water volume did not impact species detectability, but the number of replicate samples taken at a site did, as also suggested by Furlan *et al.* (2016), Schmidt *et al.* (2013) and Shogren *et al.* (2017). The effect of water volume on species detectability is likely to be species-specific.

A final consideration is whether to collect replicate samples on a single site visit (which minimises travel costs) or whether to stagger sample collection over multiple visits to reduce stochastic variability in eDNA detections. Although focused on a single species, the results of Smart *et al.* (2016) suggest that the latter approach - collecting water samples on repeat site visits - is the more cost-efficient approach.

Recognising potential sources of false positives

As with many survey methods (e.g., aural detections, point-count surveys), there is potential for one or more species to be detected at an unoccupied site (i.e., false positive detections) with eDNA sampling. False positives can enter the detection process via several pathways in the field, including sample contamination, eDNA transport, and eDNA persistence (Darling and Mahon 2011; Evans *et al.* 2017).

Contamination can be minimised by ensuring protocols are in place to avoid contaminating samples and sites with a target species' DNA. Such protocols should, at a minimum, involve the sterilisation of field equipment, such as boots, and the use of single-use gloves and sampling equipment (e.g. syringes). Decontamination of field equipment for re-use (e.g. boots, buckets, trays, sampling poles, bottles) can be undertaken using a 10% commercial bleach solution (Smart *et al.* 2015). Entering the water should be avoided completely, if possible, and may necessitate extra sampling equipment (e.g. sampling poles). False positive detections at the site level can also arise via the movement of eDNA by a non-human source; for example, a predator could move an animal's body from one site to another, or leave traces of another species via its faeces.

Natural transport of eDNA downstream is also a potential source of false positives in lotic systems. Methods to reduce the chance of false positive detections in lotic systems include ensuring adequate distance between sampling sites, sampling over a short time period, or ensuring that sampling sites are not connected by waterways (Lugg *et al.* 2017). However, the exact location of an individual may be irrelevant when conducting management over large spatial scales (e.g., at the scale of entire catchments). Similarly, sampling along a waterway in multiple locations can provide insight into this movement via estimation of eDNA concentrations of target species from samples (samples further away from the location of the target species will likely display lower eDNA concentrations).

False positive detections can also arise in the field via prolonged eDNA persistence after the extirpation of a species at a site. Detecting a species in conditions that rapidly degrade eDNA (e.g., high UV, low alkalinity) suggests that an organism was present not long before sampling. In conditions that favour eDNA persistence, eDNA could be from the time of sampling, or it could simply be persisting in the environment, making the timescale of inference more uncertain. This source of false positives is likely less of a concern than the potential sources

outlined above, given that many studies have found that eDNA degrades rapidly (days to weeks) in freshwater environments (Dejean *et al.* 2011; Piaggio *et al.* 2014). A species' DNA in the environment could be sourced from a live or dead organism, potentially resulting in a false positive result if the assumption is that the site is occupied by live organisms (Darling and Mahon 2011; Evans *et al.* 2017).

False positive detections can also occur in the laboratory as a result of sample contamination, or as a result of more technical aspects of eDNA analysis, such as primer specificity (Wilcox *et al.* 2013) or bioinformatic processing (Deiner *et al.* 2017). There are steps that can be undertaken to determine laboratory contamination via the use of negative controls (during DNA extraction and PCR) and/or mock communities for metabarcoding.

Accounting for imperfect detection: SODM for eDNA data

Given the above considerations, detection is likely to be imperfect with eDNA data. Accounting for imperfect detection, using any survey technique, is vital to ensure accurate estimates of site occupancy. Fortunately, a rich statistical literature on species occupancy detection models (SODM) has been developed to account for imperfect detection in wildlife survey data, and these models are a natural framework in which to analyse eDNA data (Dorazio and Erickson 2018; Strickland and Roberts 2019). SODM estimate site occupancy whilst accounting for imperfect detection, and thus are ideal for eDNA data, in which replicate samples are collected at a set of sites, and there is potential for both false positive and false negative detections (Schmidt *et al.* 2013; Lahoz-Monfort *et al.* 2016; Guillera-Arroita *et al.* 2017; Lugg *et al.* 2017).

Hierarchical SODM, which account for nested detection processes, are particularly attractive for the analysis of eDNA data. This is because multiple PCRs are often nested within multiple water samples taken at each site. For example, Hunter *et al.* (2015), Lugg *et al.* (2017) and Schmidt *et al.* (2013) use a three-level SODM considering (i) the latent presence/absence of a species, given the occupancy probability; (ii) the probability of the eDNA being contained (or 'available') within the water sample, given (i); and (iii) the probability of detecting the species' eDNA using PCR, given (ii). Using such hierarchical models enables imperfect detection to be incorporated at the different levels of the eDNA sampling process. Guillera-Arroita *et al.* (2017) have shown how these models can be extended to account for false positive detections.

SODM can also be applied to multi-species metabarcoding. For example, Valentini *et al.* (2016) used SODM to estimate detection probabilities for multiple amphibian species.

Conclusion

Environmental DNA sampling has been used in a wide variety of ecological applications to date, and important methodological considerations are beginning to emerge from this diverse literature. Here we have highlighted the need to carefully consider these methodological aspects before implementing an eDNA-based monitoring program. When applied appropriately and complemented by analytical methods that account for imperfect detection, eDNA sampling can be an effective survey method for documenting the distributions of native and non-native species in freshwater systems. Its high sensitivity and use as a species- or community-level survey tool means that it has the capacity to complement, or in some situations replace, more traditional sampling methods. However, as with any emerging technology, it is important that we evaluate the strengths and limitations of eDNA sampling relative to traditional methods, and carefully consider the scientific or management objective at hand.

Acknowledgements

This work was funded by Melbourne Water and the Australian Research Council via the Linkage Scheme (LP140100731) and San Diego Zoo Global. RT is supported by an Australian Research Council Discovery Early Career Researcher Award (DE170100601).

References

Andersen, K., Bird, K.L., Rasmussen, M., Haile, J., Breuning-madsen, H., Kjær, K.H., Orlando, L., Gilbert, M.T.P. and Willerslev, E. 2012. Meta-barcoding of ‘ dirt ’ DNA from soil reflects vertebrate biodiversity. *Molecular Ecology* 21: 1966–1979. http://doi.org/10.1111/j.1365-294X.2011.05261.x

Balasingham, K.D., Walter, R.P., and Heath, D.D. 2017. Residual eDNA detection sensitivity assessed by quantitative real-time PCR in a river ecosystem. *Molecular Ecology Resources* 17: 523–532. http://doi.org/10.1111/1755-0998.12598

Barnes, M.A., Turner, C.R., Jerde, C.L., Renshaw, M.A., Chadderton, W.L. and Lodge, D.M. 2014. Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science and Technology* 48(3): 1819–1827. http://doi.org/10.1021/es404734p

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Yu, D.W. and de Bruyn, M. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution* 29(6): 358–367. http://doi.org/10.1016/j.tree.2014.04.003

Buxton, A.S., Groombridge, J.J., Zakaria, N.B. and Griffiths, R.A. 2017. Seasonal variation in environmental DNA in relation to population size and environmental factors. *Scientific Reports* 7: 1–9. http://doi.org/10.1038/srep46294

Buxton, A.S., Groombridge, J.J. and Griffiths, R.A. 2018. Seasonal variation in environmental DNA detection in sediment and water samples. *PLoS ONE* 13(1): 1-14. https://doi.org/10.1371/journal.pone.0191737

Bylemans, J., Furlan, E.M., Hardy, C.M., McGuffie, P., Lintermans, M. and Gleeson, D.M. 2017. An environmental DNA-based method for monitoring spawning activity: a case study, using the endangered Macquarie perch (*Macquaria australasica*). *Methods in Ecology and Evolution* 8(5): 646–655. http://doi.org/10.1111/2041-210X.12709

Bylemans, J., Gleeson, D.M., Hardy, C.M. and Furlan, E. 2018. Toward an ecoregion scale evaluation of eDNA metabarcoding primers : A case study for the freshwater fish biodiversity of the Murray – Darling Basin (Australia). *Ecology and Evolution* (8): 8697–8712. http://doi.org/10.1002/ece3.4387

Clark, K., Karsch-mizrachi, I., Lipman, D.J., Ostell, J. and Sayers, E.W. 2016. GenBank. *Nucleic Acids Research* 44: 67–72. http://doi.org/10.1093/nar/gkv1276

Coissac, E., Riaz, T. and Puillandre, N. 2012. Bioinformatic challenges for DNA

metabarcoding of plants and animals. *Molecular Ecology* 21:1834-1847. https://doi.org/10.1111/j.1365-294X.2012.05550.x

Cristescu, M.E. and Hebert, P.D.N. 2018. Uses and Misuses of Environmental DNA in Biodiversity Science and Conservation. *Annual Review of Ecology, Evolution, and Systematics* (49): 209–230. http://doi.org/https://doi.org/10.1146/annurev-ecolsys-110617-062306

Cristescu, R.H., Foley, E., Markula, A., Jackson, G., Jones, D. and Frère, C. 2015. Accuracy and efficiency of detection dogs: a powerful new tool for koala conservation and management. *Scientific Reports* 5. http://doi.org/10.1038/srep08349

Dalén, L., Ötherström, A., Eijer, T. and Hapiro, B. 2007. Recovery of DNA from Footprints in the Snow. *Canadian Field-Naturalist* 121(3): 321–324. http://doi.org/http://dx.doi.org/10.22621/cfn.v121i3.482

Darling, John A. and Mahon, A.R. 2011. From molecules to management: Adopting DNA-based methods for monitoring biological invasions in aquatic environments. *Environmental Research* 111(7): 978-988. https://doi.org/10.1016/j.envres.2011.02.001

Deiner, K. and Altermatt, F. 2014. Transport distance of invertebrate environmental DNA in a natural river. *PLoS ONE* 9(2): e88786. http://doi.org/10.1371/journal.pone.0088786

Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., de Vere, N., Pfrender, M.E. and Bernatchez, L. 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology* 26: 5872–5895. http://doi.org/10.1111/mec.14350

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P. and Miaud, C. 2011. Persistence of environmental DNA in freshwater ecosystems. *PLoS ONE* 6(8): 8–11. http://doi.org/10.1371/journal.pone.0023398

Dejean, T., Valentini, A., Miquel, C., Taberlet, P., Bellemain, E. and Miaud, C. 2012.

Improved detection of an alien invasive species through environmental DNA barcoding: The example of the American bullfrog *Lithobates catesbeianus*. *Journal of Applied Ecology* 49(4): 953–959. http://doi.org/10.1111/j.1365-2664.2012.02171.x

Doi, H., Uchii, K., Takahara, T., Matsuhashi, S., Yamanaka, H. and Minamoto, T. 2015. Use of droplet digital PCR for estimation of fish abundance and biomass in environmental DNA surveys. *PLoS ONE* 10(3): 1–11. http://doi.org/10.1371/journal.pone.0122763

Dorazio, R.M. and Erickson, R.A. 2018. EDNAOCCUPANCY : An R package for multiscale occupancy modelling of environmental DNA data. *Molecular Ecology* 18: 368–380. http://doi.org/10.1111/1755-0998.12735

Eichmiller, J.J., Miller, L.M. and Sorensen, P.W. 2016. Optimizing techniques to capture and extract environmental DNA for detection and quantification of fish. *Molecular Ecology Resources* 16: 56–68. http://doi.org/10.1111/1755-0998.12421

Elbrecht, V. and Leese, F. 2017. Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment. *Frontiers in Environmental Science* 5: 1-11. https://doi.org/10.3389/fenvs.2017.00011

Evans, N.T., Shirey, P.D., Wieringa, J.G., Mahon, A.R. and Lamberti, G.A. 2017. Comparative Cost and Effort of Fish Distribution Detection via Environmental DNA Analysis and Electrofishing. *Fisheries* 42(2): 90–99. http://doi.org/10.1080/03632415.2017.1276329

Fonseca, V.G. 2018. “Pitfalls in relative abundance estimation using eDNA metabarcoding.” *Molecular Ecology Resources* 18: 923–926. http://doi.org/10.1111/1755-0998.12902

Foote, A.D., Thomsen, P.F., Sveegaard, S., Wahlberg, M., Kielgast, J., Kyhn, L.A., Salling, A.B., Galatius, A., Orlando, L. and Gilbert, M.T.P. 2012. Investigating the Potential Use of Environmental DNA (eDNA) for Genetic Monitoring of Marine Mammals. *PLoS ONE* 7(8): e41781. http://doi.org/10.1371/journal.pone.0041781

Freeland, J.R. 2017. The importance of molecular markers and primer design when characterizing biodiversity from environmental DNA. *Genome* 60: 358–374. http://doi.org/https://doi.org/10.1139/gen-2016-0100

Furlan, E.M., Gleeson, D., Hardy, C.M. and Duncan, R.P. 2016. A framework for estimating the sensitivity of eDNA surveys. *Molecular Ecology Resources* 16(3): 641-654. https://doi.org/10.1111/1755-0998.12483

Goldberg, C.S., Pilliod, D.S., Arkle, R.S. and Waits, L.P. 2011. Molecular detection of vertebrates in stream water: A demonstration using rocky mountain tailed frogs and Idaho giant salamanders. *PLoS ONE* 6(7). http://doi.org/10.1371/journal.pone.0022746

Goldberg, C.S., Turner, C.R., Deiner, K., Klymus, K.E., Francis, P., Murphy, M.A., Spear, S.F., Mckee, A., Oyler-MccCance, S.J., Cornman, R.S., Laramie, M.B., Mahon, A.R., Lance, R.F., Pilliod, D.S., Strickler, K.M., Waits, L.P., Fremier, A.K., Takahara, T., Herder, J.E. and Taberlet, P. 2016. Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution* 7: 1299–1307. http://doi.org/10.1111/2041-210X.12595

Goodenough, A.E., Carpenter, W.S., Mactavish, L., Mactavish, D., Theron, C. and Hart, A.G. 2017. Empirically testing the effectiveness of thermal imaging as a tool for identification of large mammals in the African bushveldt. *African Journal of Ecology* 56: 51–62. http://doi.org/10.1111/aje.12416

Guillera-Arroita, G., Lahoz-Monfort, J.J., van Rooyen, A.R., Weeks, A.R. and Tingley, R. 2017. Dealing with false-positive and false-negative errors about species occurrence at multiple levels. *Methods in Ecology and Evolution* 8(9): 1081–1091. http://doi.org/10.1111/2041-210X.12743

Harper, L.R., Handley, L.L., Hahn, C., Boonham, N., Rees, H.C., Gough, K.C., Lewis, E., Adams, I.P., Brotherton, P., Phillips, S. and Haenfling, B. 2018. Needle in a haystack? A

comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (*Triturus cristatus*). *Ecology and Evolution* 8(12): 6330-6341. https://doi.org/10.1002/ece3.4013

Hinlo, R., Furlan, E., Suitor, L. and Gleeson, D. 2017. Environmental DNA monitoring and management of invasive fish: comparison of eDNA and fyke netting. *Management of Biological Invasions* 8(1): 89–100. http://doi.org/10.3391/mbi.2017.8.1.09

Hodgson, J.C., Baylis, S.M., Mott, R., Herrod, A. and Clarke, R.H. 2016. Precision wildlife monitoring using unmanned aerial vehicles. *Scientific Reports* 6(1): 22574. http://doi.org/10.1038/srep22574

Hunter, M.E., Oyler-McCance, S.J., Dorazio, R.M., Fike, J.A., Smith, B.J., Hunter, C.T., Reed, R.N. and Hart, K.M. 2015. Environmental DNA (eDNA) sampling improves occurrence and detection estimates of invasive Burmese pythons. *PLoS ONE* 10(4): e0121655. http://doi.org/10.1371/journal.pone.0121655

Kelly, R.P., Port, J.A., Yamahara, K.M. and Crowder, L.B. 2014. Using environmental DNA to census marine fishes in a large mesocosm. *PLoS ONE* 9(1). http://doi.org/10.1371/journal.pone.0086175

Klobucar, S.L., Rodgers, T.W. and Budy, P. 2017. At the forefront: evidence of the applicability of using environmental DNA to quantify the abundance of fish populations in natural lentic waters with additional sampling considerations. *Canadian Journal of Fisheries and Aquatic Sciences* 74: 2030–2034. http://doi.org/10.1139/cjfas-2017-0114

Klymus, K.E., Marshall, N.T. and Stepien, C.A. 2017. Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *PLoS ONE* 12(5): e0177643. http://doi.org/https://doi.org/10.1371/journal.pone.0177643

Lahoz-Monfort, J.J., Guillera-Arroita, G. and Tingley, R. 2016. Statistical approaches to account for false-positive errors in environmental DNA samples. *Molecular Ecology*

Resources 16(3): 673–685. http://doi.org/10.1111/1755-0998.12486

Lahoz-Monfort, J.J. and Tingley, R. 2018. The technology revolution: improving species detection and monitoring using new tools and statistical methods. In S. Legge, N. Robinson, D. Lindenmayer, B. Scheele, D. Southwell, and B. Wintle (Eds.), Monitoring Threatened Species and Ecological Communities. Clayton: CSIRO Publishing.

Laramie, M.B., Pilliod, D.S. and Goldberg, C.S. 2015. Characterizing the distribution of an endangered salmonid using environmental DNA analysis. *Biological Conservation* 183: 29–37. http://doi.org/10.1016/j.biocon.2014.11.025

Leontidou, K., Verseni, C., de Groeve, J., Cristofolini, F., Vokou, D. and Cristofori, A. 2018. Taxonomic identification of airborne pollen from complex environmental samples by DNA metabarcoding: a methodological study for optimizing protocols. *Aerobiologia* 34(1): 63–74. http://doi.org/10.1101/099481

Li, J., Handley, L.L., Read, D.S. and Hänfling, B. 2018. The effect of filtration method on the efficiency of environmental DNA capture and quantification via metabarcoding. *Molecular Ecology Resources* 00: 1–13. http://doi.org/10.1111/1755-0998.12899

Liang, Z. and Keeley, A. 2013. Filtration recovery of extracellular DNA from environmental water samples. *Environmental Science and Technology* 47(16): 9324–9331. http://doi.org/10.1021/es401342b

Lugg, W.H., Griffiths, J., van Rooyen, A.R., Weeks, A.R. and Tingley, R. 2017. Optimal survey designs for environmental DNA sampling. *Methods in Ecology and Evolution* 00: 1–11. http://doi.org/10.1111/2041-210X.12951

Mächler, E., Deiner, K., Spahn, F. and Altermatt, F. 2016. Fishing in the Water: Effect of Sampled Water Volume on Environmental DNA-Based Detection of Macroinvertebrates. *Environmental Science and Technology* 50(1): 305–312. http://doi.org/10.1021/acs.est.5b04188

Minamoto, T., Naka, T., Moji, K. and Maruyama, A. 2016. Techniques for the practical collection of environmental DNA: filter selection, preservation, and extraction. *Limnology* 17(1): 23–32. http://doi.org/10.1007/s10201-015-0457-4

Piaggio, A.J., Engeman, R.M., Hopken, M.W., Humphrey, J.S., Keacher, K.L., Bruce, W.E. and Avery, M.L. 2014. Detecting an elusive invasive species : a diagnostic PCR to detect Burmese python in Florida waters and an assessment of persistence of environmental DNA. *Molecular Ecology Resources* 14: 374–380. http://doi.org/10.1111/1755-0998.12180

Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., Roset, N., Schabuss, M., Zornig, H. and Dejean, T. 2018. Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Scientific Reports* 8(1): 10361. http://doi.org/10.1038/s41598-018-28424-8

Rees, H.C., Maddison, B.C., Middleditch, D.J., Patmore, J.R.M. and Gough, K.C. 2014. The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology* 51(5): 1450–1459. http://doi.org/10.1111/1365-2664.12306

Robson, H.L.A., Noble, T.H., Saunders, R.J., Robson, S.K.A., Burrows, D.W. and Jerry, D.R. 2016. Fine-tuning for the tropics: application of eDNA technology for invasive fish detection in tropical freshwater ecosystems. *Molecular Ecology Resources* 16(4): 922–932. http://doi.org/10.1111/1755-0998.12505

Schmelzle, M.C. and Kinziger, A.P. 2016. Using occupancy modelling to compare environmental DNA to traditional field methods for regional-scale monitoring of an endangered aquatic species. *Molecular Ecology Resources* 16(4): 895–908. http://doi.org/10.1111/1755-0998.12501

Schmidt, B.R., Kéry, M., Ursenbacher, S., Hyman, O.J. and Collins, J.P. 2013. Site

occupancy models in the analysis of environmental DNA presence/absence surveys: A case study of an emerging amphibian pathogen. *Methods in Ecology and Evolution* 4: 646–653. http://doi.org/10.1111/2041-210X.12052

Serena, M. 1994. Use of time and space by platypus (*Ornithorhynchus anatinus*: Monotremata) along a Victorian stream. *Journal of Zoology* 232: 117–131. http://doi.org/10.1111/j.1469-7998.1994.tb01563.x

Shaw, J.L.A., Clarke, L.J., Wedderburn, S.D., Barnes, T.C., Weyrich, L.S. and Cooper, A. 2016. Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biological Conservation* 197: 131–138. http://doi.org/10.1016/j.biocon.2016.03.010

Shaw, J.L.A., Weyrich, L. and Cooper, A. 2017. Using environmental (e)DNA sequencing for aquatic biodiversity surveys: a beginner's guide. *Marine and Freshwater Research* 68: 20–33. http://doi.org/http://dx.doi.org/10.1071/MF15361

Shogren, A.J., Tank, J.L., Andruszkiewicz, E., Olds, B., Mahon, A.R., Jerde, C.L. and Bolster, D. 2017. Controls on eDNA movement in streams : Transport , Retention , and Resuspension. *Scientific Reports* 7: 5065. https://doi.org/10.1038/s41598-017-05223-1

Shogren, A., Tank, J.L., Egan, S.P., August, O., Rosi, E.J., Hanrahan, B.R., Renshaw, M.A., Gantz, C.A. and Bolster, D. 2018. Water flow and biofilm cover influence environmental DNA (eDNA) detection in recirculating streams. *Environmental Science & Technology* 52(15): 8530–8537. http://doi.org/10.1021/acs.est.8b01822

Sigsgaard, E.E., Nielsen, I.B., Bach, S.S., Lorenzen, E.D., Robinson, D.P., Knudsen, S.W., Pedersen, M.W., Jaidah, M. Al, Orlando, L., Willerslev, E., Møller, P.R. and Thomsen, P.F. 2016. Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. *Nature Ecology & Evolution* 1: 0004.

http://doi.org/10.1038/s41559-016-0004

Simpfendorfer, C., Kyne, P., Noble, T., Goldsbury, J., Basiita, R., Lindsay, R., Shields, A., Perry, C. and Jerry, D. 2016. Environmental DNA detects Critically Endangered largetooth sawfish in the wild. *Endangered Species Research* 30: 109–116. http://doi.org/10.3354/esr00731

Smart, A.S., Tingley, R., Weeks, A.R., van Rooyen, A.R. and McCarthy, M.A. 2015. Environmental DNA sampling is more sensitive than a traditional survey technique for detecting an aquatic invader. *Ecological Applications* 25(7): 1944–1952. http://doi.org/10.1890/14-1751.1

Smart, A.S., Weeks, A.R., van Rooyen, A.R., Moore, A., McCarthy, M.A. and Tingley, R. 2016. Assessing the cost-efficiency of environmental DNA sampling. *Methods in Ecology and Evolution* 7: 1291–1298. http://doi.org/10.1111/2041-210X.12598

Spens, J., Evans, A.R., Halfmaerten, D., Knudsen, S.W., Sengupta, M.E., Mak, S.S.T., Sigsgaard, E.E. and Hellström, M. 2017. Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol : advantage of enclosed filter. *Methods in Ecology and Evolution* 8: 635–645. http://doi.org/10.1111/2041-210X.12683

Strickland, G.J. and Roberts, J.H. 2019. Utility of eDNA and occupancy models for monitoring an endangered fish across diverse riverine habitats. *Hydrobiologia* 826(1): 129–144. http://doi.org/10.1007/s10750-018-3723-8

Strickler, K.M., Fremier, A.K. and Goldberg, C.S. 2015. Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biological Conservation* 183: 85–92. http://doi.org/10.1016/j.biocon.2014.11.038

Taberlet, P., Coissac, E., Hajibabaei, M. and Rieseberg, L.H. 2012. Environmental DNA. *Molecular Ecology* 21: 1789–1793. http://doi.org/10.1111/j.1365-294X.2012.05542.x

Thomsen, P.F., Kielgast, J., Iversen, L.L., Møller, P.R., Rasmussen, M. and Willerslev, E. 2012. Detection of a Diverse Marine Fish Fauna Using Environmental DNA from Seawater Samples. *PLoS ONE* 7(8): 1–9. http://doi.org/10.1371/journal.pone.0041732

Tillotson, M.D., Kelly, R.P., Duda, J.J., Hoy, M., Kralj, J. and Quinn, T.P. 2018. Concentrations of environmental DNA (eDNA) reflect spawning salmon abundance at fine spatial and temporal scales. *Biological Conservation* 220: 1–11. http://doi.org/10.1016/j.biocon.2018.01.030

Tingley, R., Greenless, M., Oertel, S., van Rooyen, A.R. and Weeks, A.R. 2018. Environmental DNA sampling as a surveillance tool for cane toad *Rhinella marina* introductions on offshore islands. *Biological Invasions*. http://doi.org/10.1007/s10530-018-1810-4

Turner, C.R., Miller, D.J., Coyne, K.J. and Corush, J. 2014. Improved Methods for Capture, Extraction, and Quantitative Assay of Environmental DNA from Asian Bigheaded Carp (*Hypophthalmichthys* spp.). *PLoS ONE* 9(12): 1-20. https://doi.org/10.1371/journal.pone.0114329

Turner, C.R., Uy, K.L. and Everhart, R.C. 2015. Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation* 183: 93–102. http://doi.org/10.1016/j.biocon.2014.11.017

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P.F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G.H., Geniez, P., Pont, D., Argillier, C., Baudoin, J.M., Peroux, T,. Crivelli, A.J., Olivier, A., Acqueberge, M., Le Brun, M., Møller, P.R., Willerslev, E. and Dejean, T. 2016. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology* 25(4): 929–942. http://doi.org/10.1111/mec.13428

Wilcox, T.M., McKelvey, K.S., Young, M.K., Jane, S.F., Lowe, W.H., Whiteley, A.R. and

Schwartz, M.K. 2013. Robust Detection of Rare Species Using Environmental DNA: The Importance of Primer Specificity. *PLoS ONE* 8(3). http://doi.org/10.1371/journal.pone.0059520

Wilcox, T.M., McKelvey, K.S., Young, M.K., Sepulveda, A.J., Shepard, B.B., Jane, S.F., Whiteley, A.R., Lowe, W.H. and Schwartz, M.K. 2016. Understanding environmental DNA detection probabilities: A case study using a stream-dwelling char *Salvelinus fontinalis*. *Biological Conservation* 194: 209–216. http://doi.org/10.1016/j.biocon.2015.12.023

Willoughby, J.R., Wijayawardena, B.K., Sundaram, M., Swihart,R.K. and DeWoody, J.A. 2016. The importance of including imperfect detection models in eDNA experimental design. *Molecular Ecology Resources* 16: 837-844. https://doi.org/10.1111/1755-0998.12531

Appendix B
Supplementary material for chapter 2

eDNA analysis

DNA extraction

DNA was extracted from filters using Qiagen DNeasy Blood & Tissue Kits (Spin-column protocol) in a room that is dedicated to low-quantity DNA sources with qPCR setup undertaken in a laminar flow hood. Into each filter unit, 540 µL of ATL buffer and 40 µL of proteinase K was added and then each filter sealed and incubated for 3 hours at 56 °C with constant agitation. The lysis solution was transferred into new 2 mL tubes. The manufacturers protocol was followed for the rest of the DNA extraction with some minor adjustments; 500 µL AL buffer, 500 µL ethanol, and a final elution step of 100 µL AE buffer was used for each sample. Included in every batch of DNA extractions (~24 samples) was a negative DNA extraction control (that replicated the process using a sterile Sterivex filter). For the SE Australia dataset, there were 16 negative extraction controls and for the Melbourne dataset there were 15 negative extraction controls.

