DOCTOR OF PHILOSOPHY

Assessing and managing risks from climate change in drinking water supply sources - safeguarding raw water quality through improving catchment resilience

Vorstius, Carolin

Award date:
2023

Assessing and managing risks from climate change in drinking water supply sources - safeguarding raw water quality through improving catchment resilience

Carolin Vorstius

Submitted for the degree of Doctor of Philosophy, December 2022

University of Dundee

Contents

List of Figures

List of Tables

List of Abbreviations

AAT	Annual accumulated temperature
AIC	Akaike Information Criterion
BFI	Baseflow index
BIC	Bayesian Information Criterion
BNG	British National Grid
CA	Cluster analysis
DA	Discriminant analysis
DOC	Dissolved organic carbon
DP GBR	Diffuse Pollution General Binding Rules
DWQR	Drinking Water Quality Regulator
EbA	Ecosystem-based adaptation
FA	Factor analysis
FIO	Faecal indicator organisms
GHG	Greenhouse gas
GIS	Geographic Information System
HOST	Hydrology of Soil Types
HSPF	Hydrological Simulation Program-FORTRAN
ICM	Integrated Catchment Management
INCA	Integrated Catchment model
IPCC	Intergovernmental Panel on Climate Change
IWA	International Water Association
JHI	James Hutton Institute
LC	Land class
MTPCT	Multi-target predictive clustering tree
NbS	Nature-based solution
NVZ	Nitrate Vulnerable Zone
PAM	Partitioning around medoids
PCA/PC	Principal component analysis/Principal component
PET	Potential evapotranspiration
PPE	Perturbed physics ensemble
RDA	Redundancy analysis
RQ	Research question
RSME	Root mean square error
SCaMP	Sustainable Catchment Management Programme
SCIMAP	Sensitive Catchment Integrated Mapping Analysis Platform
SDG	Sustainable Development Goal
SEPA	Scottish Environment Protection Agency
SER	Summer effective rainfall
SPR	Standard percentage runoff
SWAT	Soil and Water Assessment Tool
TOC	Total organic carbon
VIF	Variance inflation factor
WEWS	Water Environment and Water Services (Scotland) Act
WFD	Water Framework Directive
WHO	World Health Organisation

Acknowledgements

“Which is more important,” asked Big Panda, “the journey or the destination?” - “The company.” said Tiny Dragon.

James Norbury

This PhD was a journey which enabled me to grow professionally as well as personally. Without the many people accompanying and supporting me for all or part of the way, it would not have been so special or enjoyable, or even possible.

First and foremost, I am very grateful to my supervisors, Prof. John Rowan and Dr. Iain Brown, for their continued and unwavering support, feedback, and encouragement. You made a superb supervisory team – thank you!

Thank you to Scottish Water for sharing their data, thus making this project possible. More concretely, I especially thank my industry advisors at Scottish Water, Dr. Zoë Frogbrook and Dr. Fraser Leith, for additionally sharing their knowledge and expertise, and for their support over the years. Your input was vital to keep the research relevant.

Special thanks also go to Dr. Javier Palarea-Albaladejo and Prof. Steve Hubbard for their advice on statistics - and their patience in explaining it to me!

Scotland has a fantastic network of people in the water sector, and I have been privileged to meet many who were willing to discuss my research with me. Thanks very much to all of you!

I am grateful to the Scottish government for funding this project through the Hydro Nation Scholarship Programme. Thank you to the Hydro Nation Scholars team, for finding so many ways to support us. It's been a pleasure to be part of the family!

To my fellow scholars – it was great to share this journey with you all! This also goes for my fellow PhD students in Geography at the University of Dundee. It was brilliant being part of this department and thank you to everyone in it for their input and inspiration!

My family had to be part of this journey – thanks for keeping going with me! My wonderful husband Jan bravely shared all the uphill, downhill, and in between parts, and we added some baggage along the way in the form of our three beautiful children. But you three make everything lighter and brighter, and you are my pride and joy!

To my parents, thank you for always supporting me! I hope I can be to my children what you are to me.

Declaration

I declare that the thesis here submitted is original except for the source materials explicitly acknowledged, and that I am the author of this thesis and the work recorded here was done by myself. This thesis, or parts of this thesis, have not been previously submitted for the same degree or for a different degree.

Dundee, 16th December 2022

Abstract

Increasing temperatures and changes in precipitation patterns and amounts are expected to increase pressures on water resources and result in decreasing water quality and quantity. Water utilities face the challenge to anticipate and plan for deteriorations in the water quality of untreated abstraction sources (so-called raw water quality) and to manage consequent risks, a challenge exacerbated through deep uncertainty associated with climate change. This research proposes a framework how to integrate effects of climate change on raw water quality into risk assessment and to build capacity to anticipate and respond to change.

A national assessment of public water supply catchments in Scotland was used as a demonstration case for applying the framework. The analysis was conducted with data from 154 drinking water supply catchments from a period of 2011-2016. Empirical relationships between catchment characteristics and water quality were examined to understand overarching drivers, pressures, and underlying catchment sensitivities leading to impacts on raw water quality. Risk screenings were developed for two water quality indicators, colour and *E. coli*, and UKCP18-based climate and land use projections were used to identify and map catchments and areas with potential increases in risk.

The screenings identified crucial controls on water quality, high risk areas, areas of uncertainty, and allowed first suggestions for possible response options based on catchment vulnerabilities. They also provided a basis for a strategic review of and planning for the complete supply system, by providing a starting point for transferring results and insights from individual sites, and by allowing for a first appraisal of long-term sustainability of supply. By focusing on catchment integrity and resilience as a crucial part of climate change mitigation and adaptive management, this research emphasises an ecosystem-based approach as a frame for water service providers to achieve multiple objectives.

1. Introduction

"Access to safe drinking-water is essential to health, a basic human right and a component of effective policy for health protection." (WHO, 2011)

Safe and available supplies of drinking water is one of the most important 'provisioning services' from the global water cycle. It is fundamental to health and survival, but competes with other crucial uses, such as use of water to grow food, for transport, recreation and to support functioning ecosystems. Water as an integral part of the environment is in constant exchange with the land and atmosphere surrounding it, and water sources are subject to multiple pressures that impact their quality. Natural variations in water quality arise from geochemical mineralogical composition of rocks and sediments, from soil organic particulate, from biogeochemical, or microbiological processes. Chemical and biological contaminants harmful to humans can arise naturally, but humans increase contamination by concentrating natural pollutants and introducing synthetic materials such as plastics, radionuclides, pharmaceuticals, or many types of pathogens.

Ensuring that drinking water supplies are free from microbial or chemical contaminants, or have concentrations below those likely to have deleterious effects on human health, is a basic tenet of the UN Sustainable Development Goal (SDG) 6.1, access to clean water for all. Regulatory standards for drinking water set an accepted risk limit and contaminant levels above these limits are considered unsafe to drink, leading to a need for water treatment prior to human consumption. Water that is free from faecal and priority chemical contamination is considered safely managed by the World Health Organisation (WHO), if it is available when needed and from a source that is located on the premises. In 2020, 74% of the global population used safely managed drinking water (WHO, 2017a).

To ensure water safety in terms of quality, risk-based approaches are promoted by the WHO, proposing a framework for risk assessment and management through Water Safety Plans (WHO, 2017b). Water Safety Plans assess the risks to water quality from catchment to user and rank them to identify and prioritise control measures to reduce them. This relies on a thorough assessment of individual supply systems and aims at anticipating impacts to prevent them from occurring.

Climate change is expected to affect water quality directly and indirectly. Rising temperatures and changing patterns in amount and seasonality of precipitation can for example modify soil moisture, runoff generation, and consequently stream flows, further altering soil erosion, sediment transport, and associated carbon fluxes (dissolved and particulate). Ecosystem productivity and ecosystem resilience could decline, with indirect negative impacts on pollutant loads (Delpla et al., 2009). Indirect effects arising from intensified land use practices may increase pressures on natural hydrological and biogeochemical processes, leading to changes in water quality (Brown et al., 2015; Whitehead, Wilby et al., 2009). There is also the prospect of new emergent risks from either biological or chemical sources (Geissen et al., 2015). It is expected that these drivers and pressures will further challenge existing treatment infrastructure and lead to an increase in effort and costs for water treatment (Ritson et al., 2014).