Single species detection (qPCR assay)

Species-specific primers and TaqMan® probe targeting the platypus *Cytb* mitochondrial control region used in this study were from Lugg *et al*. (2017). Sequences for the primers/probe were; forward primer OAcr_F CAGCAATACCCTAGACAAGG, reverse primer OAcr_R CGCTTCAATGGCTGCGC, and MGB probe OAcr_MGB CGAACCCCATGAGTAGAAAAT. Assays were from Life Technologies (Thermo Fisher Scientific) as a custom TaqMan® gene expression assay.

Using a Roche LightCycler 480 system in a 384-well format, Real-time TaqMan® PCR assays were conducted in 10 µL reactions containing 5 µL of 2 × Qiagen multiplex PCR Master Mix (Qiagen), 0.5 µL 20× TaqMan® Gene Expression Assay (final concentration each primer/probe; 900/250 nM), 2.5 µL ddH_2O and 2 µL of DNA were prepared in triplicate. In each 384-well assay plate, control reactions containing 10, 100, 1000, 10000, 100,000 copies

of *O. anatinus Cytb* synthetic oligonucleotide (Integrated DNA technologies, Baulkham Hills, NSW, Australia), a negative qPCR control with no DNA template and all DNA extraction negative controls were included. Amplification occurred in the following conditions: 15 min at 95 °C,15 s at 95 °C and 1 min at 60 °C for 50 cycles. Crossing threshold (Ct) values were determined from the amplification profiles of PCR using the Absolute Quantification module of the LightCycler® 480 software package. To test for the presence of PCR inhibitors, a TaqMan® Exogenous Internal Positive Control VIC probe was also run for each sample. All negative extraction controls extracted at the same time as samples for each dataset were also included in assays when analysing samples for each dataset. The efficiency of all qPCR reactions was >98% and all negative controls (extraction or qPCR) were negative.

Multi-species detection method (metabarcoding)

A universal primer targeting all vertebrates (Riaz *et al.* 2011; Shehzad *et al.* 2012) was used for metabarcoding in this study. We chose a vertebrate primer (as opposed to a mammal-specific primer) due to the low number of aquatic mammals in the study area, and because the vertebrate primer has the benefit of detecting other taxonomic groups, such as fish and amphibians. The assay marker region was ~100 bp and located within the mitochondrial 12S ribosomal RNA (rRNA) gene. For the *SE Australia* dataset, we analysed each of the two water sample replicates separately. For the *Melbourne* dataset, the water sample replicates were pooled together for analysis. A two-step PCR protocol was used to prepare metabarcoding libraries, with primer sequences listed in Table S1. The first round PCR primers contained a (5′–3′) universal adaptor sequence, a 0- to 6-bp heterogeneity spacer (Fadrosh *et al*. 2014) and the marker-specific primer sequence. Reactions contained 2 µl eDNA, 1× PCR buffer, 3 mM $MgCl_2$, 0.3 µM each primer and 0.2 Units KAPA Plant DNA Polymerase (Kapa Biosystems). Reaction conditions were one cycle at 95 °C for 3 min, 40 cycles at 95°C for 20 s, 57°C (vertebrate) for 15 s, 72 °C for 15 s; and one cycle at 72 °C for 1 min. Four PCR technical replicates were performed for each sample, one for each heterogeneity spacer combination (see Table S1) with samples randomly allocated across PCR plates. DNA extraction negative controls for each sample dataset (17 and 5 for the *SE Australia* and *Melbourne* datasets respectively) and PCR negatives (96 and 12 respectively) were included and carried through to sequencing. We also included three eDNA samples from the Northern Territory (Australia) from a previous project in the *SE Australia* dataset, where platypuses do not occur, and the fish

community had already been determined using the same vertebrate markers to act as a known "mock community".

Following the first-round PCR, the four technical replicates were pooled and purified with either ExoSAP-IT PCR Product Cleanup Reagent (ThermoFisher) (*SE Australia* dataset) or Sera-Mag SpeedBead Carboxylate-Modified Magnetic Particles (GE Healthcare Life Sciences) (*Melbourne* dataset). Second-round PCR primers contained (5′–3′) the Illumina p5 or p7 binding region, an 8-bp index sequence (Fadrosh *et al.* 2014) and a universal adaptor sequence. Unique forward and reverse index combinations were used for each sample. Reactions contained 2 µl purified first PCR product, 0.5 µM of each primer and 1× Phusion Hot Start II High-Fidelity PCR Master Mix (Thermo Fisher Scientific). Reaction conditions were one cycle at 98 °C for 1 min; 15 cycles at 98 °C for 10 s, 72°C for 45 s; and one cycle at 72°C for 10 min. For the *Melbourne* dataset, the SequalPrep Normalization Plate (96) Kit (Invitrogen) was used to normalize the concentration of PCR products for each sample. Normalized samples were pooled and then subject to dual size selection to remove nonspecific fragments outside of the desired size range, using Sera-Mag SpeedBead Carboxylate-Modified Magnetic Particles (GE Healthcare Life Sciences). Sequencing was performed at the Monash Health Translation Precinct (MHTP) Medical Genomics Facility on an Illumina MiSeq platform (Illumina) using 150 bp PE chemistry. For the *SE Australia* dataset, samples were pooled at equal volumes then subject to size selection using Sera-Mag SpeedBead Carboxylate-Modified Magnetic Particles, with sequencing performed on an Illumina NextSeq platform at the MHTP Medical Genomics Facility.

Bioinformatics

Bioinformatic analyses were performed as in McColl-Gausden *et al.* (2021) using a custom analysis pipeline. A custom reference database for taxonomy assignment was built for vertebrate species known to occur in Victoria, Australia. The R package RENTREZ v1.1.0 (Winter 2017) was used to retrieve all 12S rDNA sequences present in GenBank (https://www.ncbi.nlm.nih.gov/genbank/). An *in-silico* PCR, performed with the primer sequences listed in Table S1 and the software packages OBITOOLS (Boyer *et al.* 2016) and ecoPCR (Ficetola *et al.* 2010) was then undertaken to retrieve the appropriate gene region for reference species.

Following merging of R1 and R2 paired-end sequences (using VSEARCH's fastq-mergepairs function), primer sequences were trimmed and sequencing reads dereplicated where those with an abundance of less than 10 were excluded from further analyses. Sequences were clustered into Operational Taxonomic Units (OTUs) using a pairwise identity of 100% (vsearch --id 1.0). Taxonomic assignment was performed with VSEARCH (Rognes *et al.* 2016) whereby each OTU was assigned a taxonomic identity using a minimum bootstrap support of 95% (vsearch –sintax-cutoff 0.95) by comparing against the custom vertebrate reference sequence database. As we were only interested in assignments to platypuses, which are unique within our reference database, we did not seek to resolve any other taxonomic assignments.

Simulation analysis

To investigate whether the four case studies used here have the study design structure (e.g., site, sample, and qPCR replication) necessary to accurately estimate the parameter of interest (i.e., the effect of eDNA detection method β), we conducted model simulations. For each of the dataset structures represented by the four case studies, a constant parameter set generally representative of eDNA parameter estimates in the literature (ψ, θ, p = 0.8: (Dejean *et al.* 2012; Lugg *et al.* 2017; McColl-Gausden *et al.* 2021; Tingley *et al.* 2018; Tingley *et al.* 2021) was used to simulate 500 datasets. This was done twice for each study design structure— once with β set at -2 and once with β set at 2 —to determine whether each data structure could retrieve these parameter values. Each of the 500 simulated datasets was used to fit the model described above (equations. 1-4), and the mean estimate of β was retained, resulting in 1,000 estimates of β for each study design structure (500 for β = 2; 500 for β = -2).

Results

All study design structures were able to recover the β parameter in some iterations (Figure B.1(a-d) and B.2(a-d)), although the number of site replicates was important. Study design structures were generally able to recover β = −2 more effectively than β = 2. For the *SE Australia* dataset, estimates of β were −2.019 [−2.519, −1.530] when β = −2, and 2.508 [1.190, 4.371]) when β = 2. For the *Melbourne* dataset, when β = −2 and β = 2, respectively, estimates were −2.008 [−2.234, −1.785] and 2.053 [1.560, 2.600]. For the *Harper* dataset, when β = −2 and β = 2 respectively, estimates were −2.003 [−2.238, −1.770] and 2.082 [1.499, 2.756]). For the *Bylemans* dataset, β estimates were −2.069 [−2.831, −1.326] when β = −2, and 4.087 [1.031,

9.227] when β = 2. Plots of simulated estimates of occupancy, availability, and detection probabilities (as opposed to the β parameter) can be found in Figures B.3, B.4 and B.5.

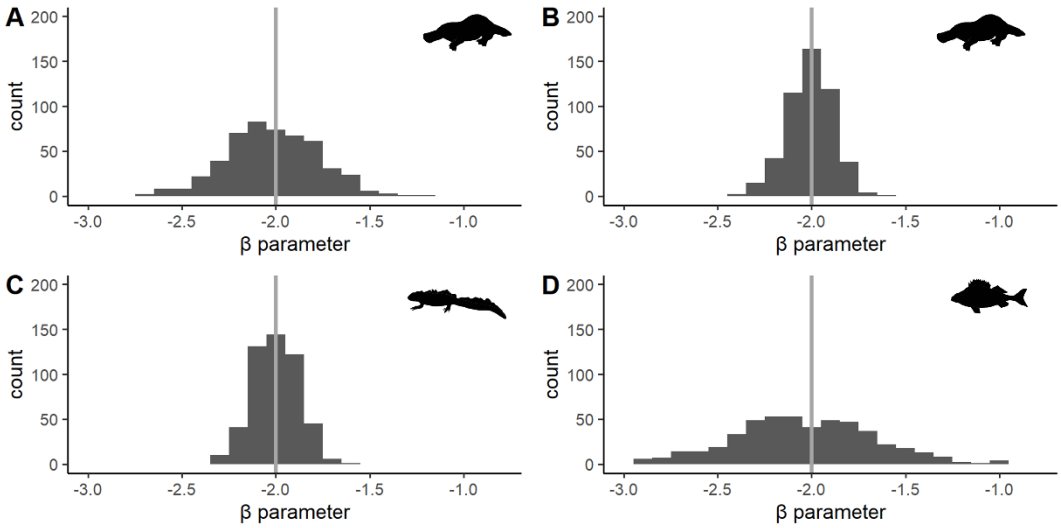

Figure B.1: Simulated estimates of β parameter where occupancy probability = 0.8, availability probability = 0.8, detection probability = 0.8 and β = −2 (vertical line). A = SE Australia dataset, B = Melbourne dataset, C = Harper et al., (2018) dataset, D = Bylemans et al., (2019) dataset.

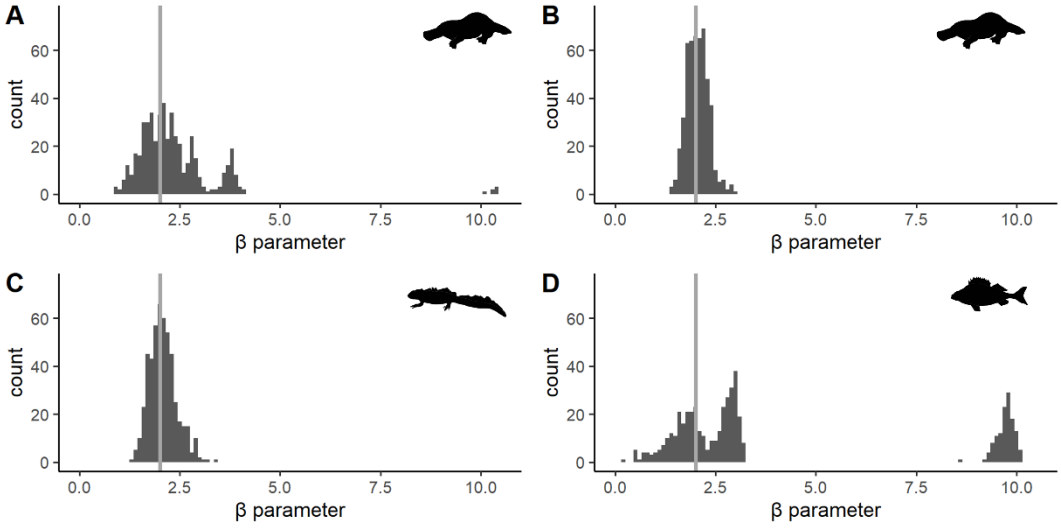

Figure B.2: Simulated estimates of β parameter where occupancy probability = 0.8, availability probability = 0.8, detection probability = 0.8 and β = 2 (vertical line). A = SE Australia dataset, B = Melbourne dataset, C = Harper et al., (2018) dataset, D = Bylemans et al., (2019) dataset.

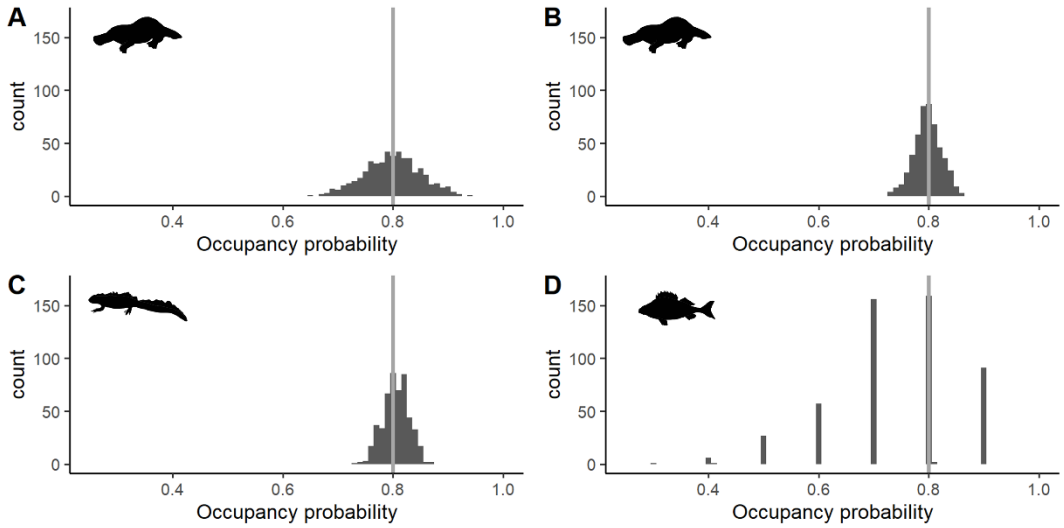

Figure B.3: Simulated estimates of occupancy probability where occupancy probability = 0.8, availability probability = 0.8, detection probability = 0.8 and β = -2. A = SE Australia dataset, B = Melbourne dataset, C = (Harper et al., 2018) dataset, D = (Bylemans et al., 2019) data.

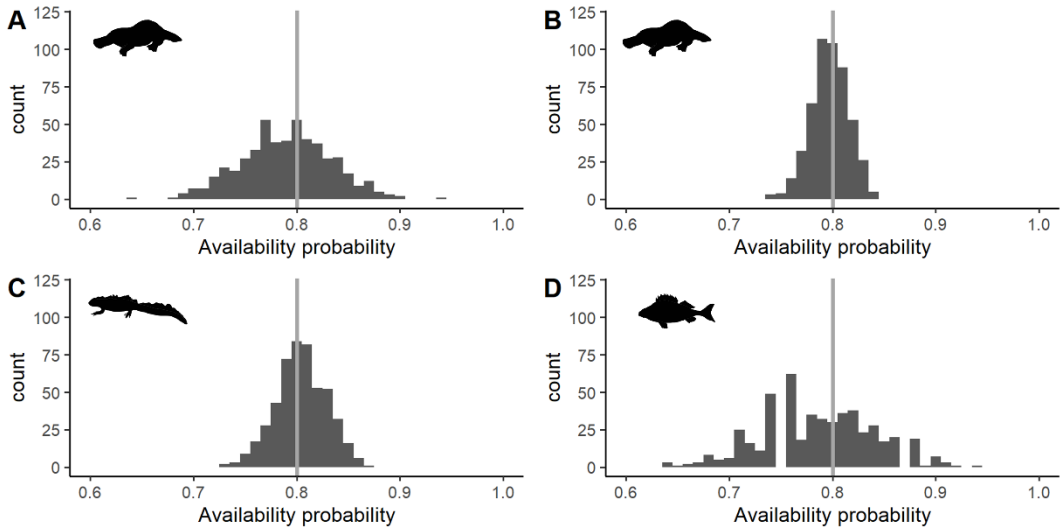

Figure B.4: Simulated estimates of availability probability where occupancy probability = 0.8, availability probability = 0.8, detection probability = 0.8 and β = -2. A = SE Australia dataset, B = Melbourne dataset, C = (Harper et al., 2018) dataset, D = (Bylemans et al., 2019) data.

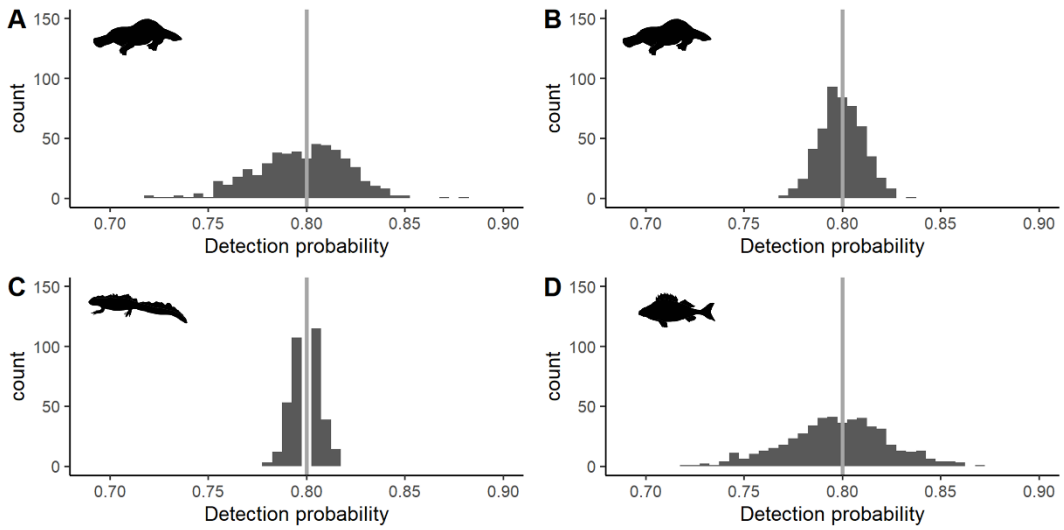

Figure B.5: Simulated estimates of detection probability where occupancy probability = 0.8, availability probability = 0.8, detection probability = 0.8 and β = -2. A = SE Australia dataset, B = Melbourne dataset, C = (Harper et al., 2018) dataset, D = (Bylemans et al., 2019) data.

Example JAGS code

```
cat("
   model {
   # Priors and model for params
   int.psi ~ dunif(0,1)         # Intercept of occupancy probability
   int.theta ~ dunif(0,1) # Intercepts availability probability
   int.p ~ dunif(0,1)     # Intercepts detection probability (1-PCR error)
   beta.lp ~ dnorm(0, 0.1)

   # 'Likelihood' (or basic model structure)
   for (i in 1:nsite){
   # Occurrence at site i
   z[i] ~ dbern(psi[i])
   logit(psi[i]) <- logit(int.psi)

for (j in 1:nrep){
   # Occurrence in sample j
```

```
    a[i,j] ~ dbern(mu.a[i,j])
    mu.a[i,j] <- z[i] * theta[i,j]
    logit(theta[i,j]) <- logit(int.theta)

for (k in 1:npcr){
    # PCR detection error process in sample k
    platy_qpcr_meta[i,j,k] ~ dbern(mu.y[i,j,k])
    mu.y[i,j,k] <- a[i,j] * p[i,j,k]
    logit(p[i,j,k]) <- logit(int.p)+ beta.lp* method[i,j,k]
    }}

    tmp[i] <- step(sum(a[i,])-0.1)
    }
    # Derived quantities
    sum.z <- sum(z[])   # Total # of occupied sites in sample
    sum.a <- sum(tmp[]) # Total # of sites with presence in <=1 of the 5 samples
    }  ",fill = TRUE)
sink()

zst <- apply(platy_qpcr_meta, 1, max)        # inits for presence (z)
zst[is.na(zst)] <- 1
ast <- apply(platy_qpcr_meta, c(1,2), max)   # inits for availability (a)
ast[is.na(ast)] <- 1
inits <- function() list(z = zst, a = ast)

# Parameters monitored
params1 <- c("int.psi", "int.theta", "int.p","beta.lp", "sum.z", "sum.a",
             "p","psi","theta")

# MCMC settings
ni <- 30000    ;   nt <- 10   ;   nb <- 10000   ;   nc <- 3
```

out_platy_compT1<- jags(platy_compT1, inits, params1, "platy_compT1.txt", n.chains = nc, n.thin = nt, n.iter = ni,n.burnin = nb)

Table B.1: Probability of detection, availability, and occupancy for each method in each scenario. NT = No Threshold, T = Threshold, Values = posterior means

Scenario (qPCR-metabarcoding)	Occupancy probability (ψ)	Availability probability (θ)	Detection probability (p) (qPCR)	Detection probability (p) (metabarcoding)
SE AUS NT-NT	0.838 (0.754, 0.910)	0.932 (0.880, 0.969)	0.833 (0.794, 0.868)	0.632 (0.717, 0.542)
SE AUS NT-T	0.837 (0.751, 0.911)	0.933 (o.882, 0.971)	0.831 (0.791, 0.867)	0.360 (0.278, 0.453)
SE AUS NT-NT pooled	0.834 (0.758, 0.906)	0.951 (0.904, 0.984)	0.815 (0.774, 0.853)	0.685 (0.572, 0.789)
SE AUS 1/3-NT	0.834 (0.748, 0.905)	NA	0.933 (0.885, 0.968)	0.591 (0.504, 0.674)
SE AUS 2/3-NT	0.778 (0.684, 0.861)	NA	0.849 (0.781, 0.907)	0.642 (0.551, 0.728)
Melbourne NT-NT	0.505 (0.442, 0.568)	0.746 (0.687, 0.804)	0.537 (0.497, 0.577)	0.463 (0.369, 0.562)
Melbourne 1/3-NT	0.500 (0.440, 0.560)	NA	0.676 (0.626, 0.724)	0.360 (0.280, 0.442)
Melbourne 2/3-NT	0.336 (0.278, 0.397)	NA	0.522 (0.456, 0.588)	0.537 (0.430, 0.645)
Bylemans NT-T Autumn	0.802 (0.512, 0.970)	0.671 (0.546, 0.788)	0.814 (0.738, 0.878)	0.738 (0.588, 0.858)
Bylemans 1/3-T Autumn	0.800 (0.527, 0.970)	NA	0.672 (0.785, 0.544)	0.510 (0.376, 0.638)
Bylemans 2/3-T Autumn	0.598 (0.300, 0.864)	NA	0.784 (0.652, 0.892)	0.701 (0.556, 0.829)
Bylemans NT-T Spring	0.700 (0.401, 0.925)	0.729 (0.596, 0.845)	0.719 (0.628, 0.802)	0.492 (0.330, 0.653)
Bylemans 1/3-T Spring	0.700 (0.398, 0.926)	NA	0.698 (0.568, 0.817)	0.364 (0.234, 0.499)
Bylemans 2/3-T Spring	0.599 (0.304, 0.861)	NA	0.594 (0.447, 0.735)	0.437 (0.287, 0.595)
Harper NT-NT	0.737 (0.533, 0.982)	0.751 (0.538, 0.988)	0.526 (0.510, 0.543)	0.636 (0.579, 0.691)
Harper NT-T	0.729 (0.522, 0.982)	0.733 (0.521, 0.982)	0.545 (0.528, 0.562)	0.539 (0.528, 0.598)
Harper 1/12-T	0.540 (0.495, 0.588)	NA	0.919 (0.871, 0.957)	0.520 (0.458, 0.579)

Harper 4/12-T	0.398 (0.354, 0.443)	NA	0.820 (0.755, 0.877)	0.705 (0.635, 0.769)
Harper 1/12-NT	0.564 (0.519, 0.610)	NA	0.879 (0.828, 0.921)	0.607 (0.547, 0.665)
Harper 4/12-NT	0.445 (0.400, 0.491)	NA	0.733 (0.666, 0.794)	0.769 (0.706, 0.827)

Two-level model

Equation 1. $$z_i \sim Bernoulli(\psi_i)$$

Equation 2. $$y_{ij}|z_i \sim Bernoulli(z_i p_{ij})$$

Equation. 3. $$logit(p_{ij}) = \alpha + \beta M_{ij}$$

Equation 1 describes the occupancy of the species at site i ($z_i = 1$ where the species is present, and $z_i = 0$ where the species is absent), given the mean probability of occurrence for that species across all sites (ψ_i). Equation 2 describes the observed detection process at site i for water sample j, conditional on the occurrence (z_i) of the species at site i. Finally, Equation 3 shows that the probability of detecting the species' DNA (p_{ij}) is a function of an overall intercept α and a coefficient describing the effect of detection method βM_{ij}, where $M_{ij} = 0$ for qPCR, and $M_{ij} = 1$ for metabarcoding. A positive parameter estimate (β) therefore indicates that the species was more likely to be detected with metabarcoding. Likewise, a negative parameter estimate (β) suggests that the species was more likely to be detected with qPCR. Prior distributions were specified on the logit scale where ψ, p and β were normally distributed with a mean of 0 and a precision of 0.1.