While water utilities can already observe trends in their water quality, and anticipate impacts from climate change, there remains deep uncertainty about the nature and magnitude of the changes (Vairavamoorthy, 2021). Climate change projections are inherently uncertain (IPCC, 2021). Predicting water quality outcomes then requires the downscaling of regional predictions to the scale of the catchment providing drinking water sources, an estimate of the effects of changes on the hydrology of the catchment, and consequences to pollutant transport and concentrations. Making predictions and understanding uncertainty around these as a basis for risk management and decision-making at the individual supply system therefore requires considerable effort (such as monitoring/data collection, data analysis, modelling, scenario building, etc.) to understand the catchment system in terms of current behaviours and risks under climate change projections and future land management practices.

Climate change brings the traditional risk assessment approach in the drinking water context that relies on identification of hazards, an estimation of their consequences, and their likelihoods, to its limits, as especially an estimation of likelihood is hardly achievable under deep uncertainty. The challenge to reliably predict a likely water quality range also means reliance on traditional end-of-pipe water treatment becomes more difficult. Planning and investment for treatment works have long lead times and relying on predictions could lead to being locked into an increasingly costly, and ultimately impossible, upgrading of treatment works to adapt to changing water quality. Alternatively, high investment to safeguard against eventual water quality degradations could prove futile if these fail to materialise. Additionally, treatment is energy intensive (Del Río-Gamero et al. 2020), so the challenge is further emphasised by obligations to reduce greenhouse gas (GHG) emissions consistent with Net Zero GHG targets.

Therefore, robust adaptation options that reduce reliance on technology should become a crucial part of a control strategy (Vairavamoorthy, 2021).

Catchment management has raised increasing interest to support favourable water quality outcomes through healthy, functioning, and resilient ecosystems (WWAP, 2018) due to the multiple benefits that can be gained for the environment, society, and economy through delivery of diverse ecosystem services (Everard & McInnes, 2013; Grizzetti et al., 2016; Keeler et al., 2012). Pioneering schemes have been developed through catchment-based partnerships to improve water quality, combining drinking water supply, nature conservation, and amenity agendas, with the aim of achieving not only an economically favourable outcome, but also delivering wider benefits for the environment and society (Appleton, 2002; Morris & Holstead, 2013). As resilient ecosystems can stabilize water quality due to a buffering capacity to pressures, a catchment resilience approach is also particularly interesting to provide a certain amount of protection from climate change impacts.

Incorporating climate change aspects into water quality management remains a challenge to drinking water providers. The need for a forward-looking perspective and anticipatory action is recognised (Garnier & Holman, 2019) and framework guidance available (WHO, 2017b), however practical approaches are lacking. These need to be able to work with current risk assessment approaches. Water Safety Plans offer the possibility to manage risk from climate change at the operational level for individual sites and systems, but water utilities also manage risks at country-level, or programme level, where water resource planners develop initiatives to secure a safe and reliable water supply, including safeguarding raw water quality. At this level, at-risk infrastructure is identified, and decisions made where to focus attention for risk reduction (MacGillivray et al., 2006), making it the appropriate level to identify overall strategies of incorporating climate change aspects and for risk-ranking/prioritising systems which are likely to need intervention to ensure threats to public health are minimised.

Risk assessment for water quality is normally informed by regular monitoring of resources and based upon existing recognised hazards and consequent risks in specific catchments, together with established procedures to react to detected changes in risk levels. Under climate change, past trends and current status cannot adequately present future risk. New approaches are needed to understand how and where climate change could start to challenge existing systems and how risks can be managed. The idea of “uniqueness of place” (Beven, 2000) recognises that each catchment system represents a complex and non-reproducible mix of conditions, but it limits attempts to understand commonalities in systems and to transfer lessons learned from in depths studies. A broader analytical framework to contextualise results and derive general inferences is needed to provide understanding of factors that make catchments vulnerable to changes in drivers, or hazards, and what the consequences are.

Statistical analysis can infer specific water quality parameters based upon catchment characteristics (Davies & Neal, 2004; Rothwell et al., 2010) and improve understanding of underlying drivers (e.g., Selle et al., 2013; Shen et al., 2011, Shi et al., 2017). Understanding the relationships between catchment characteristics and water quality, and identifying certain characteristics or combinations of characteristics that are associated with specific water quality issues is a first step to identifying different vulnerabilities. This can allow to prioritise catchments where a combination of vulnerability and exposure to a hazard increases the risk of unfavourable water quality outcomes, and allow a “catchment profiling” that supports risk-ranking as well as transferability of insights from individual catchments. Challenges and advantages of finding catchment commonalities and typologies are increasingly recognised in hydrological sciences, highlighting the need for large-scale approaches and pooling of datasets to help discriminate and categorise complex cause-effect relationships occurring in catchments in a non-stationary climate (e.g., Beven, 2016; Kundzewicz, 2018; Wagener et al., 2007). As it cannot be assumed that past trends will hold under future conditions, an additional approach to understand how changes will impact on water quality is a space for time substitution as used in geomorphology, where the spatial distribution of landform types can reflect their evolution over time, allowing a projection of development for certain areas (Huang et al., 2019). Using this concept, environmental conditions in one part of the geographical area covered may develop to resemble conditions already experienced in another part, which could allow conclusions about future development of ecosystems (Lester et al., 2014; Meerhoff et al., 2012), and associated water quality outcomes.

1.1 Research approach and aims

Incorporating climate change considerations into risk assessment for drinking water quality needs to be efficient and reflect the reality of water service providers in terms of financial constraints. Water utilities have to justify investments, and the demand for tools that work with existing procedures and structures requires approaches that can be incorporated at a programme level, following familiar patterns of risk-ranking and prioritising. This allows focusing attention on systems where risks are perceived to be highest, and to include climate change aspects in the risk assessment and management procedures at the site level. It should also build a basis for transferability of knowledge between sites, and a frame for a strategic evaluation of risk management and response options at the strategic level of the water utility.

This suggests a staged approach (Figure 1.1). The first stage establishes diagnostic water quality risk profiles for catchment-water source systems, by analysing and describing current water quality hazards, patterns of hazard exposure, and intrinsic catchment vulnerabilities, based upon available data on climate, topography, soils/lithology, and land use/land management factors. Thus, key risk relations, and their spatial and temporal dimensions, are characterised for the supply system as a whole. The second stage focuses on changing patterns of exposure and vulnerability derived from data on direct and indirect effects of climate and land use changes. This allows a first order ranking of systems most at risk from climate change and for prioritisation at programme level. At stage 3, response options can be reviewed for their potential and suitability. On site level, starting with individual high-risk catchments or groups of catchments with similar risk profiles identified at stages 1 and 2, in-depth analysis can be carried out where necessary to understand direct and indirect risks more fully, and risk control options are reviewed to ensure appropriate responses. At stage 4, catchments can be assessed as part of the large-scale resource system for strategic decisions on its sustainability. Overall strategies can be developed based on identified catchment groupings or profiles that suggest the need for e.g., increased monitoring and research to understand processes in more detail; mitigation and restoration measures that increase resilience; or stakeholder engagement to better co-ordinate anticipatory adaptation strategies. Overall sustainability can also be assessed against other policy objectives. Updates of the risk assessment with emerging evidence, new data and changed conditions are included through regular or targeted reviews and revisions of response strategies.