Appendix C
Supplementary material for chapter 3

Interim Biogeographic Regionalisation for Australia (IBRA) bioregions

Naracoorte Coastal Plain (NCP), Murray Darling Depression (MDD), Victorian Midlands (VIM), Southern Volcanic Plain (SVP), South East Coastal Plain (SCP), South Eastern Highland (SEH), Riverina (RIV), NSW South Western Slopes (NSS), Australian Alps (AUA), South East Corner (SEC), Sydney Basin (SYB) and Furneaux (FUR).

Number of sites per contracted catchment:

Number of segments in contracted catchments <=50, 3 sites sampled

Number of segments in contracted catchments >50, 4 sites sampled

Number of segments in contracted catchments >100, 5 sites sampled

Number of segments in contracted catchments >200, 6 sites sampled

Number of segments in contracted catchments >300, 7 sites sampled

Number of segments in contracted catchments >, 8 sites sampled

These values were used as a guide for sampling, in some contracted catchments it was not possible to sample the calculated number of sites. Additional samples in some contracted catchments were taken (up to 12) where additional analyses were taking place.

Table C.1: Environmental variables collected at each site

Covariate	Very poor	Poor	Fair	Good	Excellent	Potential benefit to platypus	References
Bank erosion	Extensive erosion. Little vegetation present to stabilise banks	Evidence of recent erosion. Extensive areas of bare banks	Erosion occurring in specific areas. Moderate vegetation cover	Erosion only in very small spots. Good vegetation cover. Usually gentle bank slopes.	No erosion evident. Lower banks covered with grass, reeds, or shrubs.	Minimising erosion and subsequent sedimentation supports burrow construction for platypus.	(Woon 1995; Grant and Temple-Smith 1998; Rohweder and Baverstock 1999)
Bank vegetation (immediately adjacent to water <2m)	Mostly bare ground. Occasional tree. Concrete or paved channel.	Introduced ground cover. Little native woody vegetation.	Moderate cover of native/introduced vegetation. Variation between sides - one cleared, one vegetated.	Mainly native vegetation. Little disturbance to banks.	Mainly undisturbed native vegetation (i.e., forested)	Bank vegetation minimises erosion & sedimentation which supports burrow construction, habitat & food for invertebrate prey, and shelter for platypus while foraging. Bank vegetation also conceals burrow entrances	(Woon 1995; Serena *et al.* 1998; Ellem *et al.* 1998; Grant and Temple-Smith 1998)
Verge vegetation	Bare ground or pasture/grass	Narrow band (<5m) of sparse native or	Wide corridor (5-10m) of native/introduced	Mainly native but some introduced	Mainly native vegetation on both sides (>30m wide)	Verge vegetation consolidates banks for burrow construction. It also provides, shade to prevent	(Woon 1995; Serena *et al.* 2001)

(riparian zone)	cover next to water.	introduced vegetation.	vegetation. One side cleared & other wide native vegetation.	vegetation. Wide area (>10m)		water stratification as well as food and habitat for invertebrate prey	
Banks suitable for burrowing	Low, eroded, gently sloping banks.	Limited potential burrowing sites.	Some potential burrowing sites in area. One side steep, other shallow sloping.	Extensive potential burrowing sites.	Steep, consolidated banks along both sides, >1m high	Earthen banks which have a relatively vertical profile at the water surface (>1m high) allow burrow construction	(Woon 1995; Grant and Temple-Smith 1998; Rohweder and Baverstock 1999; Serena *et al.* 2001)
Instream complexity	No snags, boulders, or vegetation over water. Silt or sandy substrate. Could be rock or concrete lined channel.	Occasional snag or rocks. No overhanging vegetation. Limited benthic complexity.	Some snags & rocks present. Some aquatic macrophytes & overhanging vegetation. Some gravel/cobbled substrate.	Lots of snags, logs, rocks. Mostly cobbled substrate. Lots of aquatic macrophytes and overhanging vegetation.	Frequent snags, logs, rocks. Cobbled, rocky and gravel substrate throughout. Extensive aquatic & overhanging vegetation.	"Excellent" category are the preferred foraging substrates for platypuses and support abundant macroinvertebrate prey	(Grant and Temple-Smith 1998; Serena *et al.* 2001; McLachlan-Troup *et al.* 2010)

Channel complexity	Straightened stream. Uniform depth (i.e., all shallow). Could be irrigation channel etc.	Slight variation in depth, not artificially straightened.	Occasional riffle or bend and variation in depth.	Variation in depth in pools and riffles. Variety of habitats (i.e., at least 2). Bends present.	Riffles and pools of varying depth. Winding channel.	Excellent category is the preferred foraging depth profile for platypuses. It reduces foraging energetics, reduces predation risk, and promotes habitat diversity to support diverse and abundant macro-invertebrates.	(Ellem *et al.* 1998; Grant and Temple-Smith 1998; Serena *et al.* 2001; Milione and Harding 2009; McLachlan-Troup *et al.* 2010)

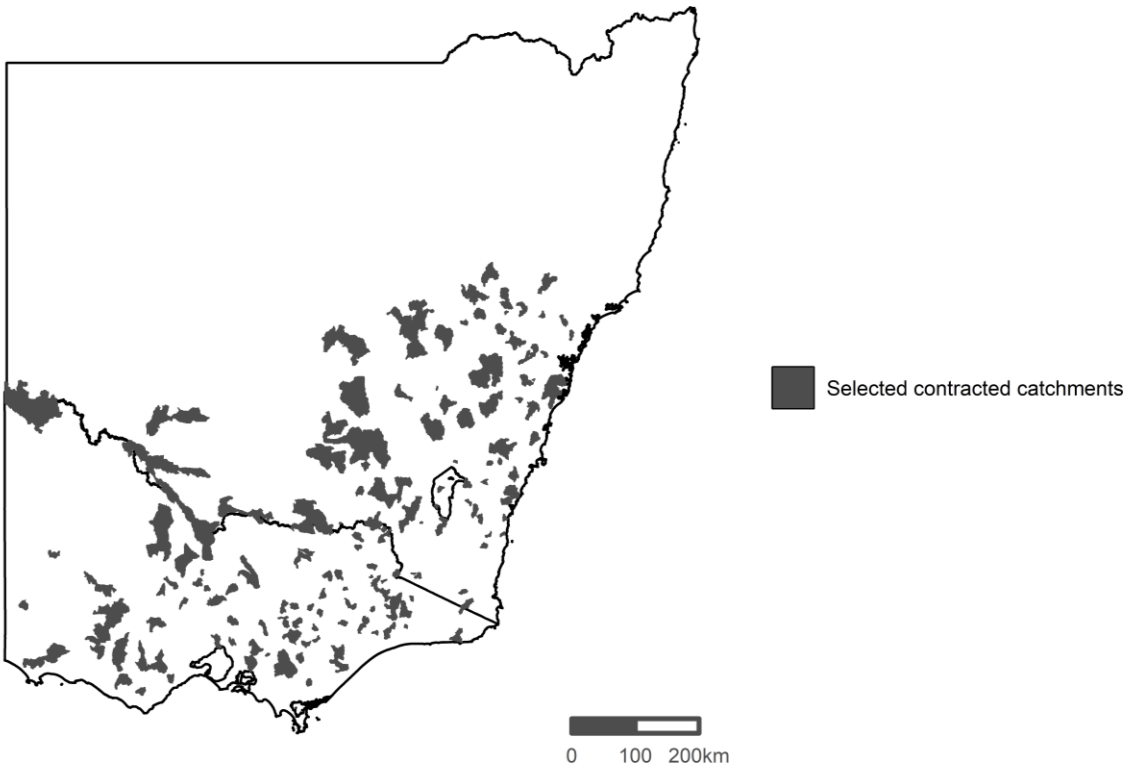

Figure C.1: Map of selected contracted catchments from stratification. Grey polygons are selected contracted catchments across Victoria and NSW

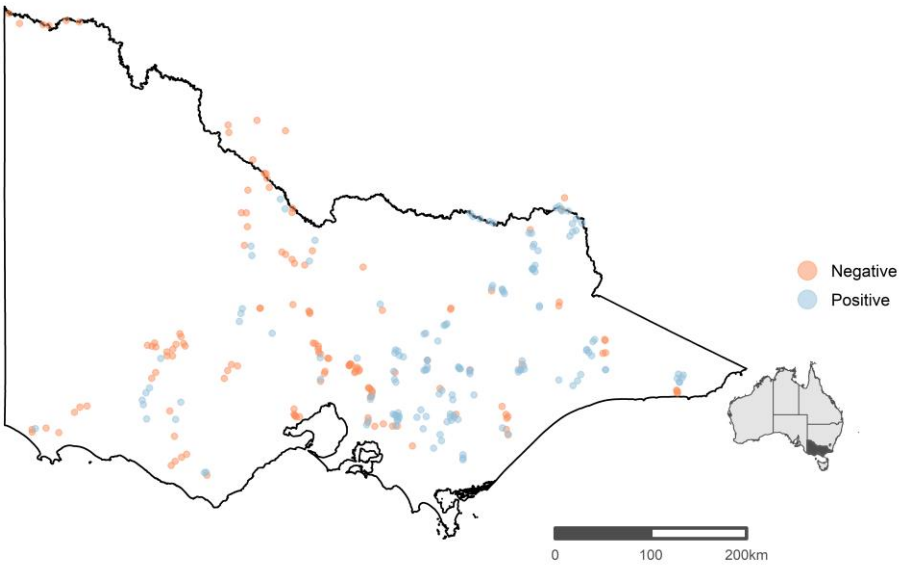

Figure C.2 Map of sampled sites for Victorian dataset with detection. Sites are either in orange (negative) or blue (positive)

JAGS code

```
model {
   # Priors and model for params
   int.psi ~ dunif(0,1)         # Intercept of occupancy probability
   int.theta ~ dunif(0,1) # Intercepts availability probability
   int.p ~ dunif(0,1)     # Intercepts detection probability (1-PCR error)

   sigma_site~dunif(0,1)
   sigma_catch~dunif(0,1)

   # convert it to a precision (1 / variance)

   tau_site<-pow(sigma_site, -2)
   tau_catch<-pow(sigma_site, -2)

   # define the random effect; one value for each site, all share the same precision

   for (i in 1:n.site){
   gamma_site[i]~dnorm(0,tau_site)
   }
   for (i in 1:n.catch){
   gamma_catch[i]~dnorm(0,tau_catch)
   }

   beta_hd~ dnorm(0, 0.1)

   beta_erosion[1] <- 0
   for (k in 2:5) {
   beta_erosion[k] ~ dnorm(0, 0.1)
   }

   beta_b_veg[1] <- 0
   for (k in 2:5) {
```

```
beta_b_veg[k] ~ dnorm(0, 0.1)
}

beta_b_bur[1] <- 0
for (k in 2:5) {
beta_b_bur[k] ~ dnorm(0, 0.1)
}

beta_lu3 ~ dnorm(0, 0.1)
beta_lu4 ~ dnorm(0, 0.1)
beta_lu5 ~ dnorm(0, 0.1)
beta_ro ~ dnorm(0, 0.1)

#quad
beta_lu4_quad~ dnorm(0, 0.1)
beta_lu5_quad~ dnorm(0, 0.1)
beta_runoff_quad~ dnorm(0, 0.1)

# 'Likelihood' (or basic model structure)
for (i in 1:n.site){
# Occurrence in pond i
z[i] ~ dbern(psi[i])
logit(psi[i]) <- logit(int.psi)+

beta_lu3 * lu3[i] +
beta_lu4 * lu4[i] +
beta_lu4 * lu4[i] +
beta_hd*hd[i]+
beta_erosion[erosion[i]]+
beta_b_veg[b_veg[i]]+
beta_b_bur[b_bur[i]]+
beta_ro*runoff[i]+

beta_lu4_quad*pow(lu4[i],2) +
```

```
    beta_lu5_quad*pow(lu5[i],2) +
    beta_runoff_quad*pow(runoff[i],2) +
    gamma_site[site[i]]+ gamma_catch[catchment_fact[i]]



    for (j in 1:n.samples){
    # Occurrence in sample j
    a[i,j] ~ dbern(mu.a[i,j])
    mu.a[i,j] <- z[i] * theta[i,j]
    logit(theta[i,j]) <- logit(int.theta)

    for (k in 1:n.pcr){
    # PCR detection error process in sample k

    y[i,j,k] ~ dbern(mu.y[i,j,k])
    mu.y[i,j,k] <- a[i,j] * p[i,j,k]
    logit(p[i,j,k]) <- logit(int.p)
    }
    }
    }# end model

    ",fill=TRUE)
sink()

zst <- apply(GAPS_detection_array, 1, max)        # inits for presence (z)
zst<-ifelse(is.na(zst),1,zst)
ast <- apply(GAPS_detection_array, c(1,2), max)   # inits for availability (a)
ast<-ifelse(is.na(ast),1,ast)
inits <- function() list(z = zst, a = ast,int.psi = 0.5)

# Parameters monitored
params <- c("int.psi", "int.theta", "int.p",
            "p", "theta","psi","z",
```

```
        "beta_hd","beta_erosion","beta_b_veg","beta_b_bur",
        "beta_ro","beta_lu3","beta_lu4","beta_lu5",
        "beta_runoff_quad","beta_lu4_quad","beta_lu5_quad")
ni <- 350000  ;   nt <- 1000  ;   nb <- 50000  ;   nc <- 3



# Call WinBUGS and summarize posterior
out_gaps_wo_flow7<- jags(gaps_wo_flow7, inits, params, "gaps_wo_flow7.txt", n.chains = nc, n.thin
= nt, n.iter = ni, n.burnin = nb)
```

GAPS dataset residual plots

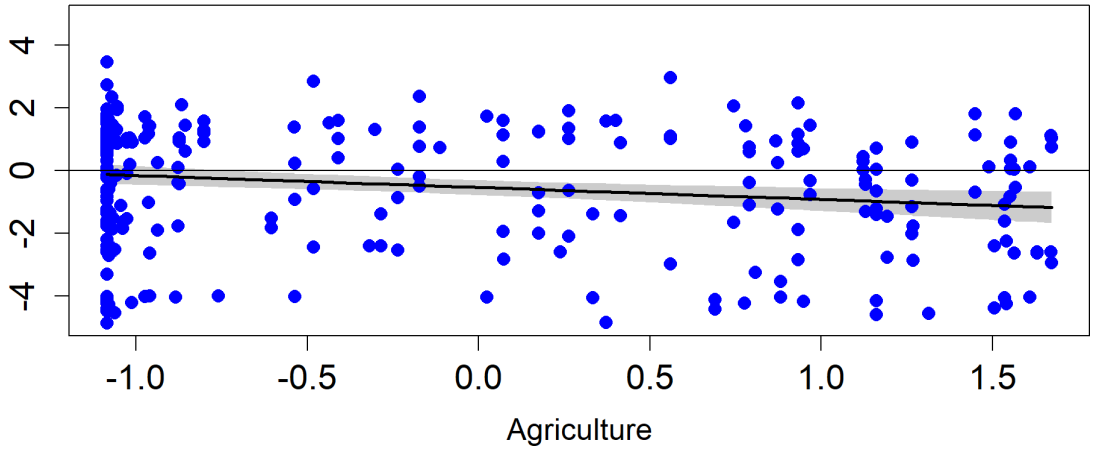

Figure C.3: Plot of Dunn-Smyth residuals (smooth curves and 95% confidence bands) against missing covariate (agriculture).

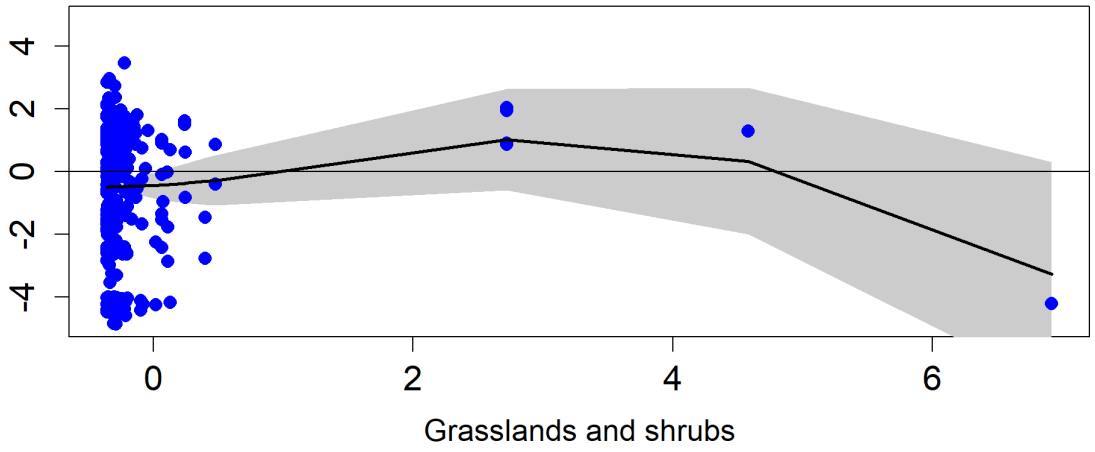

Figure C.4: Plot of Dunn-Smyth residuals (smooth curves and 95% confidence bands) against missing covariate (Grasslands and shrubs)

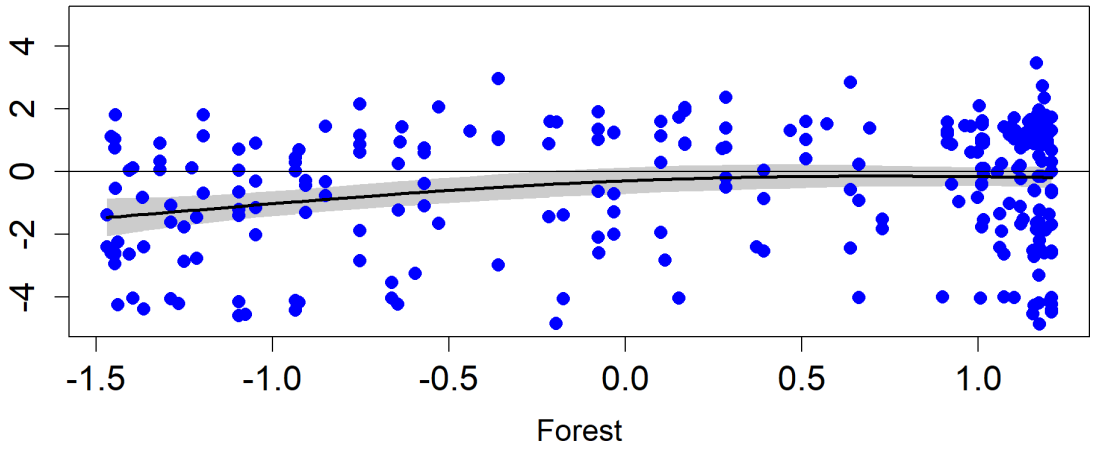

Figure C.5: Plot of Dunn-Smyth residuals (smooth curves and 95% confidence bands) against missing covariate (Forest)

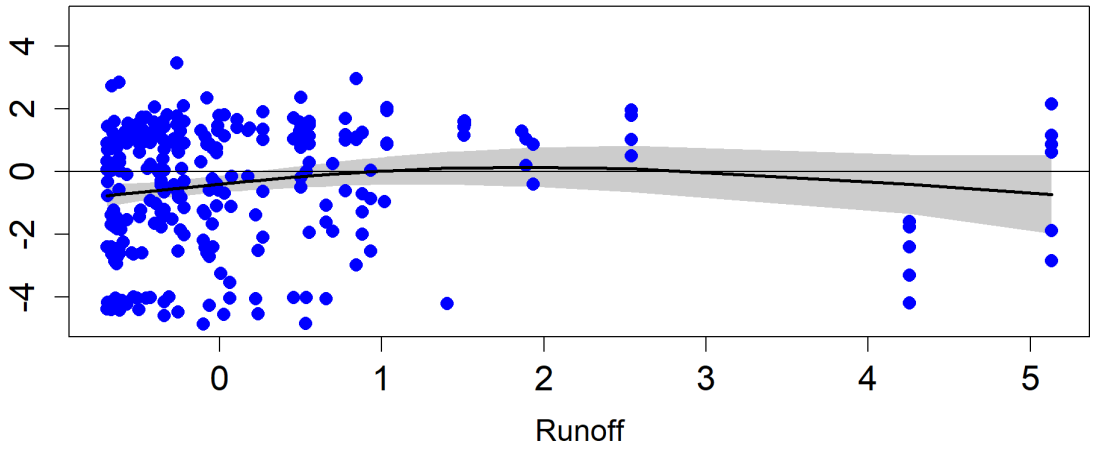

Figure C.6: Plot of Dunn-Smyth residuals (smooth curves and 95% confidence bands) against missing covariate (runoff)

Victorian dataset residual plots

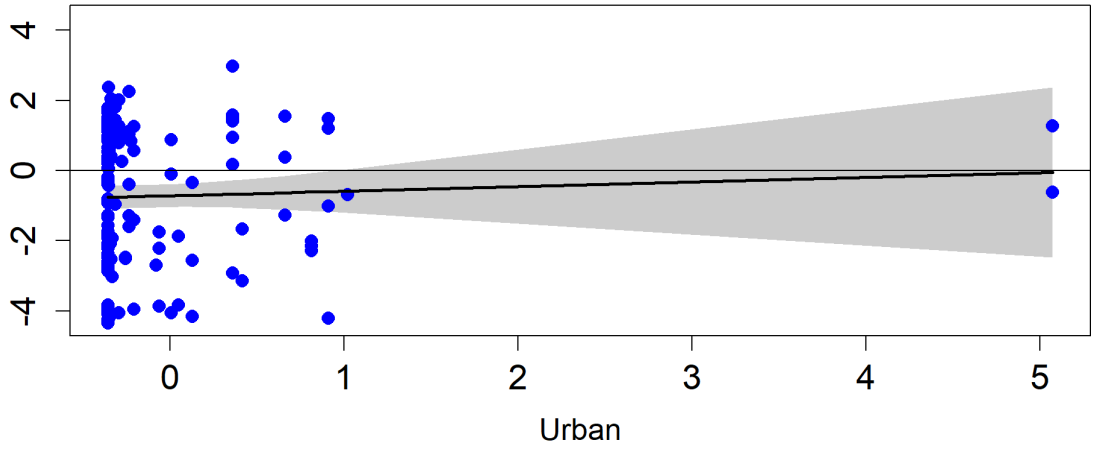

Figure C.7: Plot of Dunn-Smyth residuals (smooth curves and 95% confidence bands) against missing covariate (Urban)

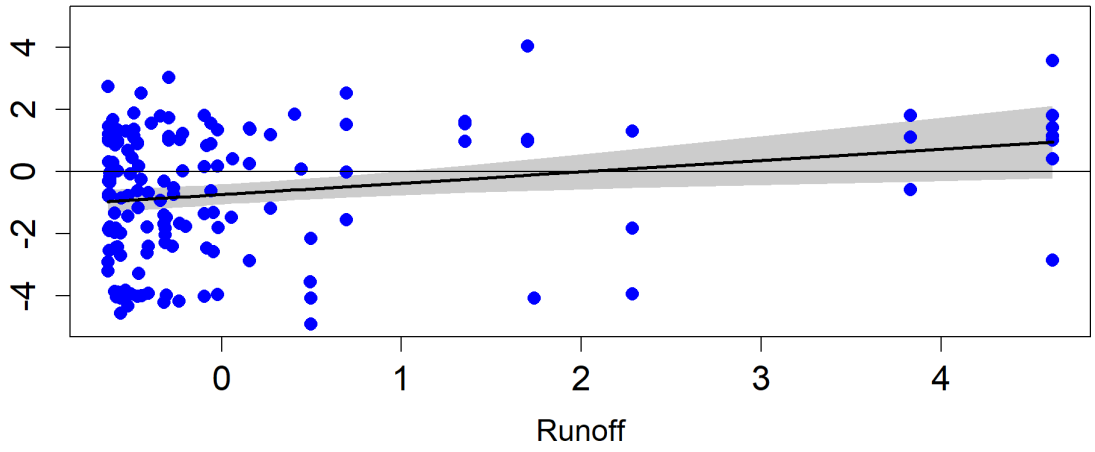

Figure C.8: Plot of Dunn-Smyth residuals (smooth curves and 95% confidence bands) against missing covariate (Runoff)

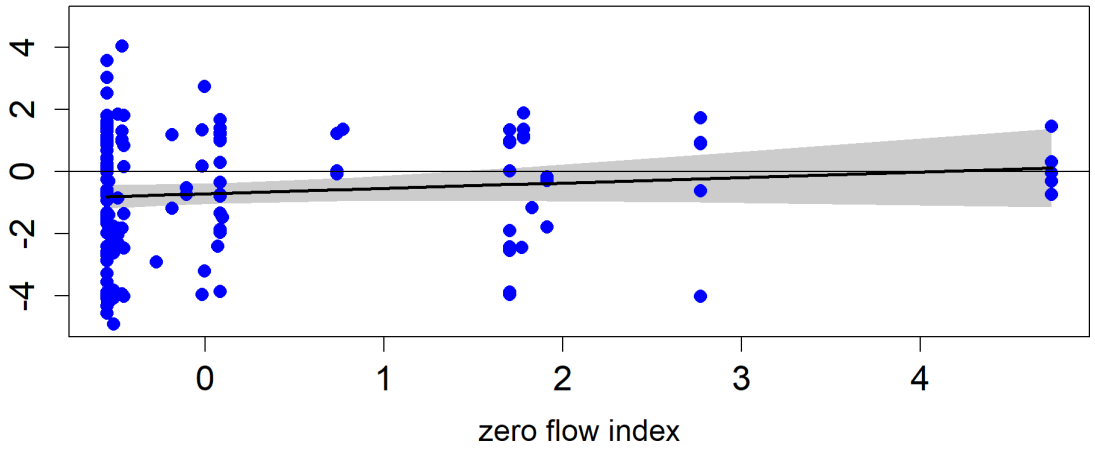

Figure C.9: Plot of Dunn-Smyth residuals (smooth curves and 95% confidence bands) against missing covariate (Zero-flow)

Table C.2: Confusion matrix of observed vs predicted occupancy (z) for each site

	Observed		
Predicted		*Present*	*Absent*
	Present	*250*	*34*
	Absent	*22*	*198*

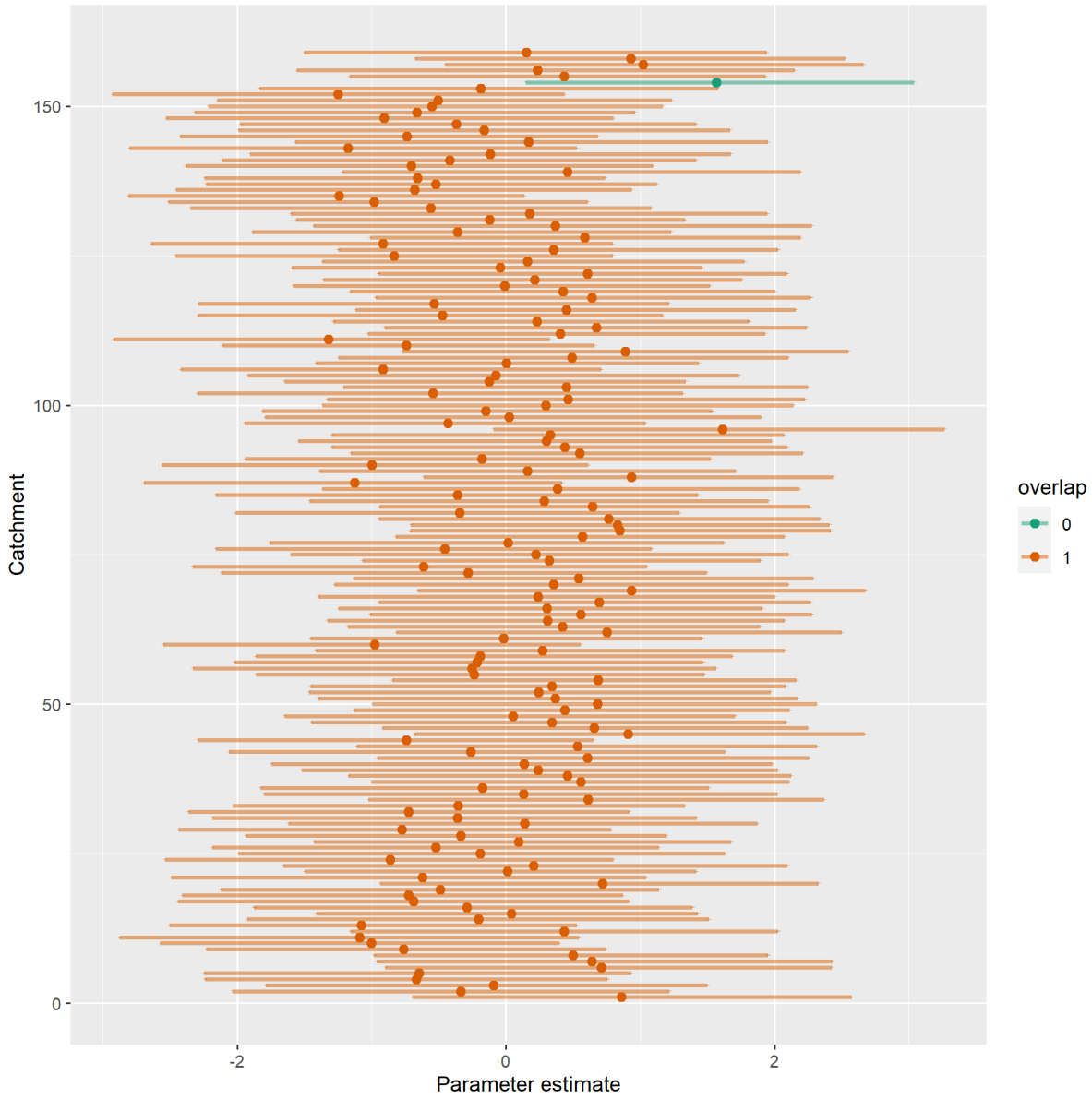

Figure C.10: Parameter estimates for the random effect of a catchment on occupancy probabilities. Points are the parameter mean estimate and lines are the 95% credible intervals. Green lines do not overlap zero orange lines overlap zero.

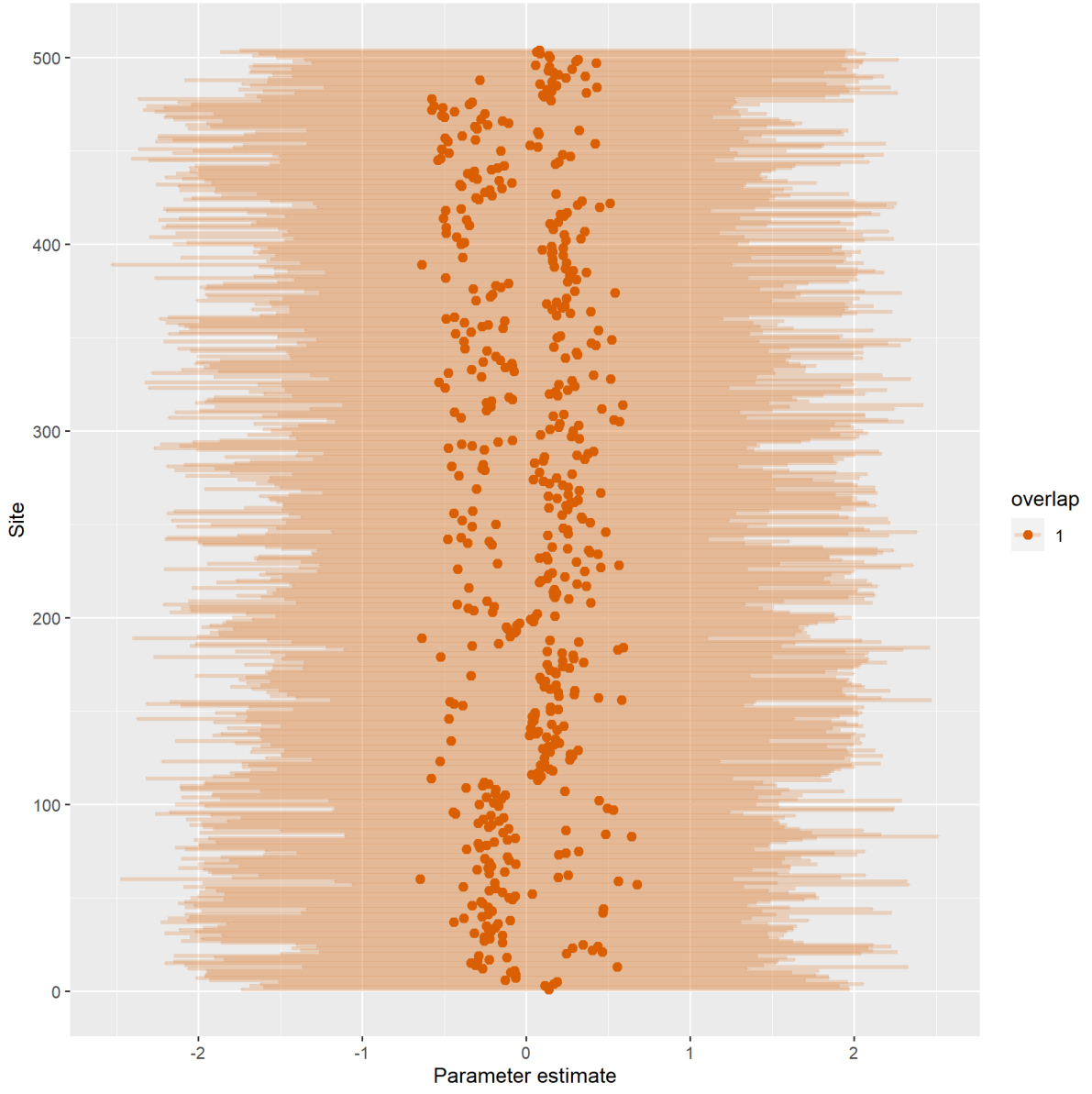

Figure C.11: Parameter estimates for the random effect of a site on occupancy probabilities. Points are the parameter mean estimate and lines are the 95% credible intervals. Green lines do not overlap zero orange lines overlap zero.

Response of occupancy to environmental covariates

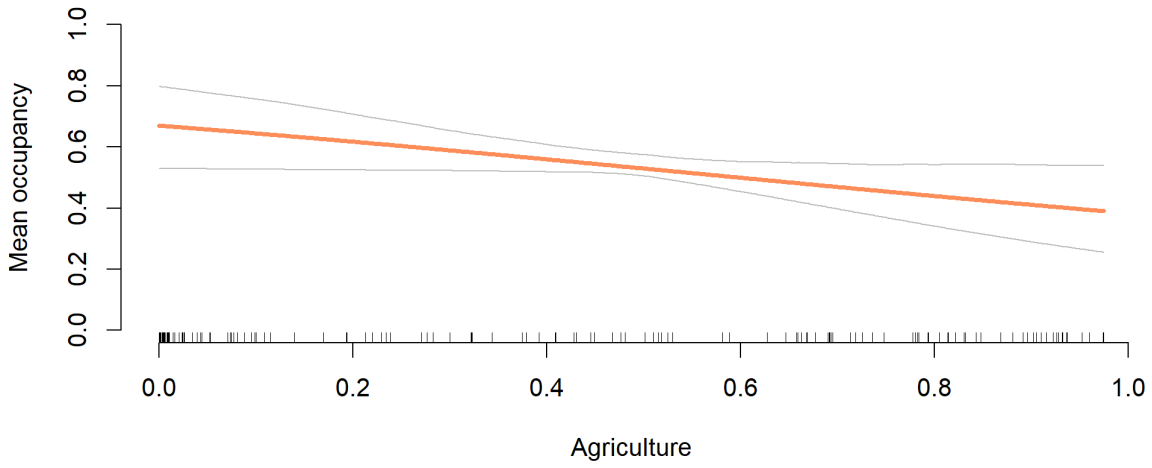

Figure C.12: Estimated platypus occupancy probability as a function of the proportion of a catchment covered by grasslands and shrubs including mean (orange line) and 95% credible intervals (grey lines)

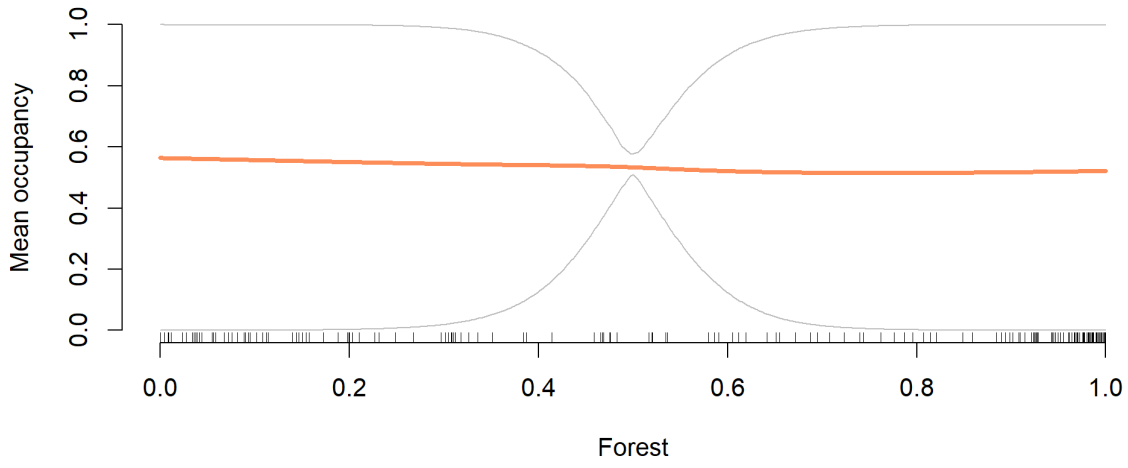

Figure C.13: Estimated platypus occupancy probability as a function of the proportion of a catchment covered by forest including mean (orange line) and 95% credible intervals (grey lines)

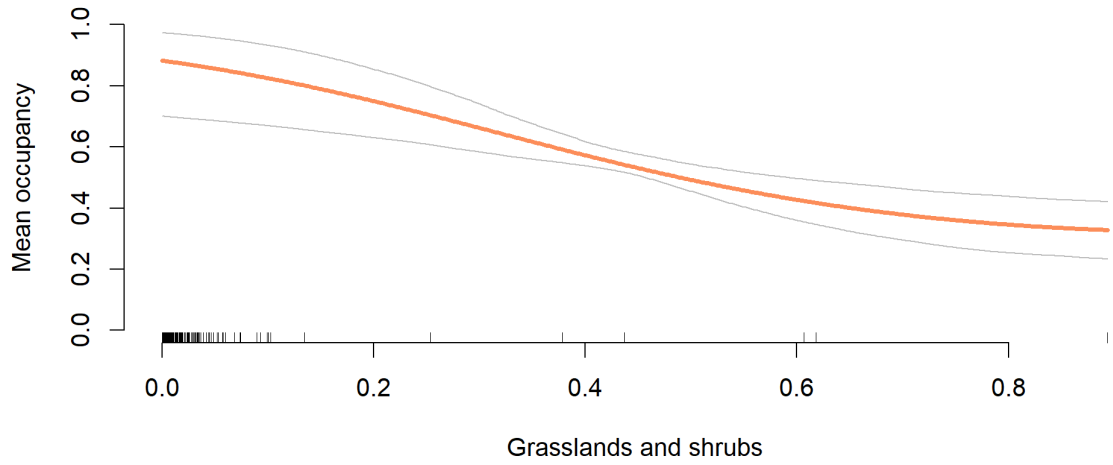

Figure C.14: Estimated platypus occupancy probability as a function of the proportion of a catchment covered by grasslands and shrubs including mean (orange line) and 95% credible intervals (grey lines)

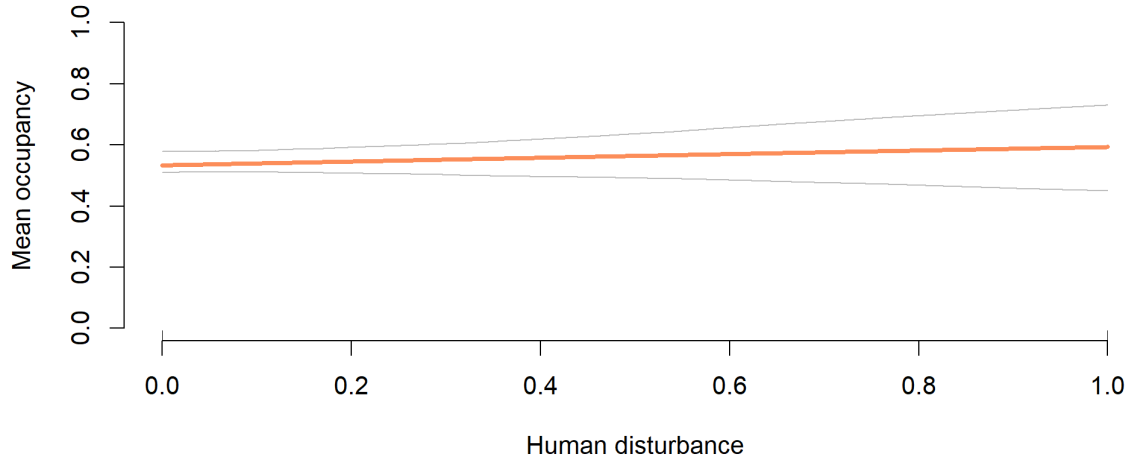

Figure C.15: Predicted platypus occupancy probability as a function of human disturbance including mean (orange line) and 95% credible intervals (grey lines)

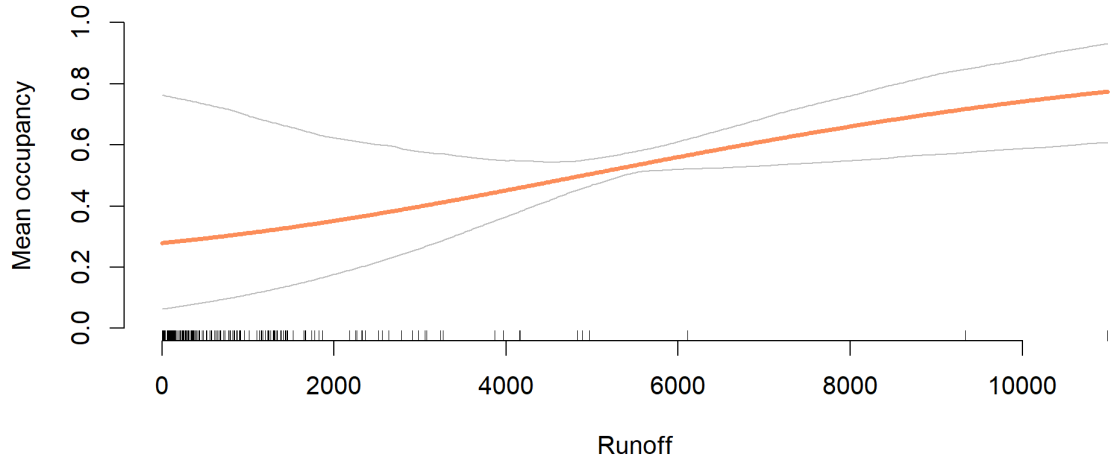

Figure C.16: Estimated platypus occupancy probability as a function of the proportion of the amount of runoff in a contracted catchment including mean (orange line) and 95% credible intervals (grey lines)

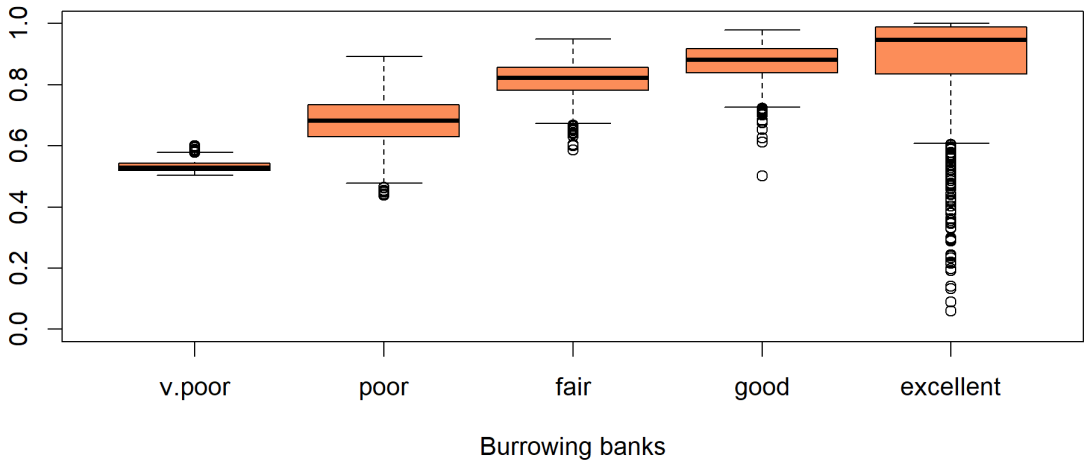

Figure C.17: Estimated platypus occupancy probability as a function of the condition of banks for burrowing (reference category=very poor)

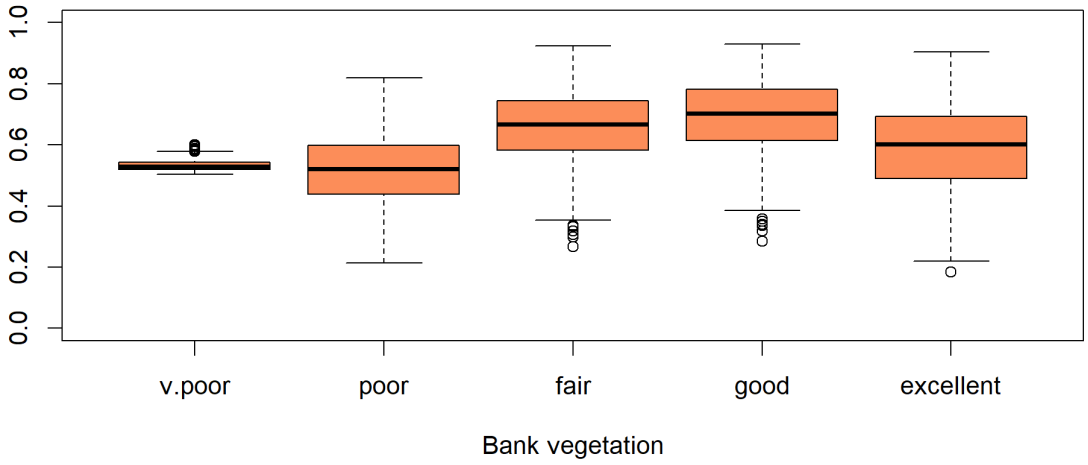

Figure C.18: Estimated platypus occupancy probability as a function of the condition of bank vegetation (reference category=very poor)

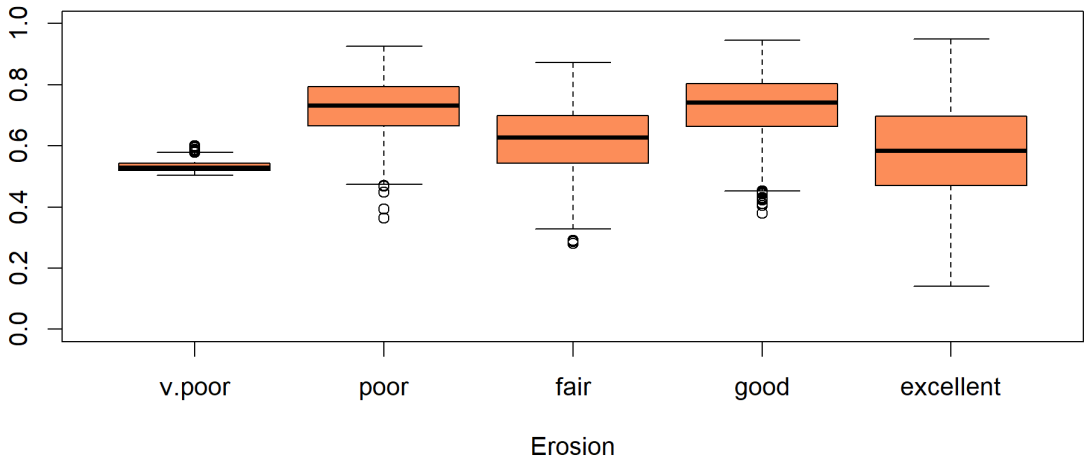

Figure C.19: Estimated platypus occupancy probability as a function of the condition of bank erosion (reference category=very poor)

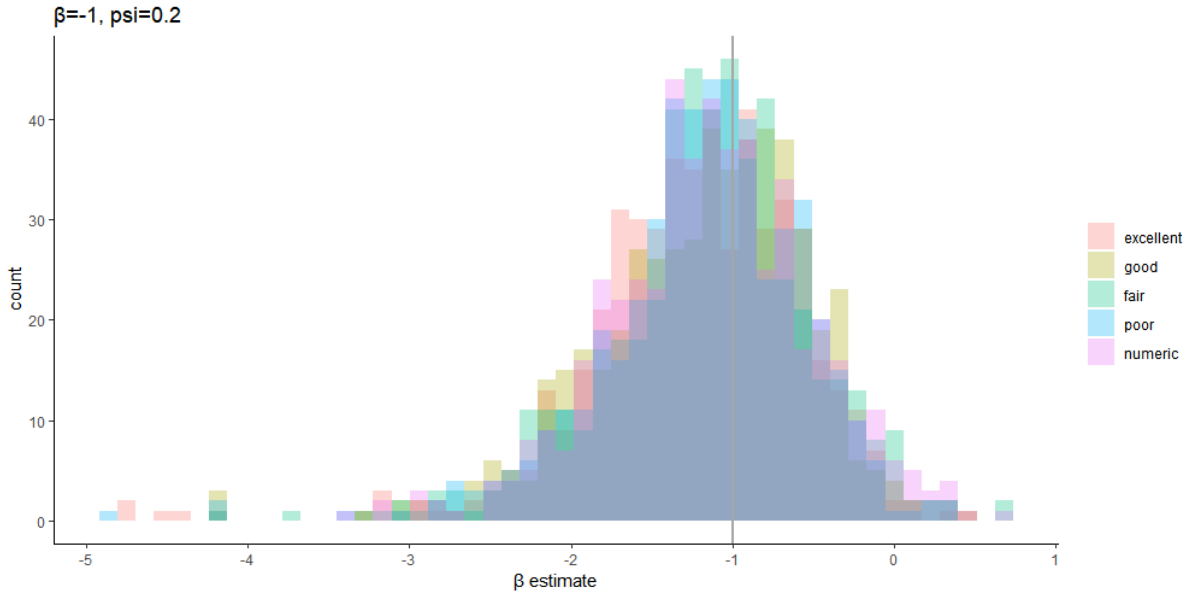

Figure C.20: Simulated estimates of β parameter where occupancy probability = 0.2, availability probability = 0.9, detection probability = 0.2 and β = −1 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = −1.