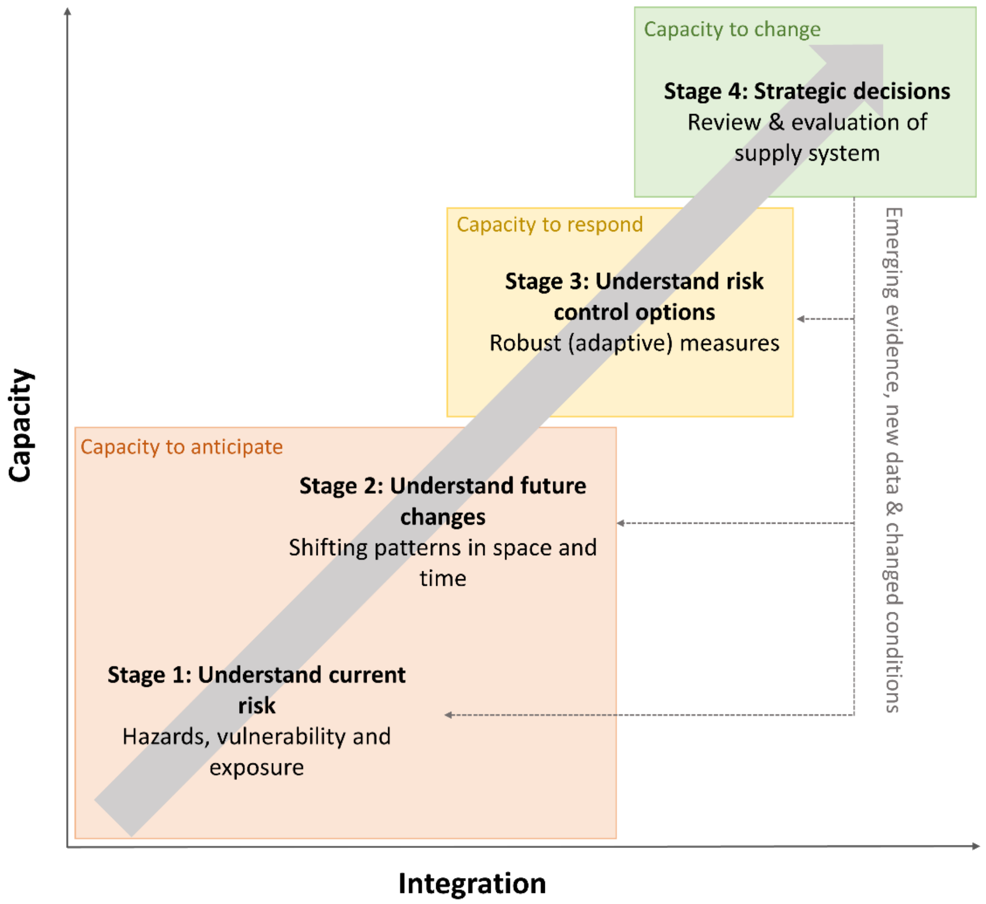


Figure 1.1: Schematic illustration of proposed approach. Moving through stages 1- 4 successively advances integration of climate change impacts into existing water supply risk assessment procedures and builds capacity to adapt to impacts of climate change. Stages 1 and 2 support anticipation of impacts, stage 3 supports evaluation of responses, and stage 4 supports evolvement to adapt.

Appropriate tools for each of the stages are required. Risk screening offers scope for the large-scale approach needed at stages 1 and 2, drawing on pooled data for empirical modelling, while at stage 3 tools are required to support a more detailed analysis and risk control option identification and implementation. At stage 4, approaches and tools are needed that support connecting multiple overarching objectives and goals.

The research aims to test the proposed staged approach in a national context for Scotland, based on Scotland's public water supply. Thereby, the following research questions (RQs) are explored:

RQ 1: How is climate change likely to impact raw water quality of Scotland's public water supplies?

The basis for exploring this question is that water quality is intrinsically linked to the source catchment characteristics, natural or modified through human activity. Therefore, changes that catchments experience will translate into changes in water quality. Understanding which catchment properties influence water quality provides insight into catchment sensitivities, and how this makes them vulnerable to drivers and pressures. Looking at how pressures are projected to change then provides a basis for estimating potential impacts on catchments and thus water quality. To this end, the thesis explores the following aspects:

Assessment 1.1: What are current water quality concerns regarding drinking water, how do catchment characteristics relate to these, and can water quality "profiles" be identified, i.e., can catchments be grouped or categorized according to similar patterns in current water quality and related pressures?

This question relates to stage 1 in the proposed approach. Moving on to stage 2:

Assessment 1.2: How are current issues likely to be impacted by climate change due to catchment sensitivities, and what are consequences for other water quality issues, including emerging concerns?

Drawing on results from these assessment steps under RQ 1, the thesis will further explore:

RQ 2: How can risks to water quality from climate change be managed in a drinking water context?

The assessment will identify gaps in knowledge and help to distinguish aleatory and epistemic uncertainties. To answer RQ 2, the thesis will discuss what current limits to knowledge are; if, where, and how these limits might be pushed, and what tools could help to further identify, shape, and implement adequate response strategies, thus enabling stage 3 and 4.

In this context, the thesis aims to reflect on:

RQ 3: What role does a catchment approach play in mitigating and adapting to climate change?

There has been an increasing focus on catchment approaches to stabilize and improve water quality. The thesis will use the results of RQ 1 and RQ 2 to discuss the benefits of ecosystem-based adaptation, the challenges to implementing catchment management approach, and ways to take better advantage of the benefits that also support advancing multiple policy objectives.

The national assessment of Scotland is used to demonstrate the approach and draw conclusions on how a more systematic and strategic inclusion of climate change considerations into water quality management could be achieved. It also allows to more broadly reflect on the merits of

- a dataset with extensive spatial coverage for drawing conclusions about temporal dimensions of risk;
- risk-screening as a first-order risk ranking tool to facilitate the inclusion of future and emerging issues into established risk assessment procedures;
- catchment profiling to contextualise results and support programme-level risk management and decision-making;
- different types of modelling approaches (empirical, risk-based, process-based) for different stages of risk assessment

Scotland relies on surface water sources especially for larger supply systems, rather than on groundwater which is a very important source of drinking water in many countries globally, and especially for smaller supplies (Howard et al., 2006). Nevertheless, the proposed framework is adaptable to different, catchment-based supply systems. Using Scotland as a demonstration case benefits from exploring a heterogenous landscape, diverse water resource management approaches, and policy recognition for improved risk-based approaches to help secure and maximise multiple benefits in a changing world (Scottish Government, 2011; Water Resources (Scotland) Act 2013). Scotland has the ambition to become a ‘Hydro Nation’, which seeks to use its knowledge and expertise to benefit the economy through innovative management practices (Muscatelli et al., 2020). Public-sector organisations such as Sottish Water, as the provider of public water supply and wastewater services, are key partners in this endeavour, pushing their role towards contributing to overarching policy objectives that underpin Scotland’s sustainable development. This research seeks to advance this by providing insights into approaches that support risk assessment, management, and decision-making for multiple benefits under deep uncertainty.

1.2 Structure

The thesis is split into 7 chapters (Figure 1.2). Chapter 2 gives further background on relevant concepts and terms in the context of climate change and water quality, risk assessment, catchment management, and includes an introduction to Scotland. Chapter 3 describes the approach to the national assessment, including data sources, preparation, and summary statistics for both catchment data and water quality data, and explains how these data are used in the research. Chapter 4 covers the analysis into relationships between catchment characteristics and water quality and explores possible catchment profiling, constituting Assessment 1.1. It includes a review of relevant literature in the field of catchment - water quality relationships, an explanation of the methodologies used for the analysis and discusses the results, also with regard to their implications for the further work within this research. Chapter 5 and 6 both further analyse data on specific water quality indicators, colour and *E. coli*, aiming to provide Assessment 1.2 and resulting in risk screenings for these particular indicators. These chapters include some further background on current knowledge around sources and processes for these water quality indicators, and explain rationale and methods for the risk screenings. Chapter 7 more broadly discusses the results from chapter 4, 5 and 6, to reflect on lessons learned from the national assessment for all RQs.

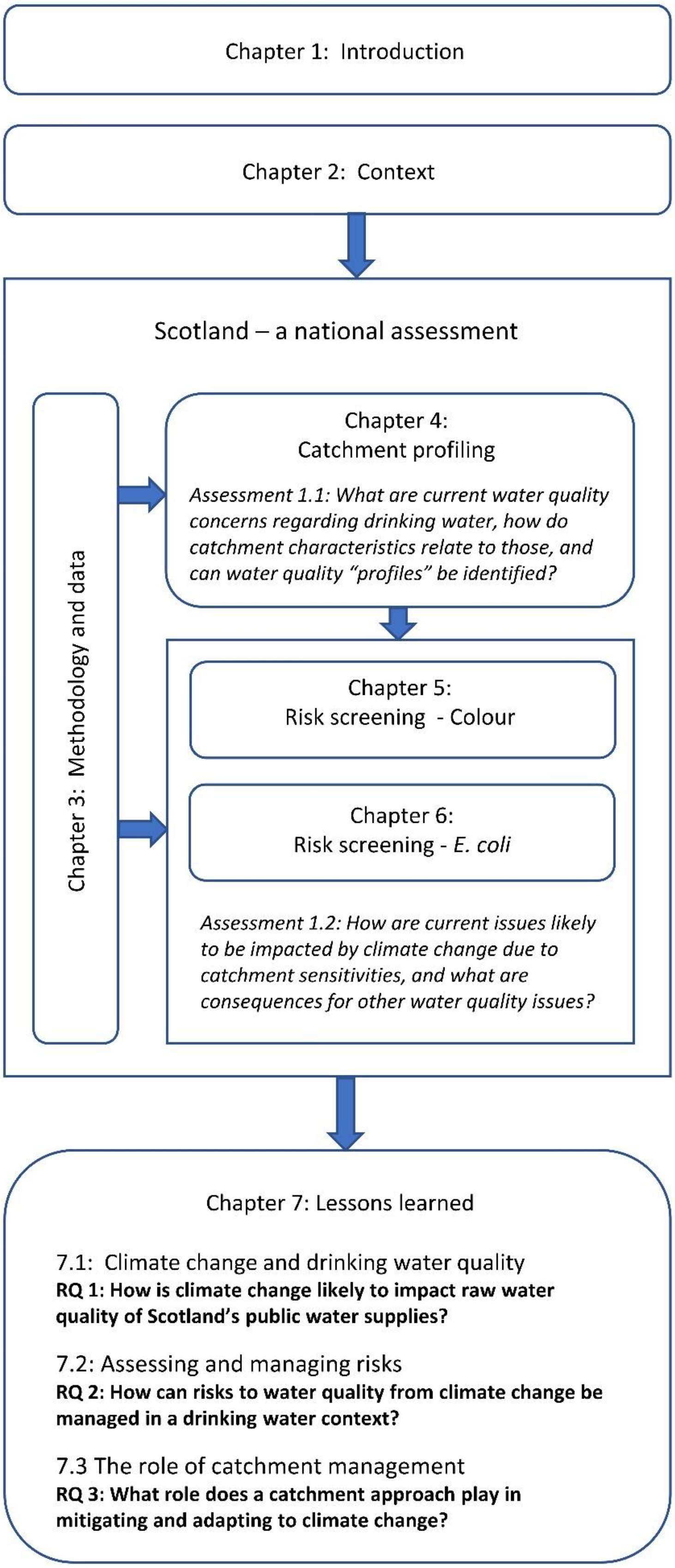


Figure 1.2: Schematic illustration of the content and structure of the thesis. Chapter 1 introduces the topic and research questions (RQ). Chapter 2 provides a wider context and the background for the national assessment used to test the proposed framework for integrating climate change impacts into raw water quality risk assessment. Chapter 3 – 6 present the approach and results. Chapter 7 discusses the assessment in the context of the outlined RQs.

2. Scientific, policy and socio-economic context

The research sits in the context of risk assessment for managing climate change impacts on drinking water quality, using data from Scotland's national public water supply utility. This section provides necessary background on important concepts, terms, and literature in relevant fields. It starts with a summary of projected climate change and implications for water quality (2.1). It continues with an overview of risk assessment in general and within the climate change and drinking water context, and introduces approaches and tools that can support risk assessment (2.2). It reviews measures to protect water quality within the catchment and concepts around building resilient catchments (2.3), and finally explains the regulatory framework, natural and man-made conditions, and climate change impacts for water quality in Scotland to provide the context for the national assessment (2.4).

2.1 Climate change and water quality

Over the past decades it has become increasingly clear that global warming driven by emissions from human activities is happening and that we are experiencing changes unprecedented over millennia. The nature and impact of global warming will vary across regions, but all regions are expected to experience more hot extremes and less cold extremes, and most regions are expected to see intensified extreme precipitation, flooding, and drought events (IPCC, 2021). This has profound impacts on hydrological regimes with consequences for water quality.

Water quality is determined by the catchment - the area of land from which water drains into a stream, river, or lake - and the properties of and complex interactions between atmosphere, land, and water within this area. Substances are exchanged with the air, soil, flora, and fauna and reach water. Human activity also influences the presence and concentration of substances either by introducing them directly or by influencing natural processes and changing the equilibrium of exchange (Boyd, 2015). Water quality will vary throughout the year and between places as it is influenced by the hydrological processes that transport water. The water quality at a point in a stream or lake, at a given time, is a result of the flow paths, including direct runoff, shallow flow-through (water residing in the ground a few days) and groundwater base flow (water residing in the ground for years) (Weatherhead & Howden, 2009). There are numerous physical, chemical, and biological variables to describe water quality, and what constitutes “good” quality and which variables are used to describe it depends on the purpose (Boyd, 2015). In terms of drinking water, good water quality is usually defined as being safe for human consumption (see 2.2.2).

2.1.1 Drivers for water quality

If water quality doesn’t meet the requirements, it’s considered polluted, and origins of pollution can be local as well a global, natural, or man-made. Today, even remote or protected areas of the world are influenced by anthropogenic pressures. Emissions can be transported through the atmosphere, one well established effect being acidification of catchments through sulphur, and increasingly through nitrogen oxide from vehicle emissions, although heavy metals are also deposited in catchments and reach surface waters (Soulsby et al., 2002). Over the past years, it has also increasingly been realised that microplastics have been distributed globally through aquatic environments, with severe impacts on living organisms (Hamid et al., 2018).

Local pressures include pollution through discharges from urban and industrial sources. Urbanisation is associated with deterioration in water quality due to increased sediment load, heavy metal and nutrient input, and bacterial pollution (Schoonover & Lockaby, 2006). Because of the high percentage of impervious areas, pollutants on surfaces can get washed directly into surface waters. Pollution occurs during storm events and sewer overflow, through treated or untreated outfall and through diffuse sources (Borja et al., 2006). Industries such as petrol, chemicals, paper, textiles, food processing or construction can cause pollution through discharges (Borja et al., 2006). Old and abandoned mines can discharge polluted water rich in sulphate, iron, or aluminium into surface waters (Robins, 2002).

Agriculture has a direct impact on stream water quality through point and diffuse pollution. Arable cultivation and improved pasture are associated with increased nutrient runoff and have been identified as one of the major inputs for nitrogen (Allcock & Buchanan, 1994; Bouraoui & Grizzetti, 2014). Pesticides and bacterial pollution also originate from agricultural activities (Reichenberger et al., 2007, Hooda et al., 2000).

Natural forests are generally associated with good water quality and some forest ecosystems, particularly high elevation cloud forests, can even provide a higher quantity of water (Dudley & Stolton, 2003). By contrast, plantation afforestation can have a number of negative impacts on water quality (Calder, 2007). Forests can deteriorate stream acidity by scavenging pollutants from the atmosphere, by increasing nutrient uptake from the soil, through soil erosion associated with tree-planting, and drainage operations (Soulsby et al., 2002). Roads established for woodland management can also change water balance relations and hydrological response, creating new interception, throughflow and runoff dynamics (van Dijk & Keenan, 2007). Forestry operations involving the use of heavy machinery and using poor practice in site preparation, planting, thinning, and harvesting, and frequent disturbance through firewood and litter collection or overgrazing, can damage the soil, increase runoff, and mobilise sediment and other pollutants (van Dijk & Keenan, 2007). In contrast, afforestation with native woodland is thought to have almost no implication on stream acidity, although catchments with managed forests have shown lower levels of pollution with faecal coliforms than unmanaged forests, possibly due to a lowered biodiversity (Schoonover & Lockaby, 2006).