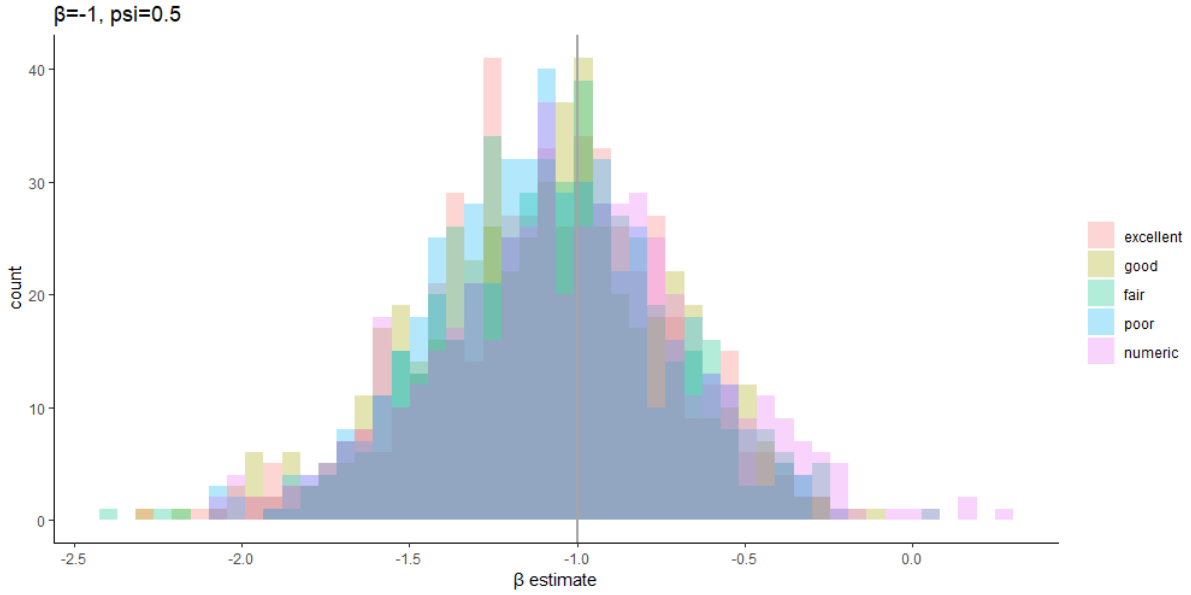

Figure C.21: Simulated estimates of β parameter where occupancy probability = 0.5, availability probability = 0.9, detection probability = 0.2 and β = −1 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = −1.

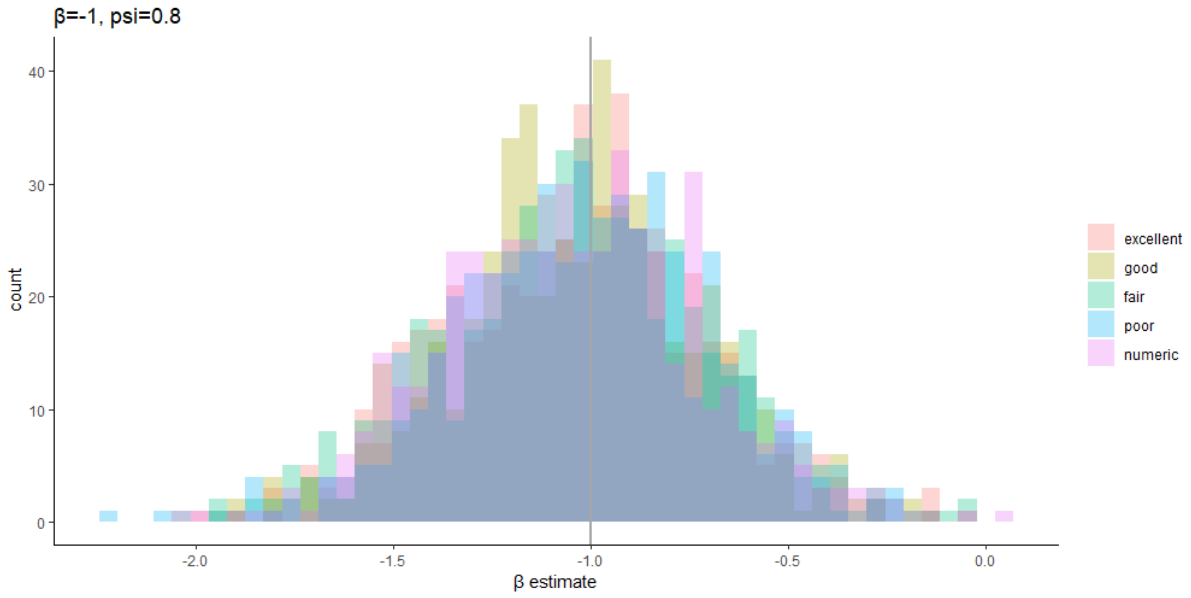

Figure C.22: Simulated estimates of β parameter where occupancy probability = 0.8, availability probability = 0.9, detection probability = 0.2 and β = −1 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = −1.

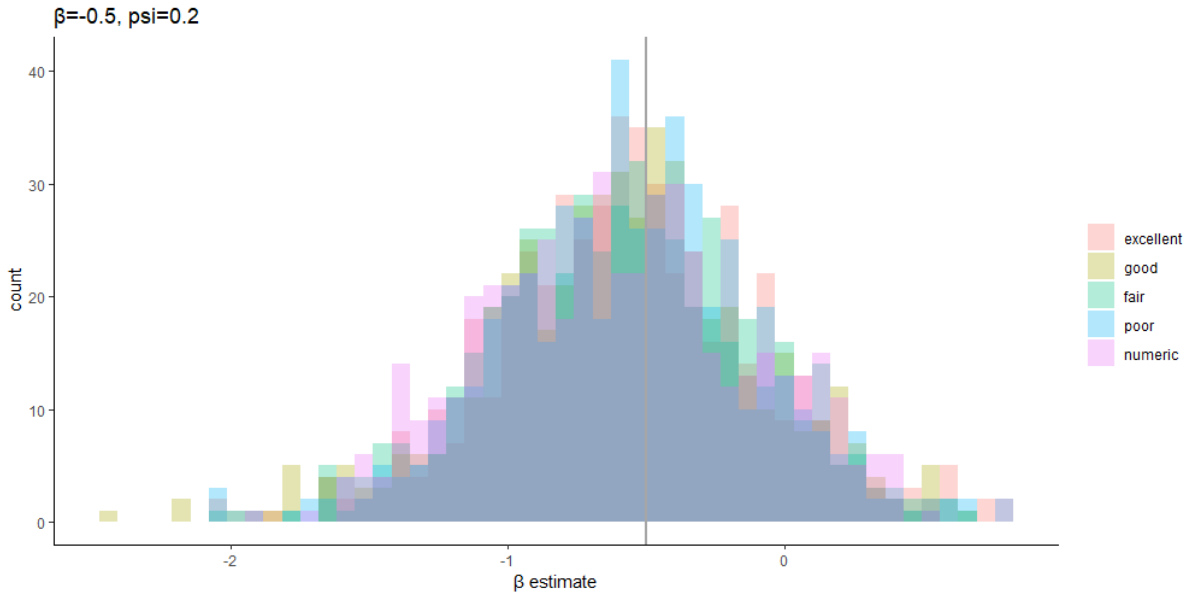

Figure C.23: Simulated estimates of β parameter where occupancy probability = 0.2, availability probability = 0.9, detection probability = 0.2 and β = −0.5 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = −0.5.

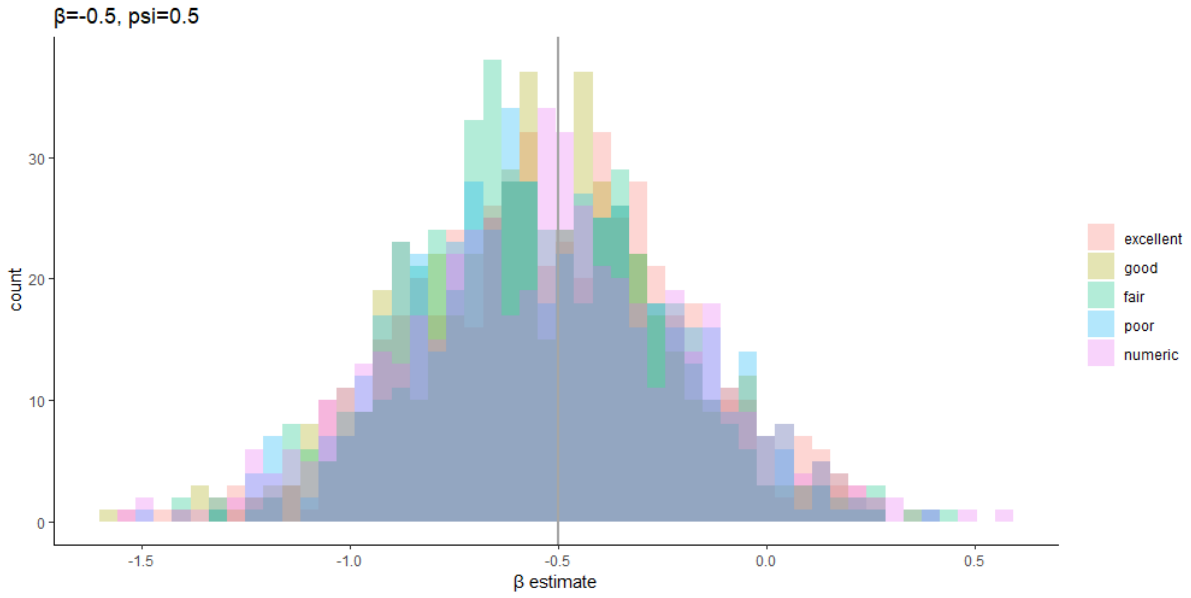

Figure C.24: Simulated estimates of β parameter where occupancy probability = 0.5, availability probability = 0.9, detection probability = 0.2 and β = −0.5 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = −0.5.

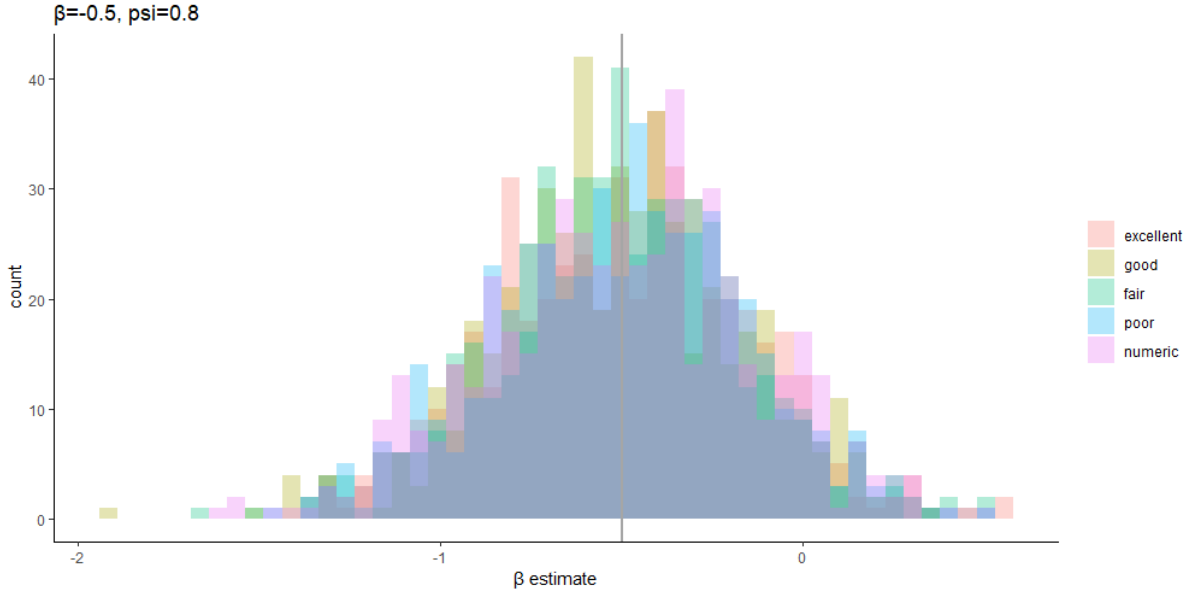

Figure C.25: Simulated estimates of β parameter where occupancy probability = 0.8, availability probability = 0.9, detection probability = 0.2 and β = −0.5 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = −0.5.

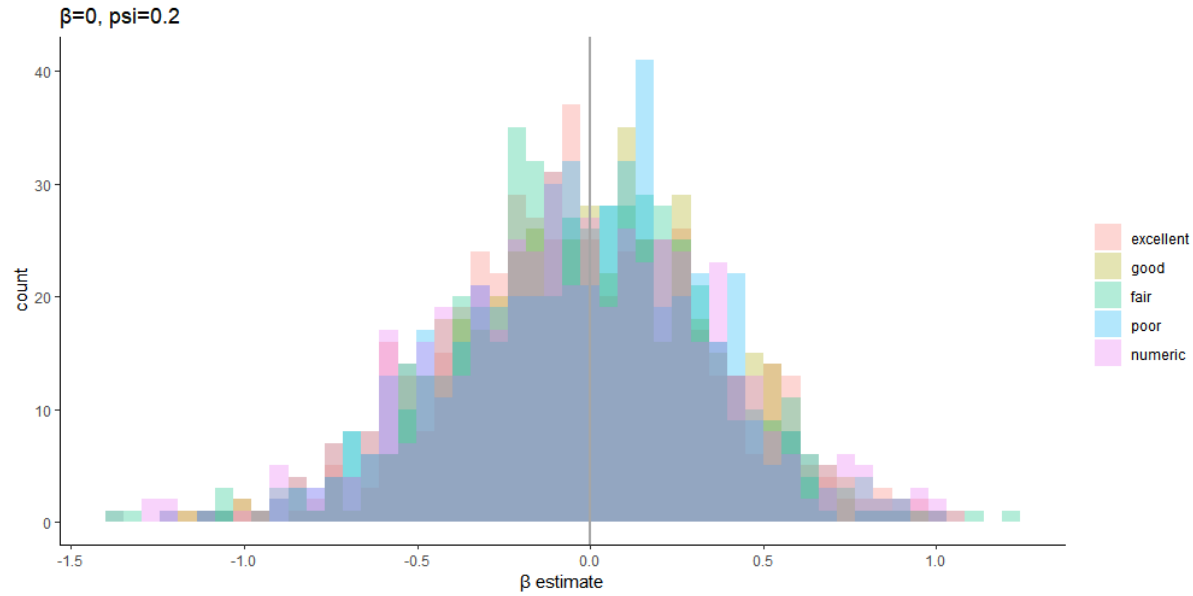

Figure C.26: Simulated estimates of β parameter where occupancy probability = 0.2, availability probability = 0.9, detection probability = 0.2 and β = 0 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = 0.

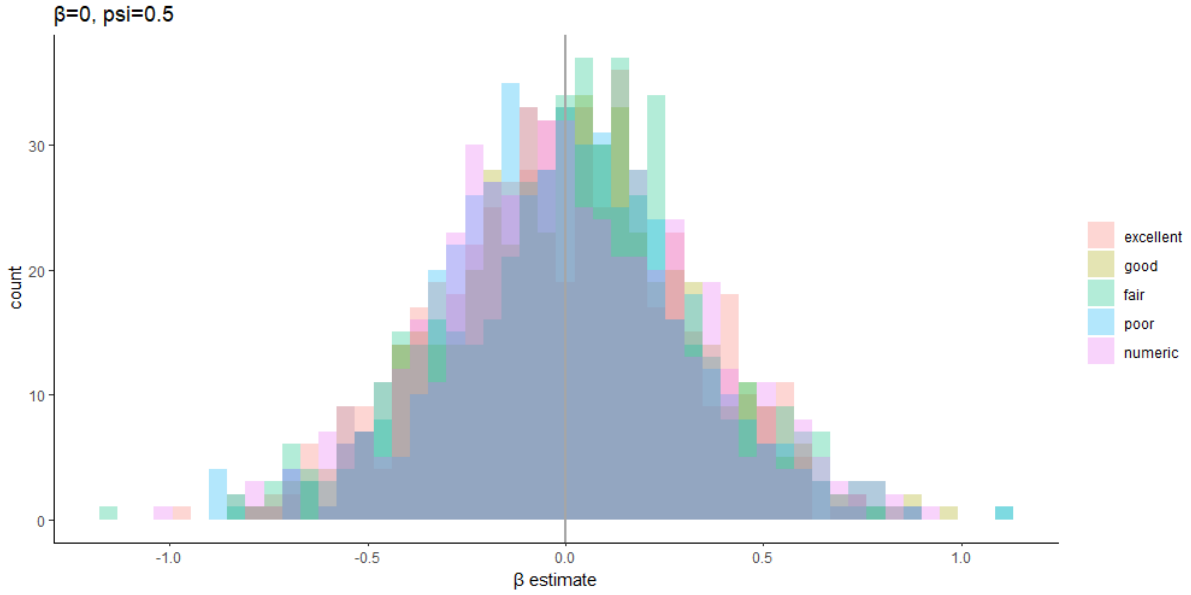

Figure C.27: Simulated estimates of β parameter where occupancy probability = 0.5, availability probability = 0.9, detection probability = 0.2 and β = 0 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = 0.

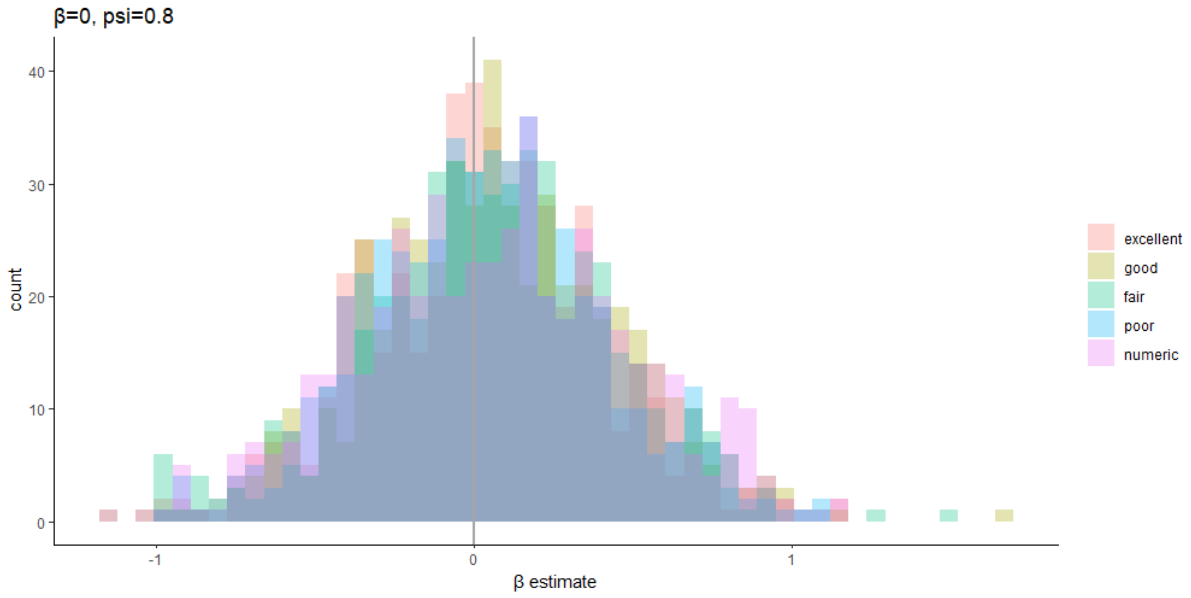

Figure C.28: Simulated estimates of β parameter where occupancy probability = 0.8, availability probability = 0.9, detection probability = 0.2 and β = 0 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = 0.

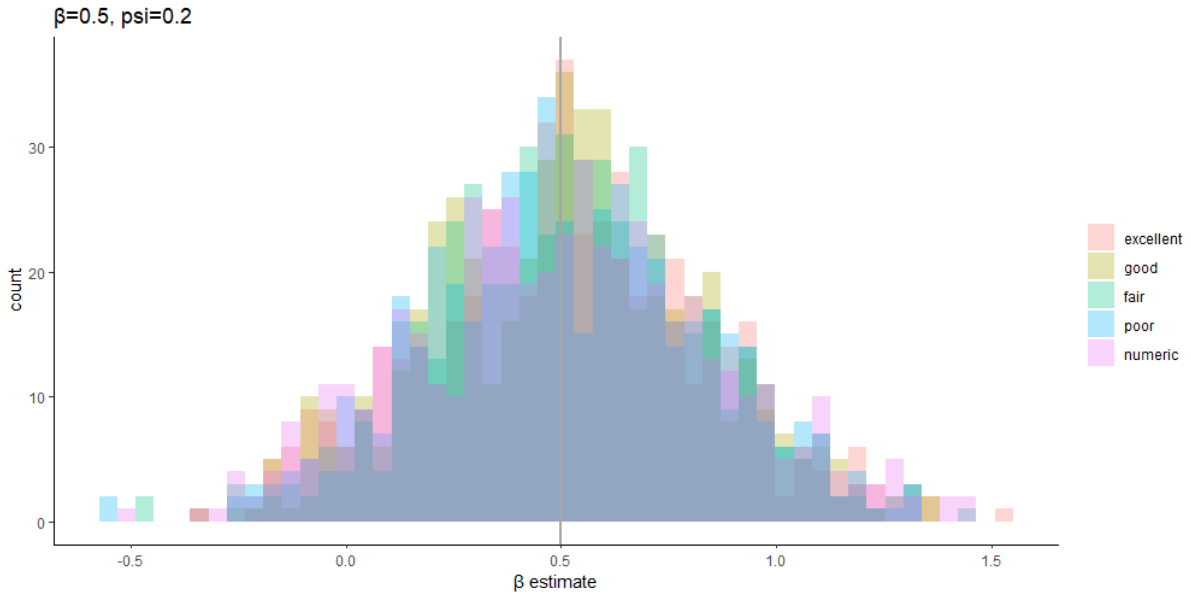

Figure C.29: Simulated estimates of β parameter where occupancy probability = 0.2, availability probability = 0.9, detection probability = 0.2 and β = 0.5 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = 0.5.

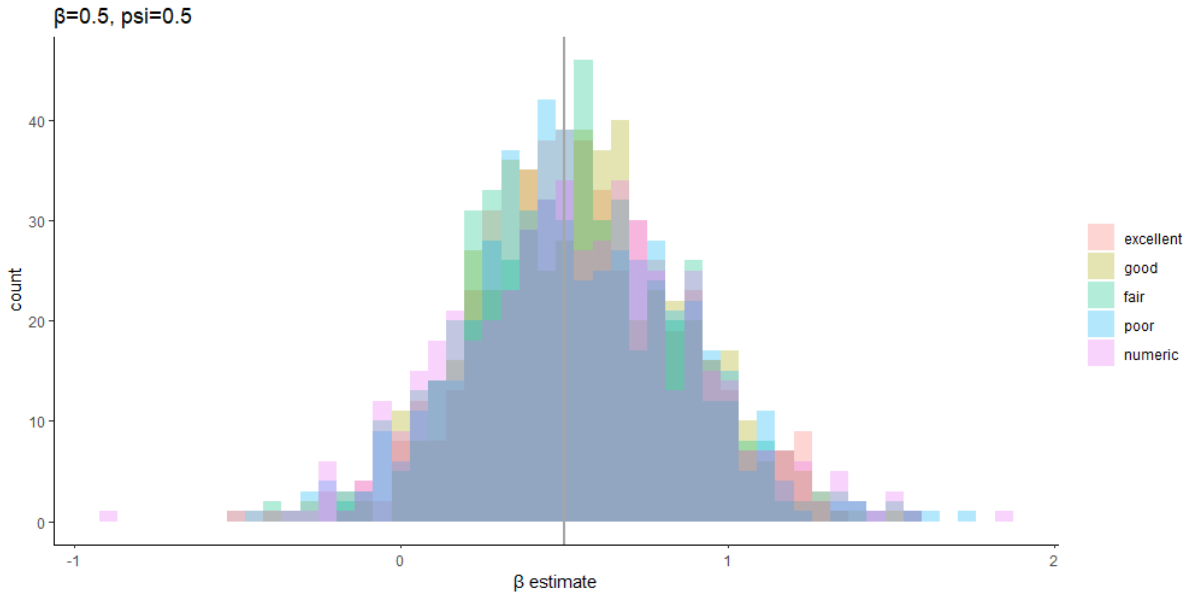

Figure C.30: Simulated estimates of β parameter where occupancy probability = 0.5, availability probability = 0.9, detection probability = 0.2 and β = 0.5 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β =0.5.