A general degradation of natural habitats, especially wetlands, caused by a multitude of impacts including for example the introduction of invasive species or habitat fragmentation, reduces ecosystems’ natural ability to purify water, also leading to water quality degradation (Hefting et al., 2013). Bogs and fens for example act as nutrient retention areas, particularly for phosphorus and inorganic nitrogen originating from fertilizer use (Bragg, 2002). Artificial drainage of peatlands on the other hand can lead to high concentrations of dissolved organic carbon and of brown ferric hydrites (Heathwaite et al., 1993).

Pressures on and pollution sources for catchments and water sources and their impacts for water quality are summarised in Figure 2.1.

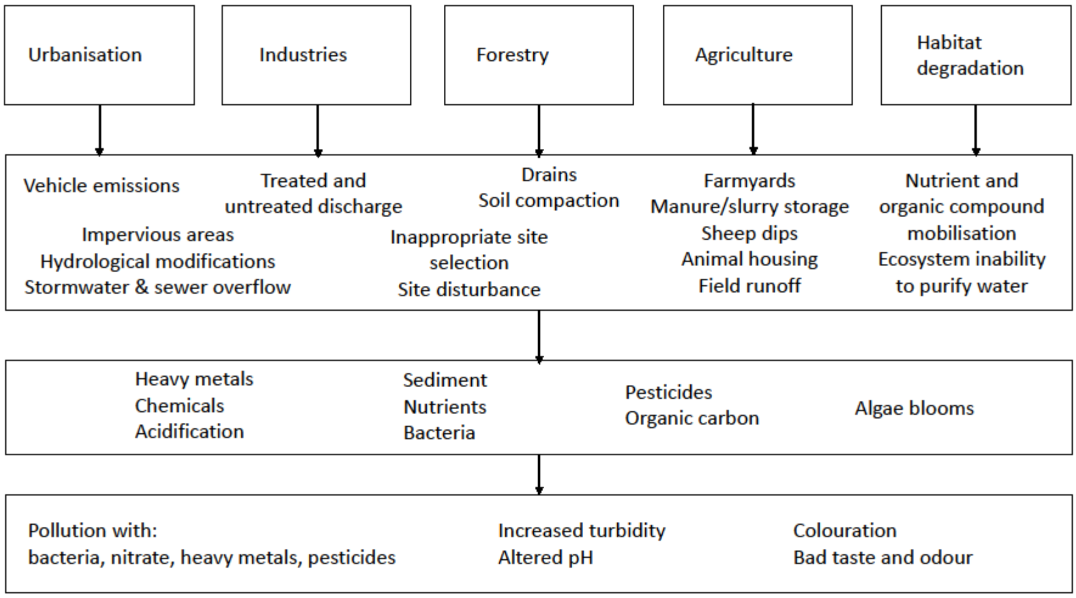


Figure 2.1: Illustration of drivers resulting in catchment pressures, causing polluting substances to reach surface water bodies, and consequences for raw water quality.

2.1.2 Future changes and their impacts on raw water quality

Each of the last four decades has been successively warmer than the one before (IPCC, 2021). Global surface temperature between 2001 and 2020 has risen by 0.99°C from the period of 1850-1900 (IPCC, 2021). Impacts of these changes include poleward shifts of climate zones in both hemispheres and a lengthening of the growing season, increases in extremes such as heatwaves, droughts, heavy precipitation, floods and tropical cyclones, alteration of hydrological regimes through changes in precipitation or melting snow, and many species have experienced shifts in geographic ranges, seasonal activities, migration pattern, abundance, and species interaction. Surface temperature will continue to rise over the 21st century by at least 1°C, and potentially by up to 5.7°C, with impacts increasing in direct relation to increases in temperature (IPCC, 2021). Climate change has already caused substantial damages and losses in terrestrial, freshwater, coastal and marine ecosystems, with deterioration of ecosystems structure and function, resilience and natural adaptive capacity, and associated negative socioeconomic impacts. Increases in frequency and intensity of extremes, and changes to streamflow magnitude, has led to a loss in water security and it is projected that risks associated with water availability and water-related hazards will continue to increase (IPCC, 2021). In terms of water quality, climate change is expected to have direct as well as indirect consequences.

Changes in water temperature, soil temperature, and precipitation patterns affect concentration as well as number of substances in surface waters. An increase in water temperature will change physical and chemical processes, for example dissolution in water, solubilisation, complexation, degradation, and evaporation, leading to increased concentration of dissolved substances, decreased concentration of dissolved gases, and changes to the rate of bacteriological processes which will favour the survival of pathogens and lead to toxic algal blooms (Delpla et al., 2009; Watts et al., 2015). Low flows in summer mean that there is less water available for dilution of pollution, which may cause problems downstream of point sources. Increases in precipitation will likely lead to increased solute and sediment transport, and nutrient leaching (Watts et al., 2015). It is also possible that climate change will cause more runoff of particulate and dissolved organic carbon from peatlands. Increasing trends in dissolved organic carbon (DOC) have been observed across the Northern Hemisphere for the past decades (Monteith et al., 2007; Sawicka et al., 2017), although there is no scientific consensus what the main drivers for increased DOC release into surface waters are (Delpla et al., 2009). Discussed causes are increasing temperatures (Cole et al., 2002; Freeman et al., 2001), recovery from acidification (Evans, Chapman et al., 2006; Monteith et al., 2007), changes in hydrology (Tranvik & Jansson, 2002), and land management and peatland drainage (Worrall, Armstrong & Adamson, 2007; Worrall, Armstrong & Holden, 2007; Yallop & Clutterbuck, 2009). Apart from the problems organic carbon itself causes for drinking water quality, it is also possible that DOC facilitates the transport of dissolved lead, titan, and vanadium in peatlands after storm events (Rothwell et al., 2007). Extreme weather events are known to cause water quality incidents (Khan et al., 2015; Young et al., 2015) and more frequent extremes are therefore concerning for drinking water providers.

Indirect impacts of climate change are effects on urban, industrial, or agricultural activities that in turn affect water quality through distribution as well as intensity of pressures. Important changes in land use include a likely change in land used for agriculture (Brown et al., 2010, Rounsevell et al., 2006), with associated changes in pressures induced by agriculture. For Europe, forest areas are also predicted to increase (Rounsevell et al., 2006). Growing populations will increase pressures from urbanisation, e.g., with more frequent or more severe sewer overflow events (AECOM, 2015), but also from industry and agriculture. Carbon emissions mitigation measures can also affect water quality and quantity, for example the expansion of renewable energy schemes such as hydropower (Rosenberg et al., 1997).

2.2 Risk assessment and management

The term, or the concept of, 'risk' lacks a universal definition, and what risk is understood to mean and comprise often depends on the disciplinary context (Wassénius & Crona, 2022). In the simplest form, it is defined as the product of likelihood of an event multiplied by its consequences (Adger et al., 2018; Gormley et al., 2011), leading to risk matrices that allow the scoring or ranking of risks (Figure 2.2).

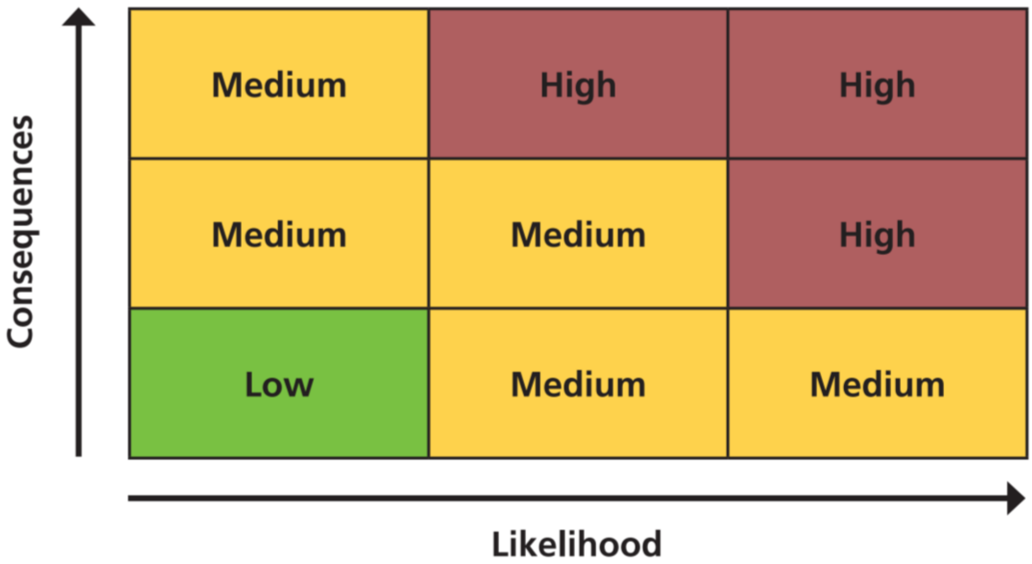


Figure 2.2: Example of a risk assessment matrix, used to assign risk levels according to their likelihood and severity of consequences (reproduced from Gormley et al., 2011).

A risk assessment under this definition then in sequence: identifies a hazard (an agent that may cause adverse effects), assesses the consequences, and assesses the probability, to characterise the risk. Especially in a climate change context however, this way of defining risk has been criticised for its limitations (Adger et al., 2018; Challinor et al., 2018).

2.2.1 Climate change context

Traditional approaches to risk assessment, based on probabilities derived from historic data, are no longer adequate in a climate and global change context where "stationarity is dead" (Milly et al., 2008) and uncertainty high. Uncertainty can be defined as "any departure from the (unachievable) ideal of complete determinism" (Walker et al., 2003). Walker et al. (2003) distinguish five levels of uncertainty between complete certainty and total ignorance, with level 4 (multiple plausible alternatives for which no likelihoods can be assigned) and level 5 (unknown future) classed as deep uncertainty. A distinction is also made between aleatory and epistemic uncertainty, with aleatory uncertainty stemming from natural variability, and epistemic uncertainty due to lack or ambiguity of knowledge (Maier et al., 2016). While it can be argued that all uncertainty is epistemic as lack of, or inadequacy of, knowledge (Der Kiureghian & Ditlevsen, 2007), it is practicable to distinguish between uncertainty that can be reduced through, for example, increased data sharing, collection and analysis, and uncertainty that is practically not reducible.

I. The concept of risk

In view of this, the concept of risk in a climate change context has evolved over the last decades and is much influenced by the Intergovernmental Panel on Climate Change (IPCC) and its reports. From its 5th Assessment Report (2014) on, risk is defined as "the potential for adverse consequences for human or ecological systems". This broad definition acknowledges that probabilities cannot always be quantified and that ways must be found to assess and manage risks despite uncertainties and complexities. In a climate change impact context, risks "result from dynamic interactions between climate-related hazards with the exposure and vulnerability of the affected human or ecological system". Hazard is defined as "the potential occurrence of a natural or human-induced physical event that may cause loss of life, injury, or other health impacts, as well as damage and loss to property, infrastructure, livelihoods, service provision, and environmental resources". Exposure is "the presence of people; livelihoods; species or ecosystems; environmental functions, services, and resources; infrastructure; or economic, social, or cultural assets in places and settings that could be adversely affected". Vulnerability is "the propensity or predisposition to be adversely affected. [It] encompasses a variety of concepts and elements including sensitivity or susceptibility to harm and lack of capacity to cope and adapt". Risk is thus the result of the interactions of hazard, exposure, and vulnerability (Figure 2.3).

Figure 2.3: Risk characterisation by the IPCC, showing risk as the interaction of hazard, vulnerability, and exposure (reproduced from IPCC, 2022).

II. Vulnerability and adaptation

While these terms are defined as above in the latest IPCC Assessment Report, they are sometimes used with different conceptions. Vulnerability was originally defined as being composed of exposure, sensitivity, and adaptive capacity. In this context, exposure is “the nature and degree to which a system is exposed to significant climatic variations”. Sensitivity is “the degree to which a system is affected, either adversely or beneficially, by climate variability or change. Adaptive capacity is “the ability of a system to adjust to climate change (including climate variability and extremes) to moderate potential damages, to take advantage of opportunities, or to cope with the consequences”. Exposure has thus shifted in its conception, from focusing on impact (“the degree…”) to spatial conceptualisations (“the presence…”) and has become a part of the risk concept. Sensitivity and adaptive capacity remain important components of vulnerability, although focusing on harmful impacts. It is also distinguished between outcome, or end-point vulnerability, and contextual, or starting-point vulnerability. Contextual vulnerability describes the present inability to cope, while outcome vulnerability describes the outcome of a sequence of analyses and consequences that remain after adaptation has taken place (IPPC, 2014).

Adaptation in turn is “the process of adjustment to actual or expected climate and its effects” (IPCC, 2014), and it involves actions spanning local, national, and regional scales from individuals to public bodies, governments, to international agencies (Adger et al., 2005). Comprehending the effectiveness of adaptation actions is hampered by several factors, among them uncertainty about the future state of the world (Adger et al., 2005).

III. Resilience

Another important concept in this context is that of ‘resilience’, a term that originally stems from ecology and refers to the capacity of ecosystems to absorb external disturbances, maintaining its functions and services. It encompasses the system’s ability to resist, not losing its processes and structures; and to recover, returning to its original state after a disturbance. The term has been taken on in social sciences and for socio-ecological systems (Tompkins & Adger, 2004; Gallopin, 2006). Especially in the context of complex social systems and climate change adaptation, the concept has been questioned and the term has been widened to include a system’s capacity to self-organise and adapt, and to adapt and cope, although the latter could be seen as impinging on the concept of adaptive capacity (McEvoy et al., 2013). There remains a lack of consensus around the term, and a gap between the theory of “increasing resilience” and understanding of how to achieve this in practice (Morecroft et al., 2012). The IPCC follows a wider definition of resilience, adapted from the Arctic Council: “The capacity of social, economic, and environmental systems to cope with a hazardous event or trend or disturbance, responding or reorganizing in ways that maintain their essential function, identity, and structure, while also maintaining the capacity for adaptation, learning, and transformation” (IPCC, 2014). Resilience remains an important concept in climate change adaptation management to avoid negative impact of climate change.

2.2.2 Drinking water context

The primary objective of a drinking water provider is to provide safe and sufficient drinking water reliably and affordably to consumers. The Bonn Charter, developed by the International Water Association (IWA), sets out an industry “best-practice” framework for water management with an emphasis on consumer interest, transparency, and stakeholder responsibilities, and formulates the overarching goal as “good, safe drinking water that has the trust of consumers” (Figure 2.4).

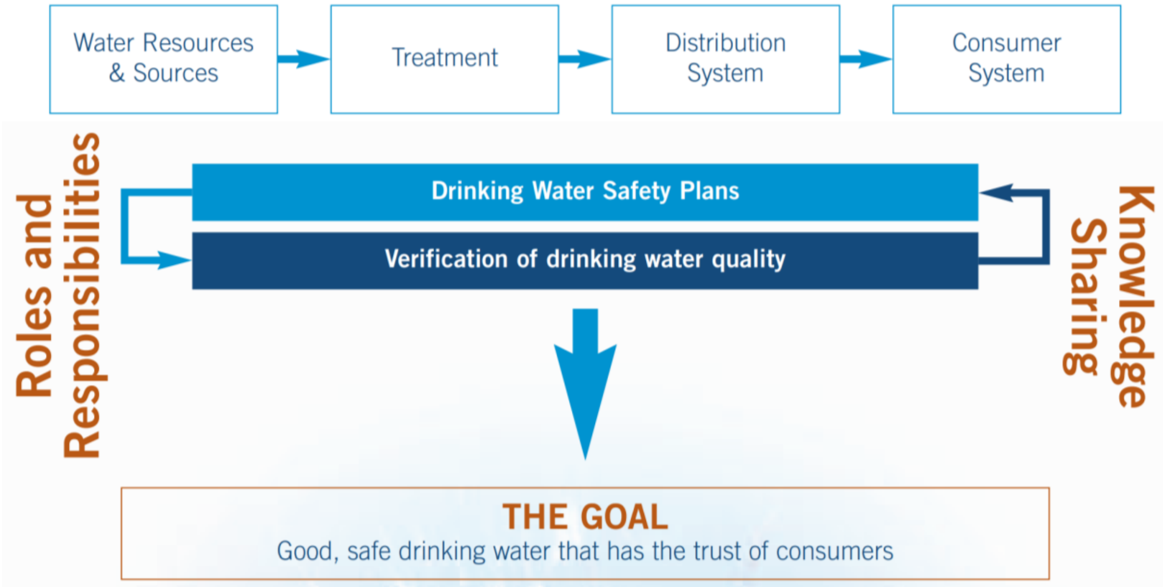


Figure 2.4: Bonn Charter framework for the delivery of safe and reliable drinking water (reproduced from IWA, 2004). The supply system is here shown as comprising of four major chain components: Water resources and sources, treatment, distribution system and consumer system, all of which are covered and risk assessed under Water Safety Plans, to achieve the goal of good, safe drinking water that has the trust of the consumers.