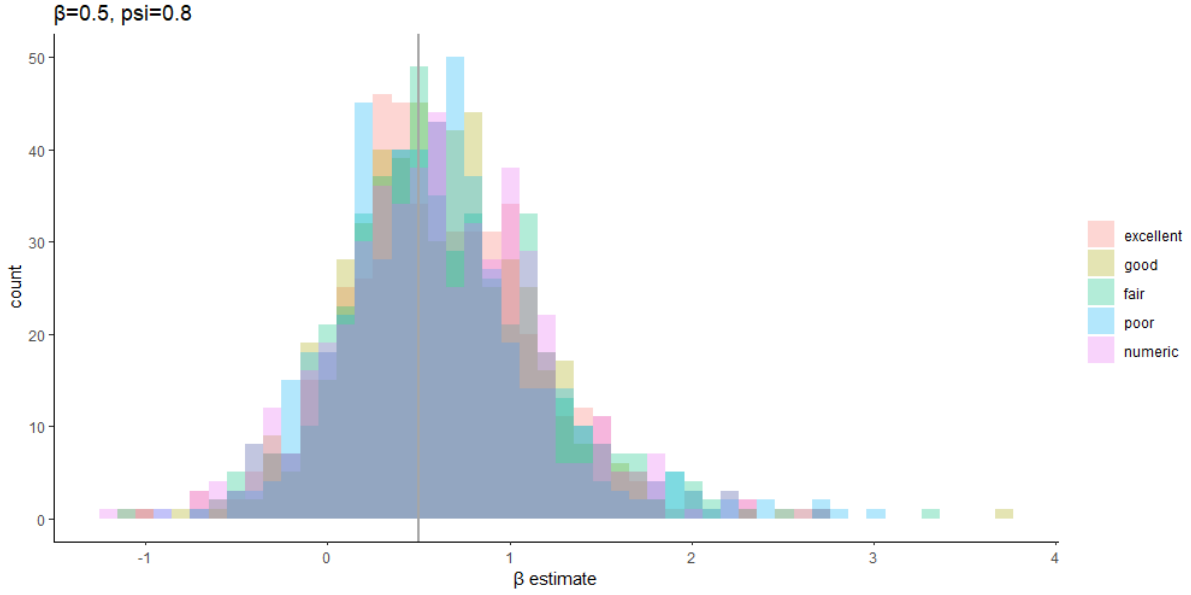

Figure C.31: Simulated estimates of β parameter where occupancy probability = 0.8, availability probability = 0.9, detection probability = 0.2 and β = 0.5 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = 0.5.

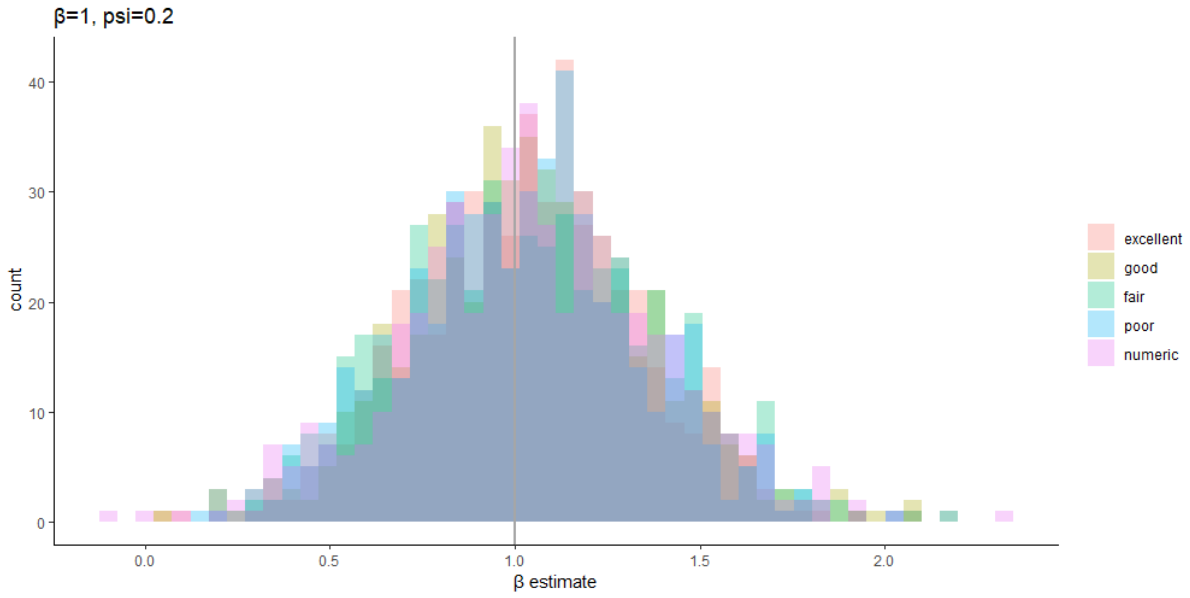

Figure C.32: Simulated estimates of β parameter where occupancy probability = 0.2, availability probability = 0.9, detection probability = 0.2 and β = 1 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = 1.

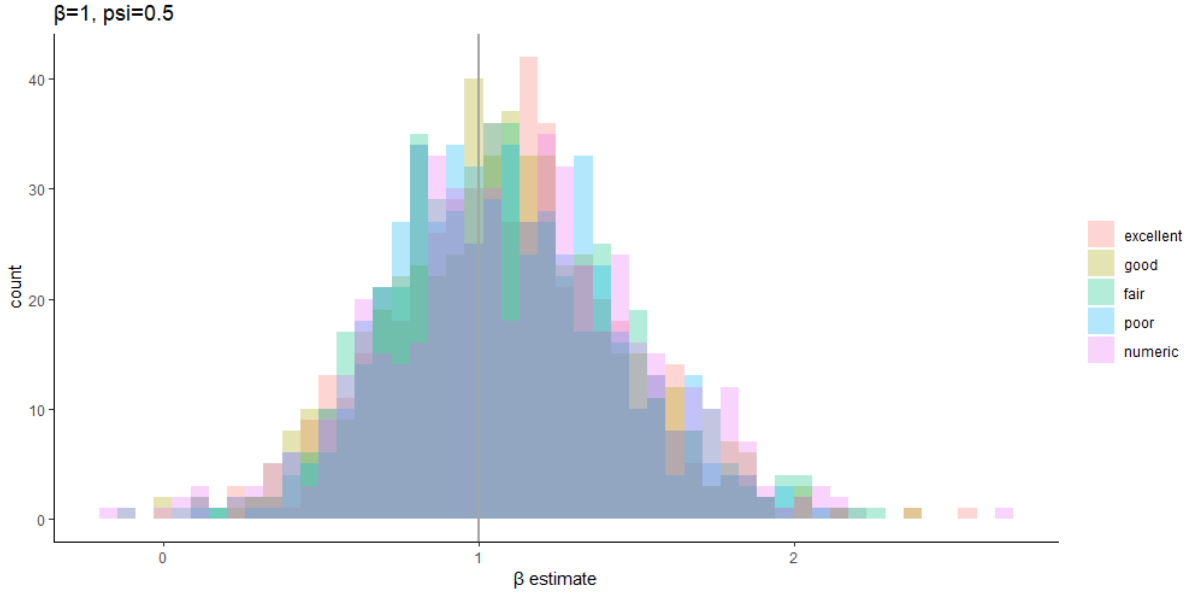

Figure C.33: Simulated estimates of β parameter where occupancy probability = 0.5, availability probability = 0.9, detection probability = 0.2 and β = 1 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = 1.

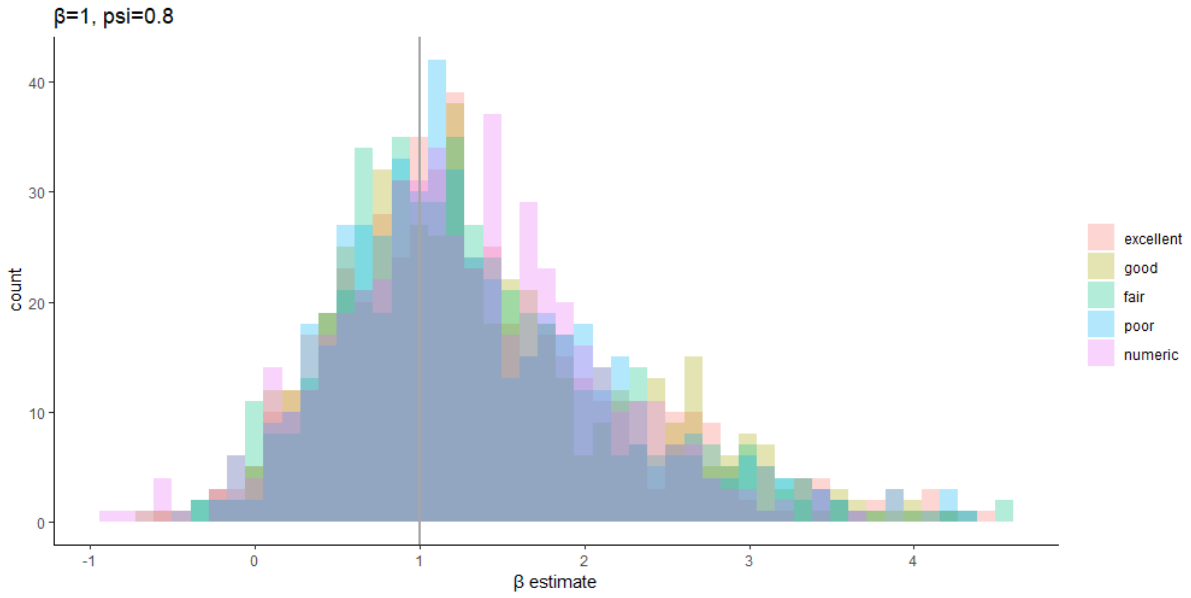

Figure C.34: Simulated estimates of β parameter where occupancy probability = 0.8, availability probability = 0.9, detection probability = 0.2 and β = 1 (vertical line). Colours represent the different data types, an ordinal covariate with 4 categories (excellent, good, fair, and poor) and a numeric category. All categories were set as β = 1.

Appendix D
Supplementary material for chapter 4

Model 2

Equation 1. $$z_{ij} \sim Bernoulli(\psi_{i,j})$$

Equation 1 describes the occupancy of the species at site *j* during visit *i* (z_{ij}=1 where the species is present, and z_i=0 where the species is absent), given the mean probability of occurrence for that species across all visits ($\psi_{i,j}$).

Equation 2. $$logit(\psi_{ij}) = logit(intercept.\psi_j) + \beta_{after1,i,j} + \beta_{after2,ij} + \beta_{cov,i,j}$$

Equation 2 shows that the probability of species occurrence at site *j* is affected by the sites before-fire occupancy probability (ψ_j), whether the visit is since fire ($after1, i, j$ and $after2, i, j$, binary indicator variables) and different covariates $\beta_{cov,i,j}$ (see Table D.1).

Equation 3. $$a_{ijk}|z_{ij} \sim Bernoulli\ (z_{ij}\theta_{ijk})$$

Equation 3 describes the presence of the species DNA in a water sample ($a_{ijk}|z_{ij} = 1$ where a species DNA is present, and $a_{ijk}|z_{ij}$ =0 when it is absent, during visit *i*, at site *j*, for water sample *k)*.

Equation 4. $$cloglog(\theta_{ijk}) = cloglog(intercept.\theta) + \beta_{volume} * volume_{i,j,k}$$

Equation 4 shows that the probability of species DNA being available to detect in water sample *k* is affected by the volume of water collected during visit *I*, site *j* and in water sample *k*. A cloglog link function that rescales the linear predictor to the interval [0, 1] was used.

Equation 6. $$y_{ijkl}|a_{ijk} \sim Bernoulli(a_{ijk}p_{ijkl})$$

Equation 5 describes the observed detection data ($y_{ijkl}|a_{ijk}$=1 where the species is detected, and $y_{ijk}|z_{ij}$=0 where the species is not detected), during visit *i*, at site *j* for water sample *k* and in replicate *l*. This is a function of the occurrence (z_{ij}) of the species at site *j* during visit *i*, as well as the probability of detecting an individual (p_{ijkl}) at site *j* in water sample *k* in replicate *l*, given it is available to detect (a_{ijk}).

Table D.1: covariates used in model two

Covariate	**Pre-fire**	**Post-fire 1**	**Post-fire 2**
High severity (proportion of watershed with high severity burnt)	0	Proportion	Proportion
Low severity (proportion of watershed with low severity burnt)	0	Proportion	Proportion
Rainfall (amount of rainfall in the watershed post-fire)	0	rainfall	rainfall
volume	volume	volume	volume

Table D.2: Severity class classifications for different vegetation types

Community	**Canopy defoliated**	**Canopy intact/partially defoliated**
Forest	High severity (Severity classes: high canopy scorch (>80%) and canopy consumption))	Low severity (Severity classes: low canopy scorch (<20% scorch) and moderate canopy scorch (20-80%))
Woodland	High severity (Severity classes: high canopy scorch (>80%) and canopy consumption))	Low severity (Severity classes: low canopy scorch (<20% scorch) and moderate canopy scorch (20-80%))
Shrubland	High severity (Severity classes: high canopy scorch (>80%) and canopy consumption))	Low severity (Severity classes: low canopy scorch (<20% scorch) and moderate canopy scorch (20-80%))
Grassland/pasture	Burnt (i.e., within the fire perimeter- High severity)	Unburnt (i.e., outside the fire perimeter)

Table D.3: Summary data for water volume in ml collected at each site

	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Visit 1 sample 1	0	150	300	300.7	500	600
Visit 1 sample 2	0	120	300	298	500	600
Visit 2 sample 1	0	50	230	245.6	400	1140
Visit 2 sample 2	0	0	240	235.1	400	600
Visit 3 sample 1	0	125	300	268.8	400	600
Visit 3 sample 2	0	125	300	269.5	400	600

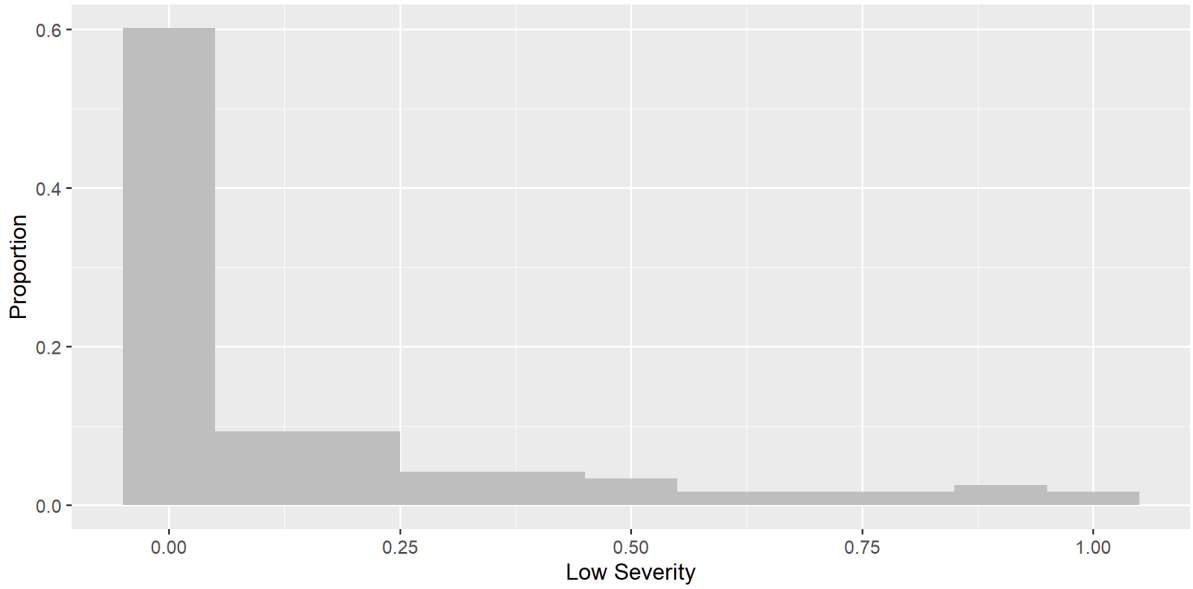

Figure D.1: Histogram of proportion of watersheds covered by low severity fire.

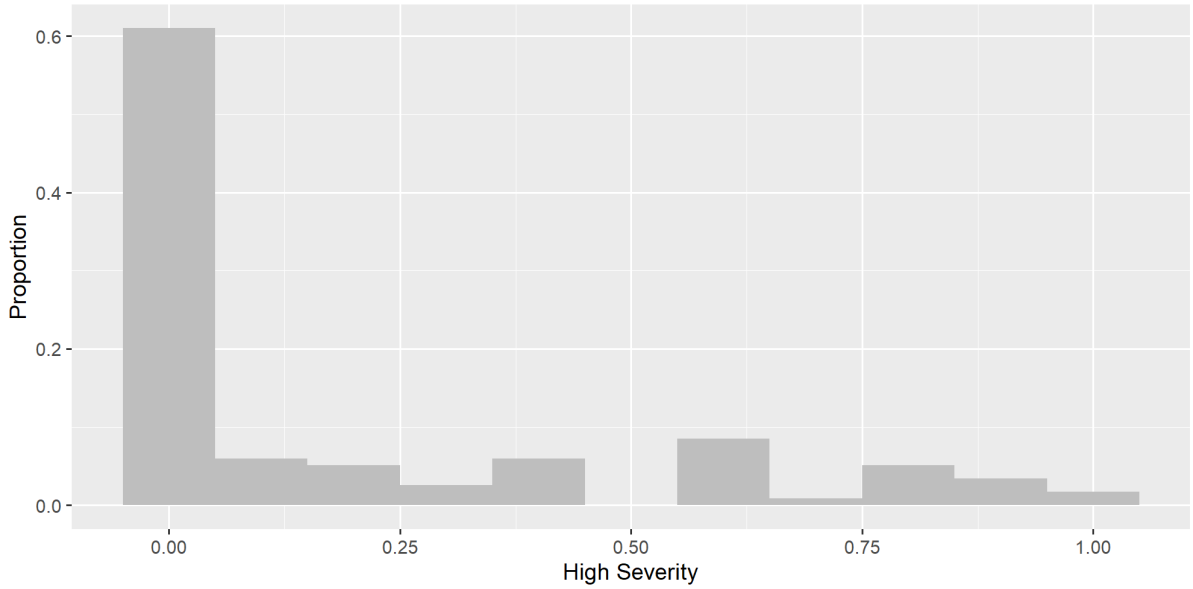

Figure D.2: Histogram of proportion of watersheds covered by high severity fire.

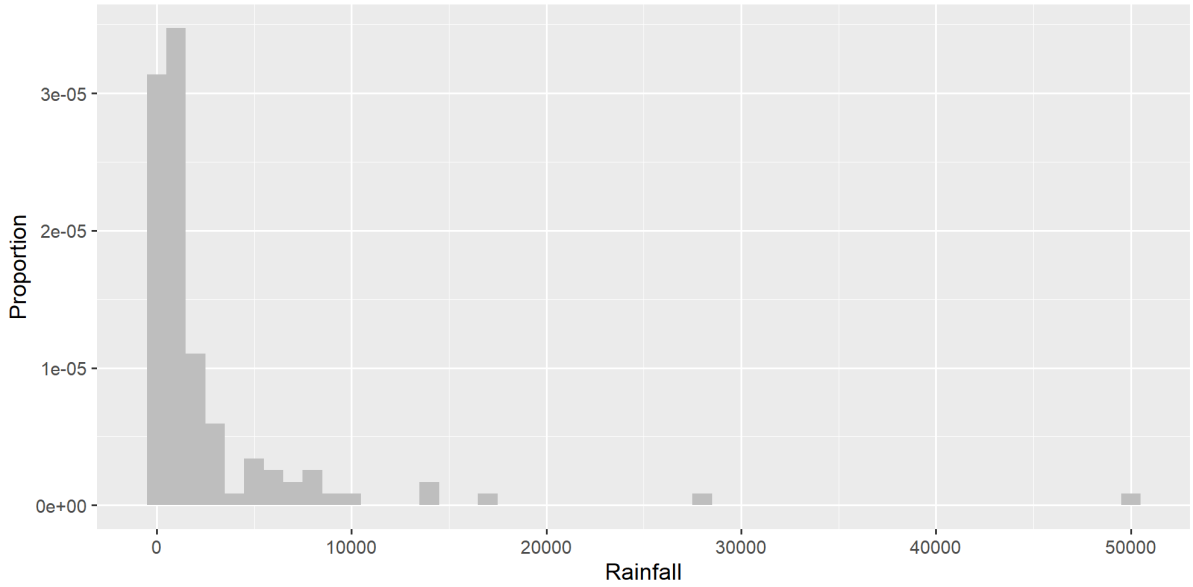

Figure D.3: Proportion of watersheds with amount of rainfall across a watershed post-fire.

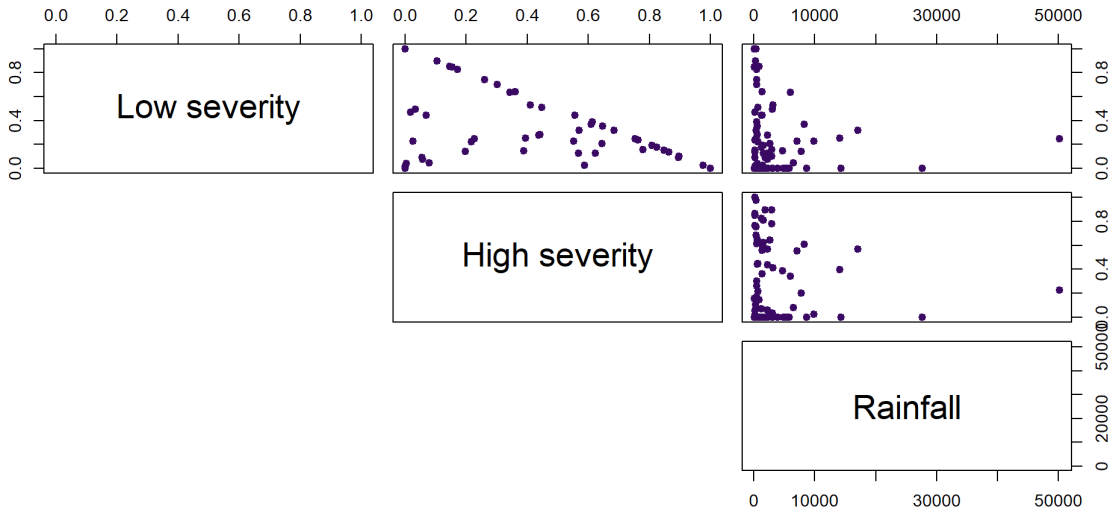

Figure D.4: Scatter matrix of proportion of watershed covered by low severity fire, high severity fire, and the amount of rainfall in a watershed post-fire using data from sites sampled in all three time periods.

JAGS code for Chapter 4

```
model {

    # Priors

    # - Occupancy

    for(j in 1:nsite){

    int.psi[j] ~  dunif(0,1)

    }

    # - eDNA availability

    int.theta ~ dunif(0,1) # Intercepts availability probability

    #detection

    int.p ~ dunif(0,1)

    # - occupancy

    betaBA1 ~ dnorm(0, 0.1)

    betaBA1burnt ~ dnorm(0, 0.1)

    betaBA2 ~ dnorm(0, 0.1)

    betaBA2burnt ~ dnorm(0, 0.1)

    betaBurnt ~ dnorm(0, 0.1)

    beta_volume~ dnorm(0, 0.1)



    # Ecological process

    for(i in 1:nvisit){

    for(j in 1:nsite){

    logit(pocc[i, j]) <-logit(int.psi[j])   + betaBA1 * I_after1[i, j]   + betaBA2 * I_after2[i, j]
+betaBurnt* I_burnt[i, j]+

    betaBA1burnt * I_after1[i, j] * I_burnt[i, j] + betaBA2burnt * I_after2[i, j] * I_burnt[i, j]

    z[i, j] ~ dbern(pocc[i, j])
```

```
    for (k in 1:nrep){
    # Occurrence in sample j
    a[i,j,k] ~ dbern(mu.a[i,j,k])
    mu.a[i,j,k] <- z[i, j] * theta[i,j,k]
    cloglog(theta[i,j,k]) <- cloglog(int.theta)+beta_volume*volume[i,j,k]
    # Observation process
    for(l in 1:npcr){
    logit(p[i, j, k, l])<- logit(int.p)
    mu.p[i, j, k, l] <-  a[i,j,k] *p[i, j,k,l]
    all_data[i, j, k, l] ~ dbern(mu.p[i, j, k,l])
    }
    }
    }
    }
    ",fill = TRUE)
sink()

zst<-matrix(data=1,nrow=3,ncol=187)
ast<-array(1, c(3,187, 2))

ast <- apply(all_data, c(1,2,3), max)   # inits for availability (a)
ast[is.na(ast)] <- 1

inits <- function() list(z = zst, a = ast)
# Parameters monitored
```

```
params1 <- c("int.psi","int.theta", "int.p",
             "betaBA1","betaBA2","betaBA1burnt",
             "betaBA2burnt","betaBurnt", "beta_volume","z")


# MCMC settings
ni <- 30000   ;   nt <- 10   ;   nb <- 10000   ;   nc <- 3
#ni <- 3000   ;   nt <- 10   ;   nb <- 1000  ;   nc <- 3
out_fire_platy_updated_s_sev_v<-  jags(fire_platy_updated_s_sev_v,  inits,  params1,
"fire_platy_updated_s_sev_v.txt", n.chains = nc, n.thin = nt, n.iter = ni,n.burnin = nb)
```

Table D.4. Confusion matrix of observed vs predicted occupancy. A threshold of 0.5 was used to convert the predictions to binary values.