Under the Sustainable Development Goal 6, Target 6.1 is to achieve “universal and equitable access to safe and affordable drinking water for all [by 2030]”. While there is no international benchmark on what constitutes affordable, it is understood that paying for safe drinking water “must not limit people’s capacity to acquire other basic goods and services guaranteed by human rights, such as food, housing, health, clothing and education” (United Nations General Assembly, 2015). In many countries, independent regulators oversee affordability of water services (WHO, 2017a). So, water utilities need to focus on making sure there is always enough (reliably sufficient) water that always meets the required quality (reliably safe), and to do this as economically efficiently as possible (affordably). Especially the latter part means that a risk-based approach is increasingly used, whereby risks from failing the objective are balanced against the costs of securing it.

I. Standards for drinking water quality

In terms of ‘reliably safe’, the World Health Organisation states that “safe drinking-water, as defined by the Guidelines, does not represent any significant risk to health over a lifetime of consumption, including different sensitivities that may occur between life stages” (WHO, 2017c). The levels of contaminants that meet these requirements, or in other words, that constitute an acceptable level of risk, are usually set out in water quality standards. There are no internationally set standards, but the WHO has produced guidelines that set levels for a wide range of water quality parameters, and most countries have standards for drinking water quality that are aligned to the WHO guidelines (WHO, 2017a). Such regulatory standards are the ‘reliably safe’ margins and water utilities are hence not necessarily concerned about improving water quality to the maximum possible, but to the required minimum.

The main reason for treating drinking water is health related, and dangers to health can arise from microbiological pollution, and pollution with chemicals or radioactive substances. Apart from health however, water might also be treated for aesthetic and acceptability reasons to improve appearance, odour, or taste.

Microbiological pollution, including bacteria (e.g., *Escherichia coli*, S*almonella*), viruses (e.g., Hepatitis, Noroviruses, Rotaviruses) and parasites (e.g., *Cryptosporidium*), mainly originate in drinking water from human or animal faecal pollution. They reach water bodies in a variety of ways, e.g., through badly sited latrines and septic tanks, spills of slurries from farms, runoff from agricultural land or hard surfaces, or discharges from treatment works (Fawell & Nieuwenhuijsen, 2003). Sometimes, organisms also grow in source water or in water distribution systems (WHO, 2011). Waterborne diseases are still a significant cause of death in many parts of the world (Kanamori et al., 2016; Ma et al., 2022; Prüss-Ustün, 2016), and outbreaks of infectious diseases are the most common health risk associated with drinking water.

Except for nitrate, chemical pollution often only becomes a health risk after an extended exposure of several years (WHO, 2011), as exemplified in Bangladesh and West Bengal, India, where arsenic naturally occurs in groundwater sources (Chowdhury et al., 2000). Chemicals might enter the water sources from natural deposits in rocks and soils, from industrial sources or human dwellings, such as processing factories, mines, and sewage, or from agricultural activities. Some chemicals, such as trihalomethanes, also occur in the disinfection process (WHO, 2011). Naturally occurring chemicals of health concern are arsenic, barium, boron, chromium, fluoride, selenium, and uranium. Some chemicals that are not of health concern may however affect the acceptability as drinking water, such as iron or manganese. Chemicals from industrial sources that are of concern are mainly heavy metals and solvents and include for example cadmium, mercury, benzene, dichlormethane, or tetrachlorethane, among others. Chemicals from agricultural sources are mainly pesticides, and nitrate. Nitrate and nitrite are linked to methaemoglobinaemia in infants and toxicity in adults (Sharpley et al., 2009).

Dangers from radioactive substances in drinking water is usually much smaller than from microbial or chemical pollution. Radionuclides can enter water naturally or from man-made sources, the latter of which are usually easier to control and have lower radiation. Naturally occurring radionuclides are therefore of higher concern (WHO, 2011).

Water that has an unpleasant appearance, odour or taste is likely to be rejected by the consumer, even if it is safe to drink. This can be caused by biological sources or processes, chemicals, or during water treatment. Examples include cyanobacteria (causing colouration and turbidity after filtration, or compounds detectable by taste), chlorine (impacting taste and odour), iron and manganese (causing colouration) or sulphate (noticeable taste, as well as causing a laxative effect in high doses on unaccustomed customers) (WHO, 2011).

There is no international binding standard of what constitutes safe water. The setting of water quality standards is perceived to be a national affair, and many countries have adopted a set of standards or guidelines for drinking water providers to protect the health of their citizens and ensure provision of water that is safe and acceptable for drinking (see Table 2.1 for examples).

Table 2.1: Selected substances of interest in drinking water, with their sources, health implications, and standards set in the WHO drinking water guidelines (WHO, 2017c) and in the EU drinking water directive (Directive (EU) 2020/2184).

Substance	Source	Implication for drinking water	WHO guidelines	EU standard
Arsenic	Widely found in the Earth's crust. Mainly low concentrations in water, but groundwater can have elevated levels.	Can cause dermal lesions, peripheral neuropathy, skin cancer, bladder cancer, lung cancer. Effects on cardiovascular system in children.	10 µg/l	10 µg/l
Benzene	Vehicle emissions are the main source for benzene in the environment.	Acute exposure effects the central nervous system. Lower concentrations can cause haematological changes, including leukaemia. It is carcinogenic to humans.	10 µg/l	1 µg/l
Cadmium	Occurs in water through wastewater discharges or diffuse pollution from fertilizers or air pollution.	Accumulates in the kidneys, but no evidence of carcinogenicity.	3 µg/l	5 µg/l
Colour	Predominately because of coloured organic matter, but also through iron or other metals.	Might indicate pollution. Colour can also raise doubts about the safety of the water in customers and cause supply of water from another (potentially unhealthy) source.	none	Acceptable to consumers and no abnormal change
E. coli	Contamination with human or animal faeces.	Indicator for faecal contamination, so detection will/should trigger further investigation or supply stop.	0/100 ml	0/100 ml
Fluoride	Naturally present. Elevated levels through dental preparations.	Offers dental protection at concentrations of 0.5mg to 2mg per litre of water but may also cause mild dental fluorosis at concentrations of 0.9-1.2mg/l	1.5 mg/l	1.5 mg/l
Iron	Naturally present or as a result of corrosion in pipes.	No health concerns. Can cause colouration.	none	200 µg/l
Nitrate	Naturally present, but can also reach water from agricultural activities, wastewater discharges, septic tanks.	Nitrate can be reduced to nitrite in the gastrointestinal tract, especially in individuals with gastrointestinal infections. Nitrite reacts with haemoglobin, blocking oxygen release. In infants, this can cause cyanosis (blue baby syndrome).	50 mg/l	50 mg/l
Trihalomethanes	Result of chlorination of organic matter.	Some forms may be carcinogenic, but research is inconclusive. Possible increased risk of stillbirth or spontaneous abortion through Bromodichloro-methane.	The sum of the ratio of the concentration of each to its respective guideline value should not exceed 1	100 µg/l (total)

II. Water utility risk concepts

MacGillivray et al. (2006) summarise risks posed to water providers and introduce a risk hierarchy explaining that risks are managed at the strategic, programme and operational level (Figure 2.5).