	Observed		
Predicted		Present	Absent
	Present	95	3
	Absent	3	22

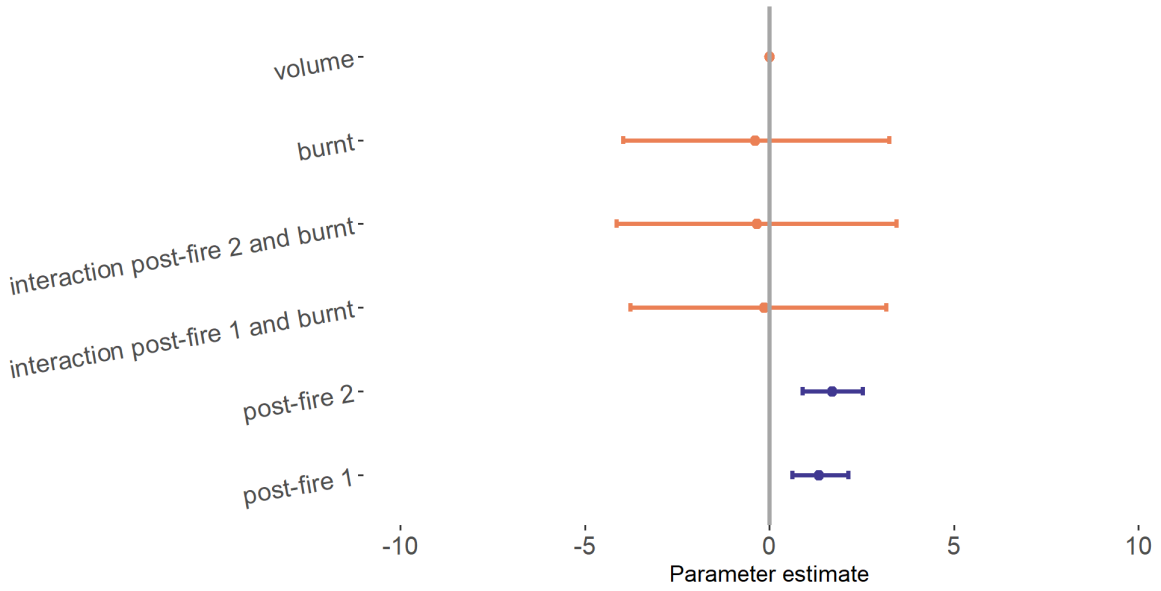

Figure D.5: β parameter estimates for the effect of a covariate on occupancy, or availability (volume only) probabilities. Points are the parameter mean estimate and lines are the 95% credible intervals. Dark purple lines do not overlap zero orange lines overlap zero. Burnt sites classified as within the burnt extent, as opposed to in the watershed.

Appendix E
Supplementary material for chapter 5

Table E.0.1: Vertebrate Primers

Forward	GATTAGATACCYCACTATGC
Reverse	TAGRACAGGCTCCTCTAG

JAGS code

```
model {
   # Priors
   for(i in 1:nspec){      # loop over species
   lpsi[i] ~ dnorm(mu.lpsi[g[i]], tau.lpsi[g[i]]) # note g-dependence now
   lp[i] ~ dnorm(mu.lp[g[i]], tau.lp[g[i]])
   beta_urban[i]~ dnorm(mu.lurban[g[i]], tau.urban[g[i]])
   beta_water[i]~ dnorm(mu.lwater[g[i]],tau.water[g[i]])
   beta_agriculture[i]~ dnorm(mu.lagriculture[g[i]], tau.agriculture[g[i]])
   beta_grass[i]~ dnorm(mu.lgrass[g[i]], tau.grass[g[i]])
   beta_forest[i]~ dnorm(mu.lforest[g[i]],tau.forest[g[i]])
   beta_runoff[i]~ dnorm(mu.lrunoff[g[i]],tau.runoff[g[i]])
   beta_z_flow[i]~ dnorm(mu.lz_flow[g[i]],tau.z_flow[g[i]])
   }




   sigma_catch~dunif(0,1)

   # convert it to a precision (1 / variance)
   tau_catch<-pow(sigma_catch, -2)

   # define the random effect; one value for each site, all share the same precision
```

```
for (i in 1:ncatch){
gamma_catch[i]~dnorm(0,tau_catch)
}

# Hyperpriors
for(g in 1:2){          # loop over groups (g)
mu.lpsi[g] <- logit(mu.psi[g])      # everything is indexed g now
mu.lp[g] <- logit(mu.p[g])
mu.psi[g] ~ dunif(0,1)
mu.p[g] ~ dunif(0,1)
tau.lpsi[g] <- pow(sd.lpsi[g], -2)
sd.lpsi[g] ~ dunif(0,5)
tau.lp[g] <- pow(sd.lp[g], -2)
sd.lp[g] ~ dunif(0,5)

mu.lurban[g] <- logit(mu.urban[g])
mu.urban[g] ~ dunif(0,1)
sd.urban[g] ~ dunif(0,5)
tau.urban[g] <- pow(sd.urban[g], -2)

mu.lwater[g] <- logit(mu.water[g])
mu.water[g] ~ dunif(0,1)
sd.water[g] ~ dunif(0,5)
tau.water[g] <- pow(sd.water[g], -2)

mu.lz_flow[g] <- logit(mu.z_flow[g])
mu.z_flow[g] ~ dunif(0,1)
sd.z_flow[g] ~ dunif(0,5)
tau.z_flow[g] <- pow(sd.z_flow[g], -2)

mu.lagriculture[g] <- logit(mu.agriculture[g])
mu.agriculture[g] ~ dunif(0,1)
```

```
sd.agriculture[g] ~ dunif(0,5)
tau.agriculture[g] <- pow(sd.agriculture[g], -2)

mu.lgrass[g] <- logit(mu.grass[g])
mu.grass[g] ~ dunif(0,1)
sd.grass[g] ~ dunif(0,5)
tau.grass[g] <- pow(sd.grass[g], -2)

mu.lforest[g] <- logit(mu.forest[g])
mu.forest[g] ~ dunif(0,1)
sd.forest[g] ~ dunif(0,5)
tau.forest[g] <- pow(sd.forest[g], -2)

mu.lrunoff[g] <- logit(mu.runoff[g])
mu.runoff[g] ~ dunif(0,1)
sd.runoff[g] ~ dunif(0,5)
tau.runoff[g] <- pow(sd.runoff[g], -2)
}


# Ecological model for true occurrence (process model)
for(i in 1:nspec){
for (j in 1:nsite) {
logit(psi[i,j]) <- lpsi[i] +
beta_urban[i]  * urban[j] +
beta_water[i]  * water[j] +
beta_agriculture[i]  * agriculture[j] +
beta_grass[i] *grass[j]+
beta_forest[i] *forest[j]+
beta_runoff[i] *runoff[j]+
beta_z_flow[i] *z_flow[j]+
gamma_catch[catchment_fact[i]]
```

```
    z[i,j] ~ dbern(psi[i,j])
    }
    }

    # Observation model for replicated detection/nondetection observations
    for(i in 1:nspec){
    for (j in 1:nsite){
    for(k in 1:nrep){
    logit(p[i,j,k]) <- lp[i]
    mu.px[i,j,k] <- z[i,j] * p[i,j,k]
    ysum[i,j,k] ~ dbern(mu.px[i,j,k])
    }
    }
    }
    # Derived quantities
    for(i in 1:nspec){          # Loop over species
    Nocc.fs[i] <- sum(z[i,]) # Number of occupied sites among the 267
    }

    for (j in 1:nsite) {           # Loop over sites
    Nsite[j] <- sum(z[,j])   # Number of occurring species at each site
    }
    }
    ",fill = TRUE)
sink()

# Initial values
zst <- apply(arr_presence, c(1,2), max)
zst[is.na(zst)] <- 1
inits <- function() list(z = zst)

# Parameters monitored
```

```
params <- c("lpsi","mu.psi", "mu.lpsi", "sd.lpsi", "mu.p", "mu.lp", "sd.lp",
        "mu.lurban","mu.lwater",
        "mu.lagriculture","mu.lgrass",
        "mu.lforest","mu.lrunoff","mu.lz_flow",
        "beta_urban","beta_water","beta_agriculture","beta_grass",
        "beta_forest","beta_runoff","beta_z_flow","psi")


ni <- 300000   ;   nt <- 100   ;   nb <- 100000   ;   nc <- 3
out_fish_data_vic_melb_n_site2_10 <- jags(fish_data_vic_melb, inits, params,
"fish_data_vic_melb.txt", n.chains = nc, n.thin = nt, n.iter = ni, n.burnin = nb, parallel = TRUE)
```

JAGS models code for species richness plots (Figures 5.4-5.5).

```
str( fish_vic_melb_introduced_spatial_sp <- list(ysum = arr_presence3, nsite =dim(arr_presence3)[2],
                                nrep = dim(arr_presence3)[3], nspec = dim(arr_presence3)[1],
                                urban=urban,
                                water=water,
                                agriculture=agriculture,
                                grass=grass,
                                forest=forest,
                                runoff=runoff,
                                z_flow=z_flow,
                                M= dim(arr_presence3)[1],
                                #native=native2,
                                catchment_fact=catchment_fact,
                                ncatch=length(unique(catchment_fact))))

# Specify model in BUGS language
sink("fish_vic_melb_introduced_spatial_sp.txt")
cat("
```

```
model {

omega~dunif(0,1)
# Priors for species-specific effects in occupancy and detection
for(k in 1:nspec){
lpsi[k] ~ dnorm(mu.lpsi, tau.lpsi)    # Hyperparams describe community
lp[k] ~ dnorm(mu.lp, tau.lp)
beta_urban[k] ~ dnorm(mu.urban, tau.urban)
beta_water[k] ~ dnorm(mu.water, tau.water)
beta_agriculture[k] ~ dnorm(mu.agriculture, tau.agriculture)
beta_grass[k] ~ dnorm(mu.grass, tau.grass)
beta_forest[k] ~ dnorm(mu.forest, tau.forest)
beta_runoff[k] ~ dnorm(mu.runoff, tau.runoff)
beta_z_flow[k] ~ dnorm(mu.z_flow, tau.z_flow)
#beta_native[k] ~ dnorm(mu.native, tau.native)
}


sigma_catch~dunif(0,1)

# convert it to a precision (1 / variance)


tau_catch<-pow(sigma_catch, -2)

for (i in 1:ncatch){
gamma_catch[i]~dnorm(0,tau_catch)
}

# Hyperpriors
# For the model of occupancy
mu.lpsi ~ dnorm(0,0.1)
```

```
tau.lpsi <- pow(sd.lpsi, -2)
sd.lpsi ~ dunif(0,5)

# For the model of detection
mu.lp ~ dnorm(0,0.1)
tau.lp <- pow(sd.lp, -2)
sd.lp ~ dunif(0, 5)


mu.urban ~ dnorm(0,0.1)
sd.urban ~ dunif(0,5)
tau.urban <- pow(sd.urban, -2) ####up to here


mu.water ~ dnorm(0,0.1)
sd.water ~ dunif(0,5)
tau.water <- pow(sd.water, -2)


mu.z_flow ~ dnorm(0,0.1)
sd.z_flow ~ dunif(0,5)
tau.z_flow <- pow(sd.z_flow, -2)


mu.agriculture ~ dnorm(0,0.1)
sd.agriculture ~ dunif(0,5)
tau.agriculture <- pow(sd.agriculture, -2)


mu.grass ~ dnorm(0,0.1)
sd.grass ~ dunif(0,5)
tau.grass <- pow(sd.grass, -2)
```

```
mu.forest ~ dnorm(0,0.1)
sd.forest ~ dunif(0,5)
tau.forest <- pow(sd.forest, -2)


mu.runoff ~ dnorm(0,0.1)
sd.runoff ~ dunif(0,5)
tau.runoff <- pow(sd.runoff, -2)

# mu.native~ dnorm(0,0.1)
# sd.native ~ dunif(0,5)
# tau.native <- pow(sd.native, -2)
#

#superpopulation process:Ntotal species sampled out of M avaiable
for(k in 1:M){
w[k]~dbern(omega)
}


# Ecological model for true occurrence (process model)
for(i in 1:nspec){
for (j in 1:nsite) {
logit(psi[i,j]) <- lpsi[i] +
beta_urban[i]  * urban[j] +
beta_water[i]  * water[j] +
beta_agriculture[i]  * agriculture[j] +
beta_grass[i] *grass[j]+
beta_forest[i] *forest[j]+
beta_runoff[i] *runoff[j]+
```

```
    beta_z_flow[i] *z_flow[j]+
    #beta_native[i] *native[j]+
    gamma_catch[catchment_fact[i]]
    z[i,j] ~ dbern(psi[i,j])
    }
    }


    # Observation model for replicated detection/nondetection observations
    for(i in 1:nspec){
    for (j in 1:nsite){
    for(k in 1:nrep){
    logit(p[i,j,k]) <- lp[i]
    mu.px[i,j,k] <- z[i,j] * p[i,j,k]
    ysum[i,j,k] ~ dbern(mu.px[i,j,k])
    }
    }
    }
    # Derived quantities
    for(i in 1:nspec){          # Loop over species
    Nocc.fs[i] <- sum(z[i,]) # Number of occupied sites among the 267
    }

    for (j in 1:nsite) {           # Loop over sites
    Nsite[j] <- sum(z[,j])   # Number of occurring species at each site
    }
    }
    ",fill = TRUE)
sink()


# Initial values
```

```
zst <- apply(arr_presence3, c(1,2), max)

zst[is.na(zst)] <- 1

inits <- function() list(z = zst)


# Parameters monitored

params <- c("mu.psi", "mu.lpsi", "sd.lpsi", "mu.p", "mu.lp", "sd.lp",

        "beta_urban","beta_water","beta_agriculture","beta_grass",

        "beta_forest","beta_runoff","beta_z_flow", "z","w","lpsi")


ni <- 30000   ;   nt <- 10   ;   nb <- 10000   ;   nc <- 3

out_fish_vic_melb_introduced_spatial_sp <- jags.basic(fish_vic_melb_introduced_spatial_sp, inits, params, "fish_vic_melb_introduced_spatial_sp.txt", n.chains = nc, n.thin = nt, n.iter = ni, n.burnin = nb, parallel = TRUE)
```

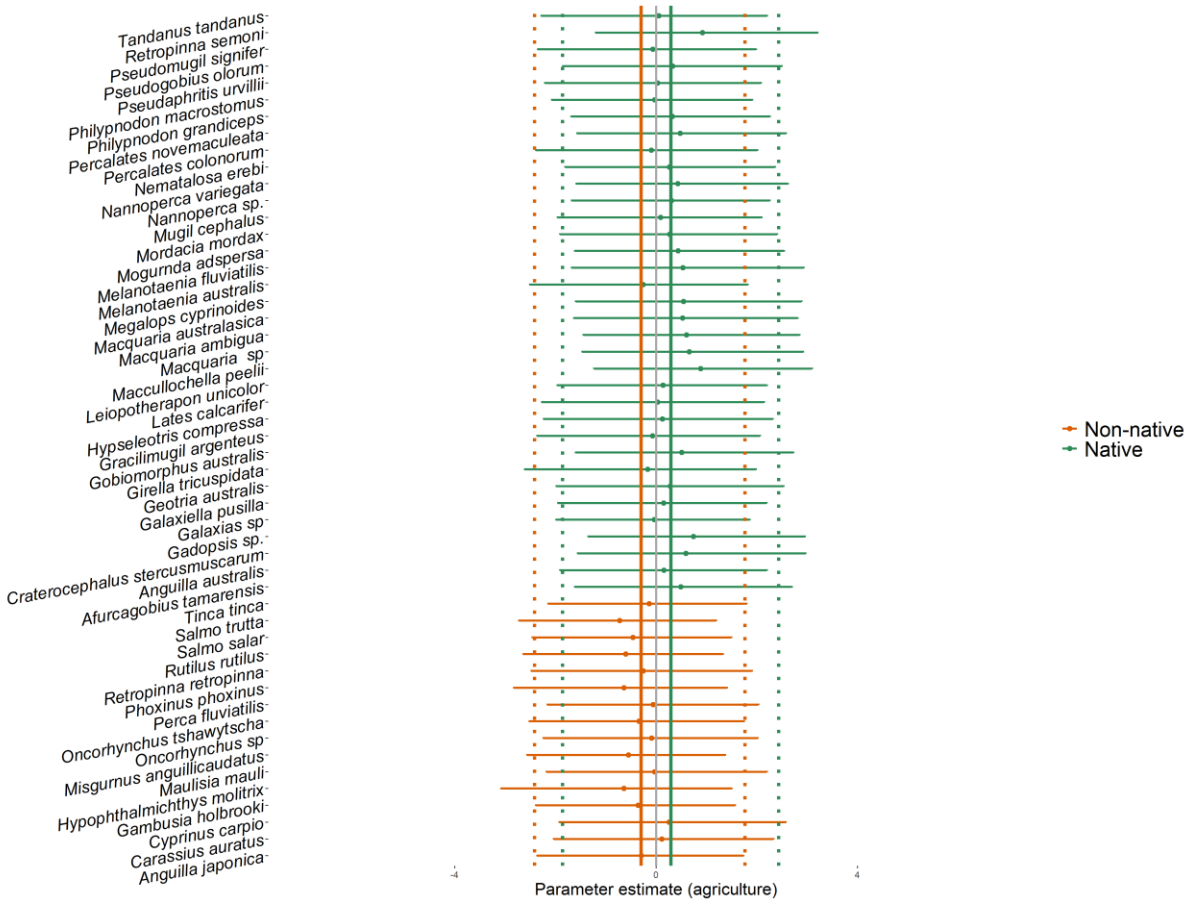

Figure E.1: β coefficients describing effects of an environmental covariate (the proportion of a contracted catchment covered by agriculture) on species occupancy. Points are the mean and lines are the 95% credible intervals. Green lines are native species and orange lines are non-native (introduced) species. Solid vertical lines are the community mean for each group and dotted vertical lines are the 95% credible interval for each group.

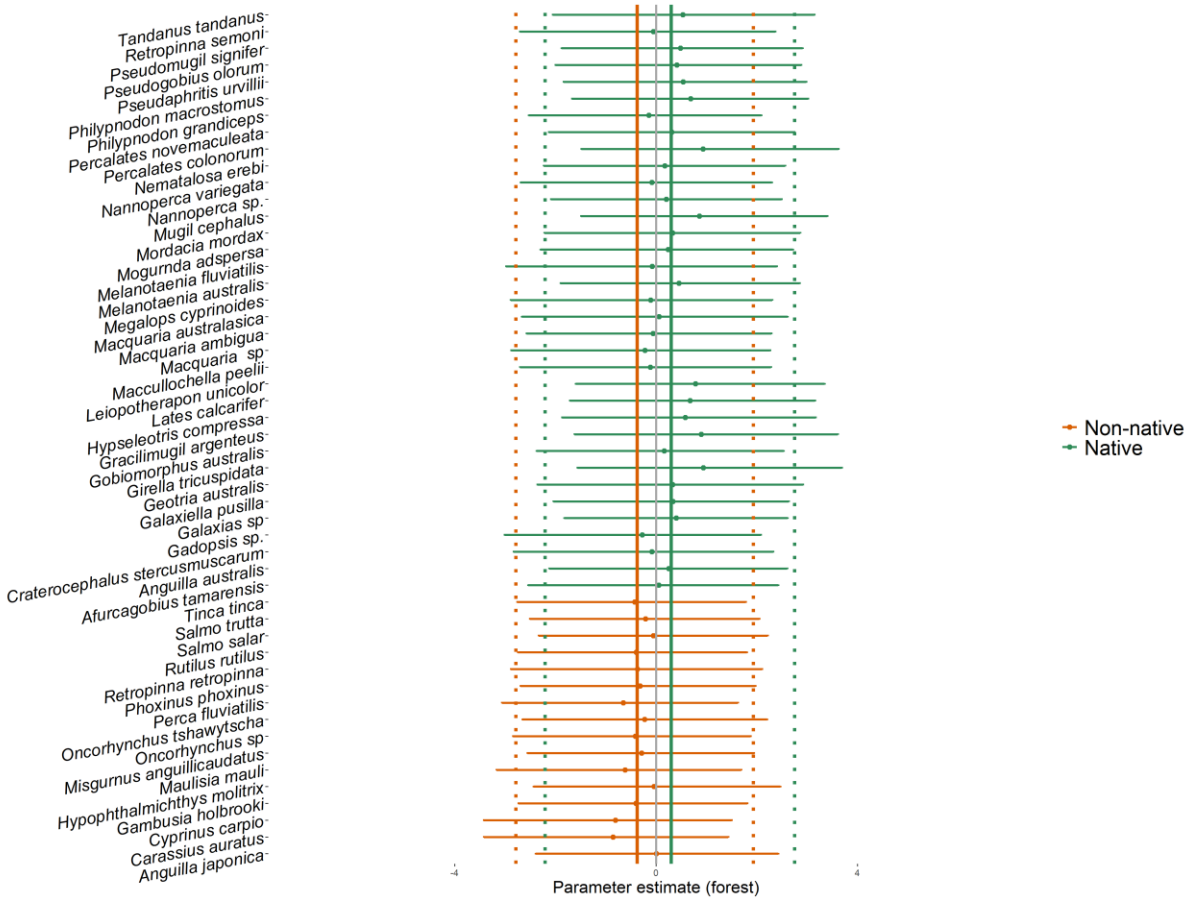

Figure E.2 β coefficients describing effects of an environmental covariate (the proportion of a contracted catchment covered by forest) on species occupancy. Points are the mean and lines are the 95% credible intervals. Green lines are native species and orange lines are non-native (introduced) species. Solid vertical lines are the community mean for each group and dotted vertical lines are the 95% credible interval for each group.

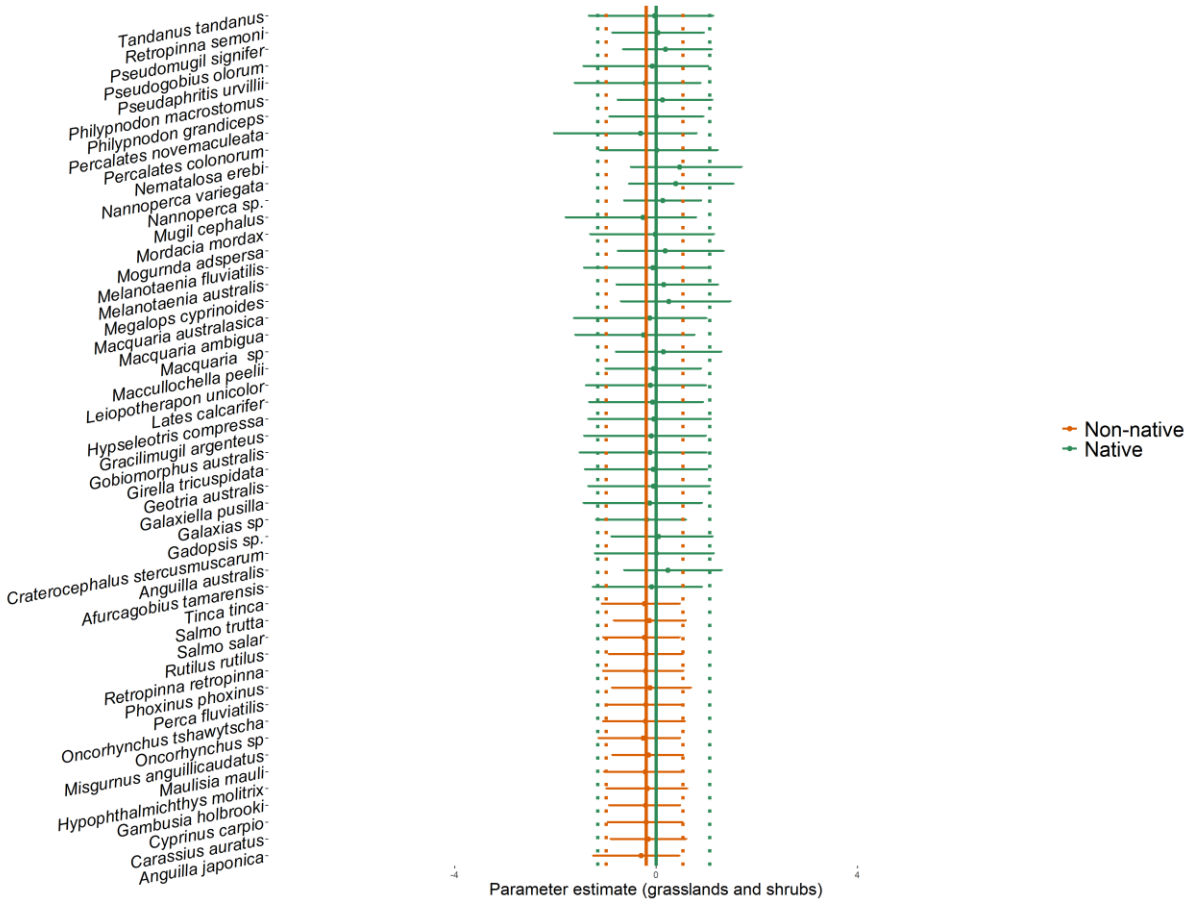

Figure E.3: β coefficients describing effects of an environmental covariate (the proportion of a contracted catchment covered by grasslands and shrubs) on species occupancy. Points are the mean and lines are the 95% credible intervals. Green lines are native species and orange lines are non-native (introduced) species. Solid vertical lines are the community mean for each group and dotted vertical lines are the 95% credible interval for each group.

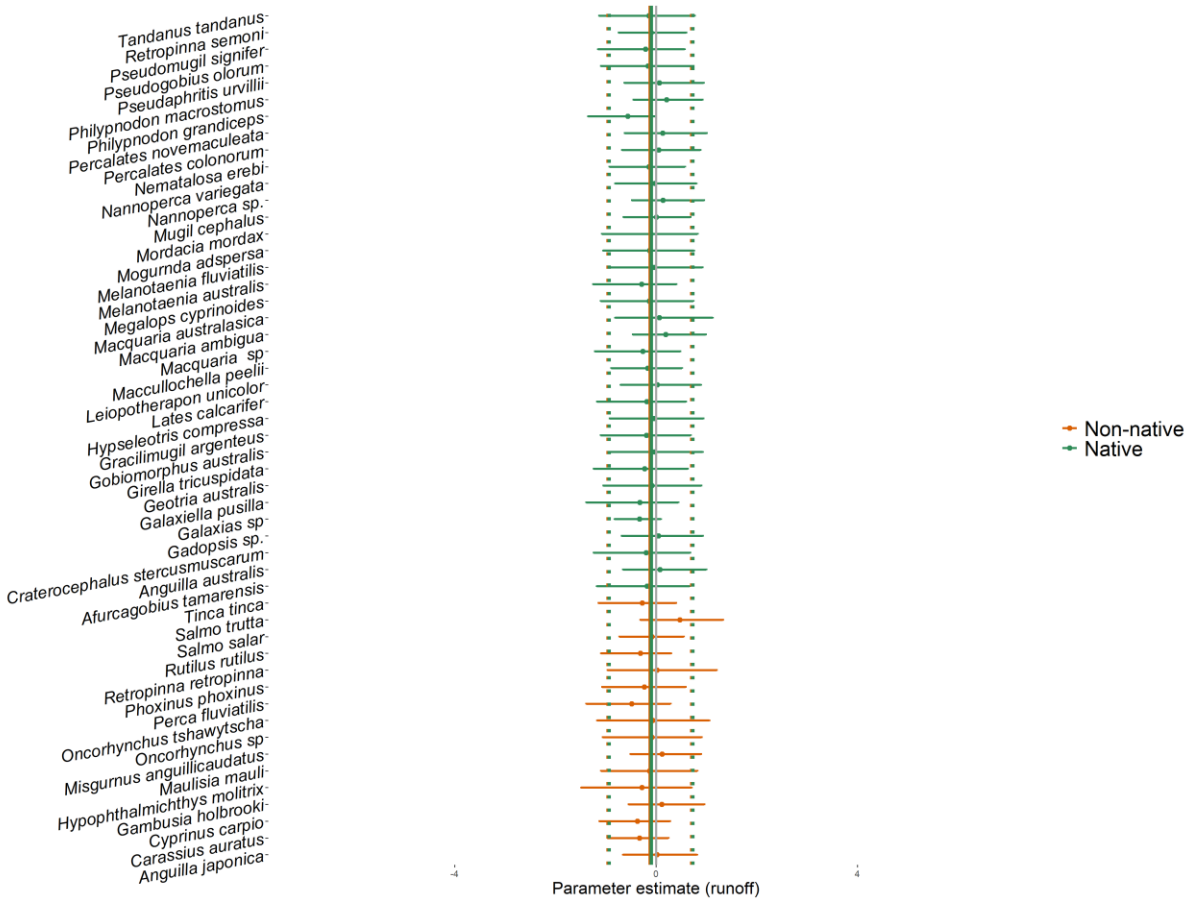

Figure E.4: β coefficients describing effects of an environmental covariate (the amount of runoff in a contracted catchment) on species occupancy. Points are the mean and lines are the 95% credible intervals. Green lines are native species and orange lines are non-native (introduced) species. Solid vertical lines are the community mean for each group and dotted vertical lines are the 95% credible interval for each group.

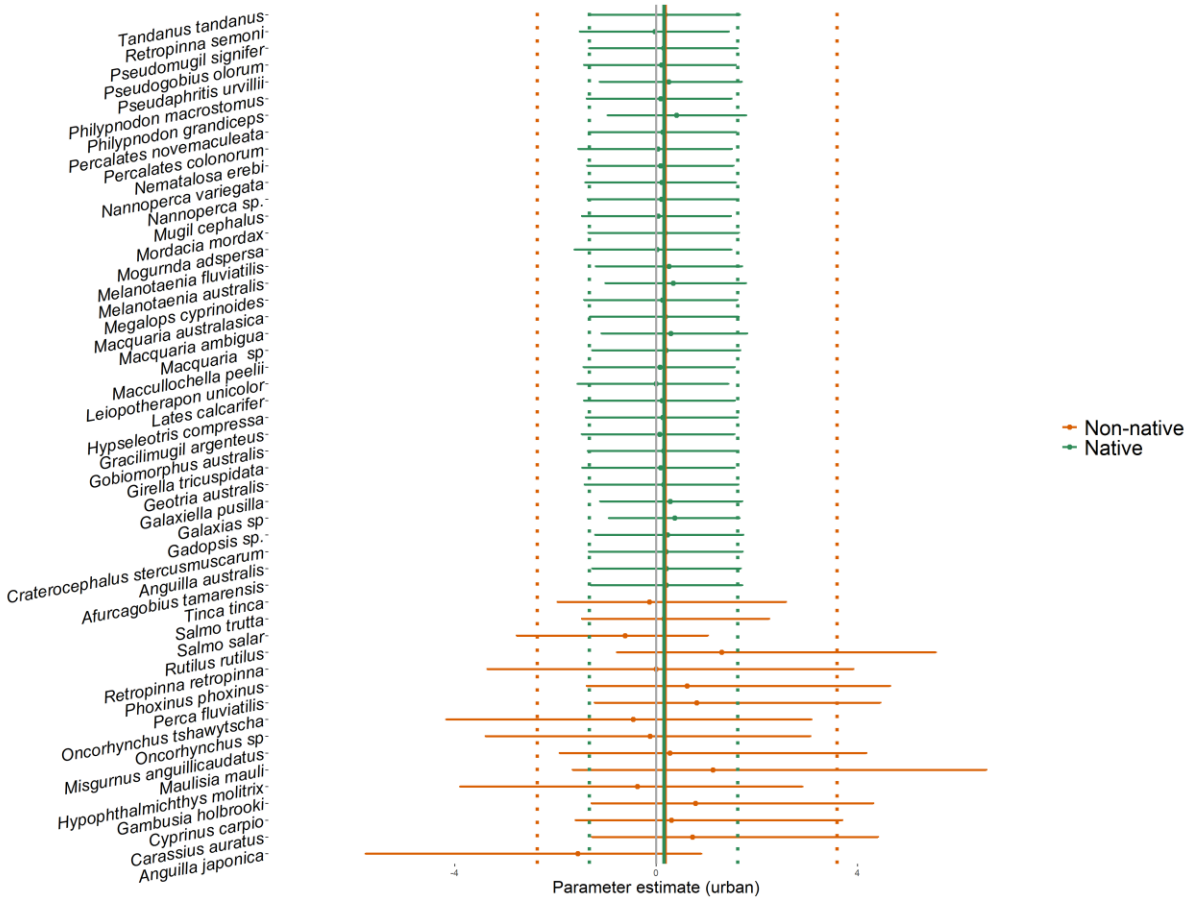

Figure E.5: β coefficients describing effects of an environmental covariate (the proportion of a contracted catchment covered by urban land uses) on species occupancy. Points are the mean and lines are the 95% credible intervals. Green lines are native species and orange lines are non-native (introduced) species. Solid vertical lines are the community mean for each group and dotted vertical lines are the 95% credible interval for each group.

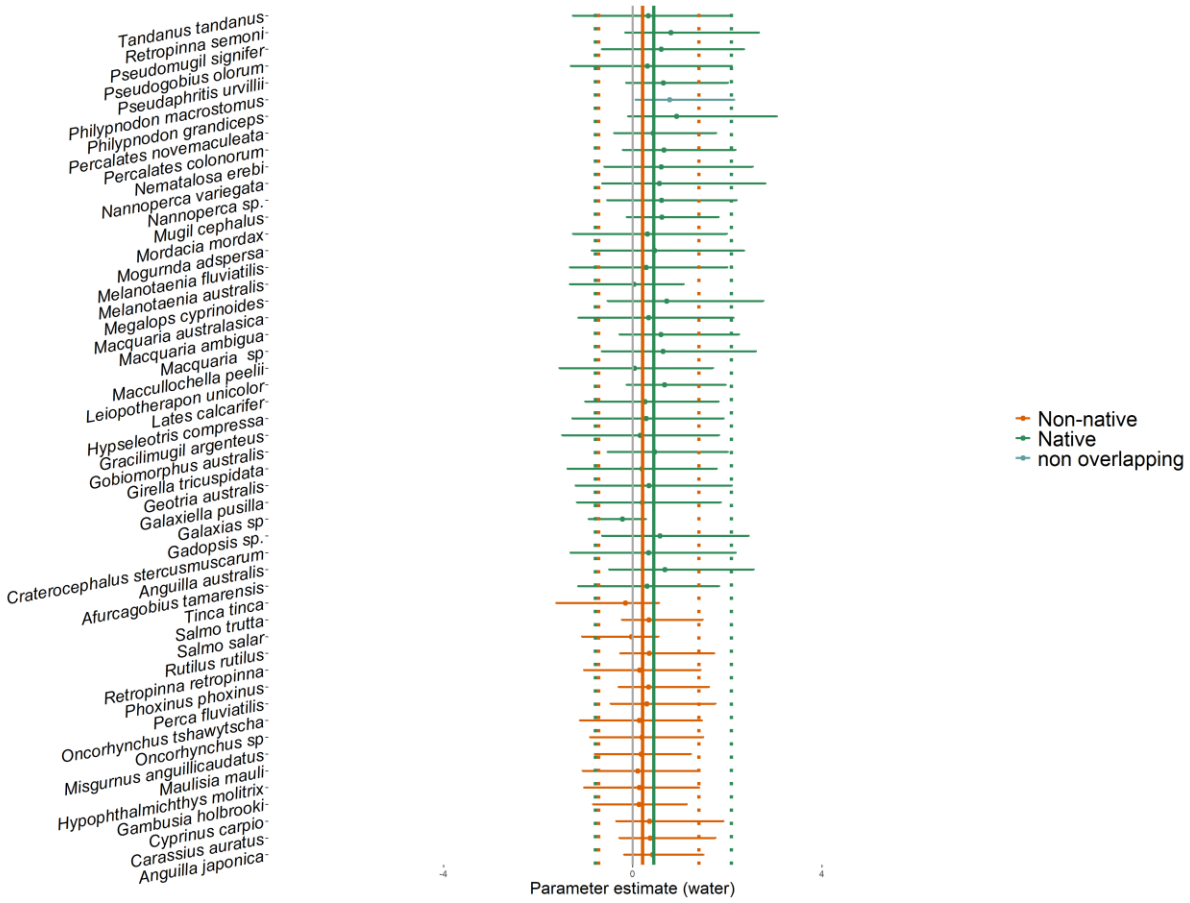

Figure E.6 β coefficients describing effects of an environmental covariate (the proportion of a contracted catchment covered by water) on species occupancy. Points are the mean and lines are the 95% credible intervals. Green lines are native species and orange lines are non-native (introduced) species. Blue lines represent species where the β estimate does not overlap zero. Solid vertical lines are the community mean for each group and dotted vertical lines are the 95% credible interval for each group.

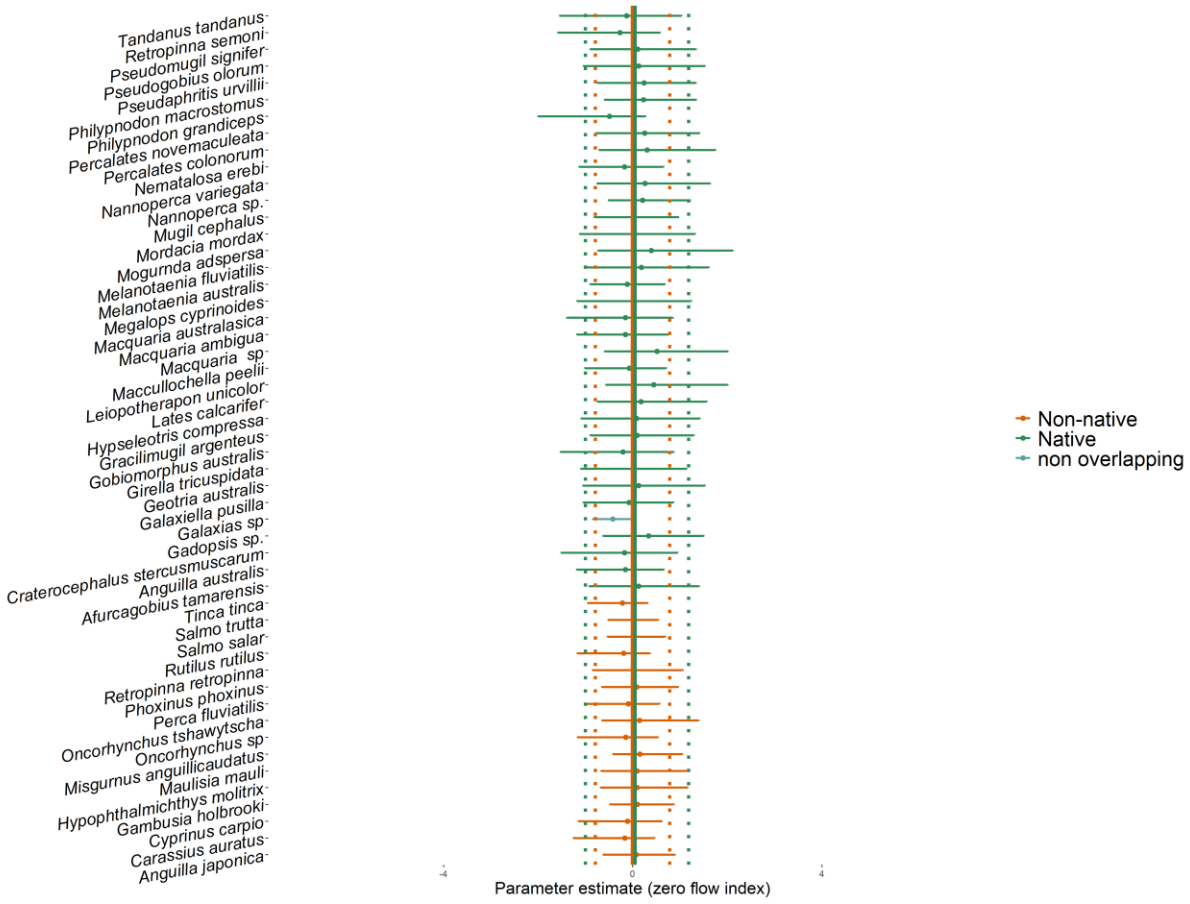

Figure E.7: β coefficients describing effects of an environmental covariate (zero-flow index) on species occupancy. Points are the mean and lines are the 95% credible intervals. Green lines are native species and orange lines are non-native (introduced) species. Blue lines represent species where the β estimate does not overlap zero. Solid vertical lines are the community mean for each group and dotted vertical lines are the 95% credible interval for each group.

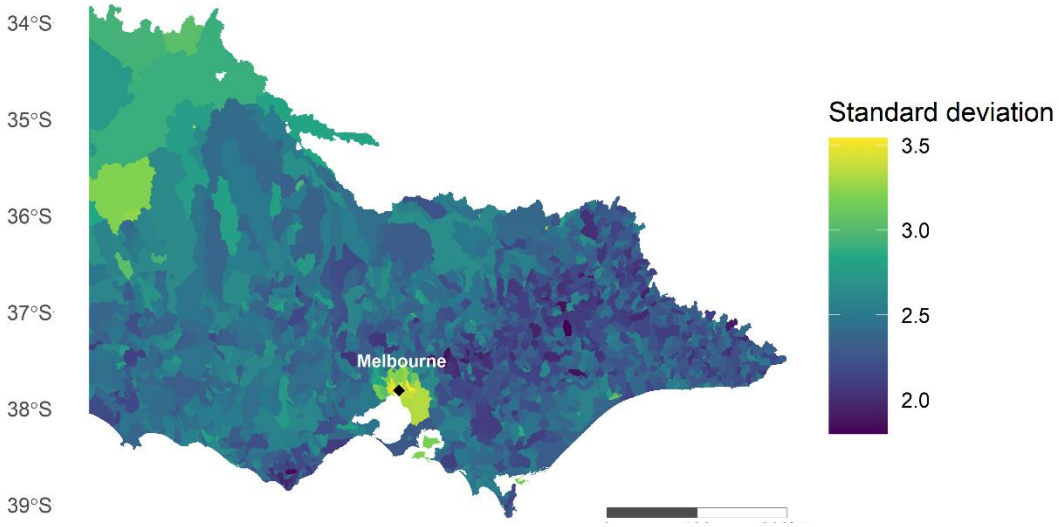

Figure E.8: Predicted introduced fish species richness at the contracted catchment scale based on the community occupancy model presented above. The posterior standard deviation estimate is used.

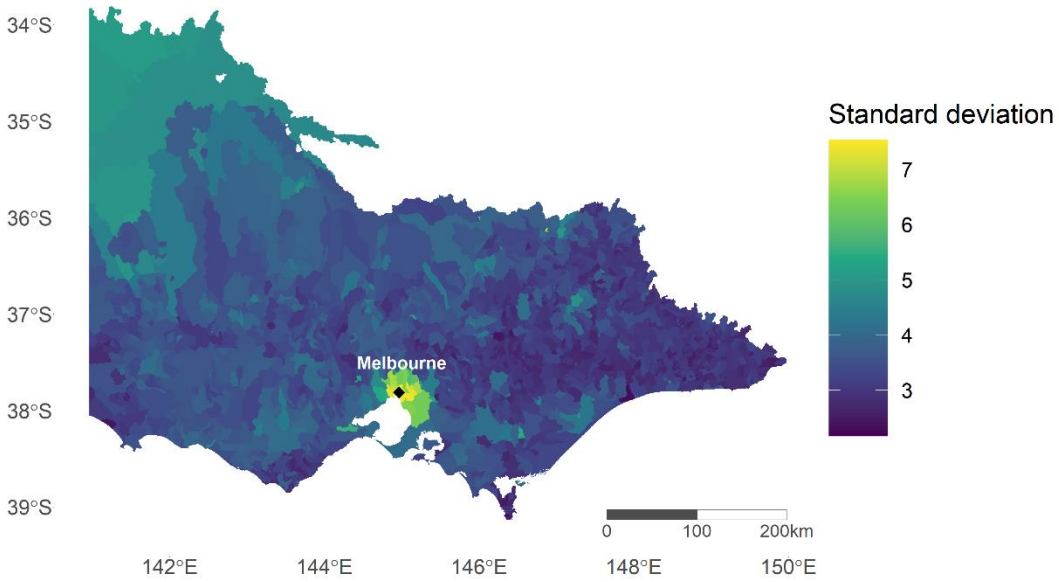

Figure E.9: Predicted native fish species richness at the contracted catchment scale based on the community occupancy model presented above. The posterior standard deviation estimate is used.

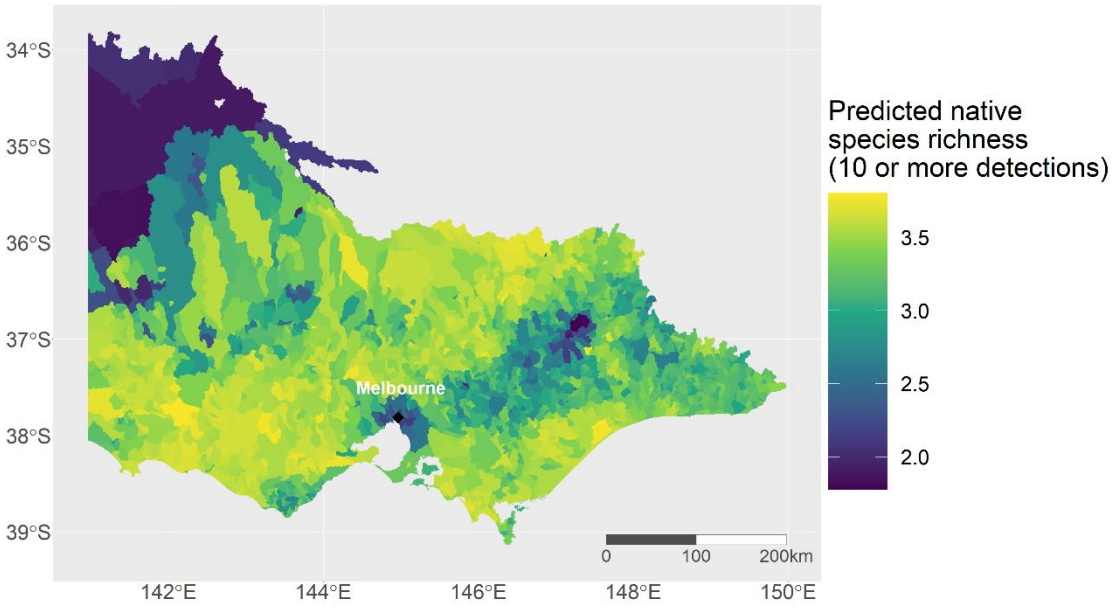

Figure E.10: Predicted native fish species richness at the contracted catchment scale based on the community occupancy model presented in the main text. The posterior mean estimate of species richness is used. Native species with 10 or more detections used for prediction.

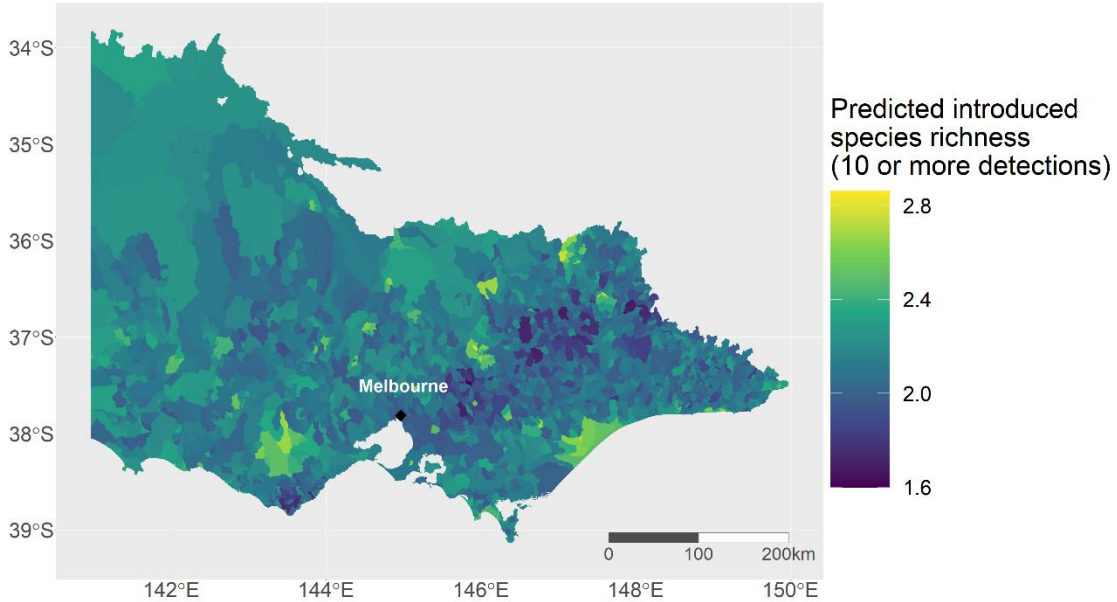

Figure E.11: Predicted native fish species richness at the contracted catchment scale based on the community occupancy model presented in the main text. The posterior mean estimate of species richness is used. Introduced species with 10 or more detections used for prediction.

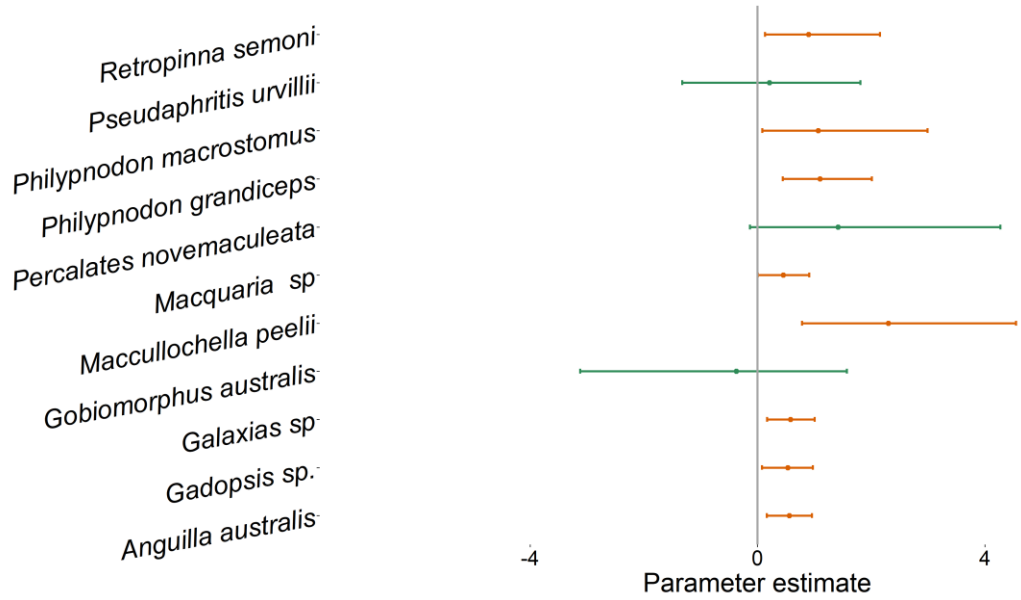

Figure E.12: β coefficients describing effect of introduced taxa richness on native taxa occupancy. Points are the mean and lines are the 95% credible intervals. Orange lines do not overlap zero and green lines overlap zero.

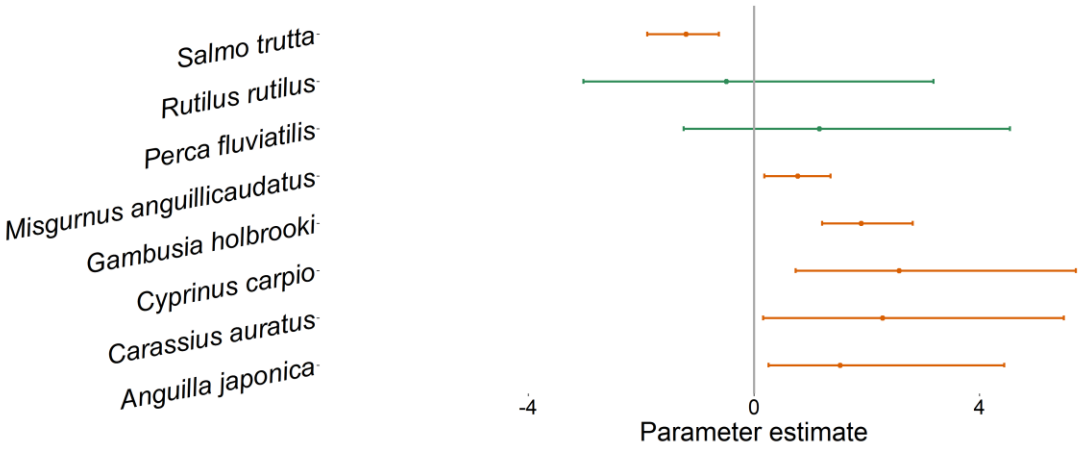

Figure E.13: β coefficients describing effect of native taxa richness on introduced taxa occupancy. Points are the mean and lines are the 95% credible intervals. Orange lines do not overlap zero and green lines overlap zero.