Figure 2.5: Conceptual illustration of the risk hierarchy for water services providers, comprising risks on strategic, programme, and operational levels (reproduced from MacGillivray et al., 2006).

At the strategic level, risk management for the water utility includes risks from regulatory systems, from competitors, risks from improving financial and operation efficiencies (business process re-engineering), from new technology, from outsourcing activities such as maintenance, distribution, billing etc., and from employee retention.

On the programme level, risk management involves analysis and management of assets, network, catchments, and vulnerability to large-scale disasters. This involves an integrated, systematic process that identifies and analyses at-risk infrastructure, often involving a system of prioritising or risk-ranking assets and network components to inform and support risk reduction efforts and focus attention on the most serious threat to system performance. Risk screening approaches seem particularly suited for this level. In the context of catchments, this often means identifying and concentrating on areas and measures that offer the highest possibility of reducing or preventing severe impacts from occurring. Tools employed for these analyses include risk mapping using Geographic Information Systems (GIS) or, where more detailed analysis is necessary, model-based approaches (MacGillivray et al., 2006).

On the operational level, site managers have to manage treatment and distribution processes to ensure uninterrupted compliant supply. Failure could stem from source contamination, human error, mechanical failure, or network intrusion. Risks from failure can be examined by looking at the following components:

- Hazard: What are the effects of failure (mainly human health implications, but also e.g., restrictions in water use)?
- Exposure: What is the size and characteristics of the population effected and for how long?
- Vulnerability: What is the relationship between the hazard and the exposure?

This can then be combined to estimate the magnitude, variability, and uncertainty of the problem, i.e., the risk. Reliability analysis then identifies potential points of failures to understand where and what kind of changes are required to reduce the likelihood of the failure from occurring. A framework for this kind of assessment are Water Safety Plans as set out by the World Health Organisation.

III. Water Safety Plans

To achieve SDG target 6.1, the WHO outlines and recommends Water Safety Plans in their Guidelines for drinking water quality (WHO, 2017c). This "holistic approach to the risk assessment and risk management" includes a number of steps aimed at covering the supply chain from the source to tap (Figure 2.6).

Water Safety Plans are developed for each individual supply system consisting of one or more water sources (surface or groundwater) and the associated catchment, the abstraction system, the treatment work, water storage systems and the distribution network. The European Union has adopted the framework by obliging member states to adopt a risk-based approach, however splitting it into separate risk assessments for catchments, supply systems (comprising abstraction, treatment, storage, and distribution) and domestic distribution systems (Art.7, Directive (EU) 2020/2184 of the European Parliament and of the Council of 16 December 2020 on the quality of water intended for human consumption (recast)).

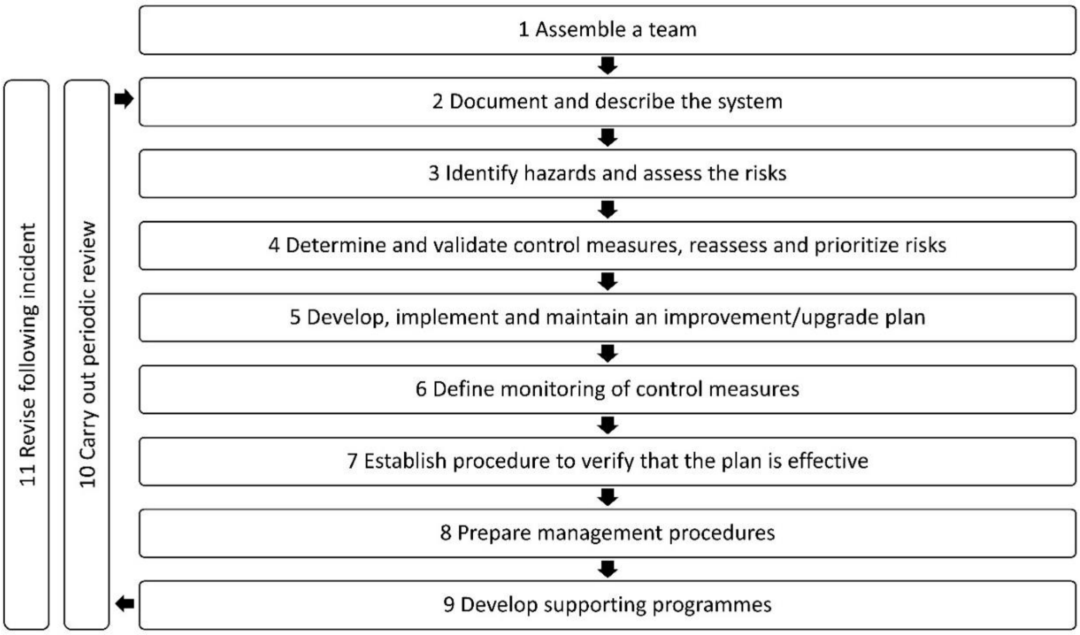


Figure 2.6: Steps for developing a Water Safety Plan, a comprehensive risk assessment for an individual supply system from source to tap (adapted from Bertram et al., 2009 & WHO, 2017c).

While the guidelines acknowledge likely impacts of climate change on water quantity and quality and the need to include these in the assessment, there is no explicit inclusion in the plans, but climate change is seen as a specific circumstance for application of the guidelines. However, the WHO has published separate guidance on “Climate-resilient water safety plans” (WHO, 2017b), describing where climate change aspects can be addressed within the plans: when assembling the team (module 1), when describing the system (module 2), when assessing hazards and associated risks under existing control measures (module 3 and 4), when planning additional control measures and long-term investment (module 5), and when preparing management procedures and supporting programmes (module 8 and 9). Interestingly, the EU directive 2020/2184 specifically mentions that risks from climate change should be considered in the risk assessment for the supply system (Art.9, Nr.2 (c)), but does not mention climate change for the catchment risk assessment.

WHO states that the description of the system should include, among others, the “reliability of source yields (considering seasonal variability and variability between years, for example due to droughts)”; the “historical water quality data and relationship with source yields”, “future climate projections that could impact the water supply”, “water quantity and quality implications of current and projected climatic conditions”, and “trends in land use and population growth impacting water resources supply or demand”. This means that drinking water providers need to know (1) the nature, scope, and uncertainty of possible changes in climate, (2) its impact on quantity and quality of water supply, and (3) the nature, scope, and uncertainty of (potential) control measures.

Climate change could lead to changing the likelihood or magnitude of a problem, or it could lead to newly emerging issues. In WSPs, inclusion of climate change introduces another, temporal, dimension. In practice, this could lead to a change of the risk assessment score, or introduce a new score such as increasing, decreasing that influences a risk-ranking/prioritisation system. For example, an issue that was formerly not prioritised could be put onto the “review” list due to being assessed as “increasing”. It could also introduce a hazard that was formerly not identified.

2.2.3 Approaches and tools

Whereas risk assessment typically involves the statistical analysis of prior events (magnitude and frequency), the challenge for the determination of future risks depends on the dynamic interplay of respective hazard, exposure and vulnerability determinants (Viner et al, 2020). While probable changes in hazards are often acknowledged, it can be more challenging to incorporate temporal dynamics in exposure and vulnerability (Jurgilevich et al., 2017). Related to this is the complexity and inter-relatedness of systems, creating “cascading risks in physical systems, ecosystems, economy and society” (Adger et al., 2018). Policymakers need to make decisions with short-term as well as long-term consequences and at the intersect of different policy-areas, and ways to include risk transmission such as improving modelling methodologies, cross-sectoral modelling, qualitative approaches like analogues and scenarios, and mixing approaches are discussed (Challinor et al., 2018). Furthermore, it is debated how to deal with uncertainty as an inherent feature of risk assessment, with attempts to reduce uncertainties around climate predictions on the one side and arguments that accurate predictions are not necessary for adaptation action on the other. Instead, it is argued that strategies are needed that perform well over a wide range of assumptions about the future and are thus insensitive to uncertainties, so called robust decision processes (Dessai et al., 2009).

I. Robustness

As in many other areas of planning, drinking water suppliers must take decisions now that concern long-term investments, such as upgrading or building new water treatment works, managing or building new reservoirs, switching to or adding new water supply sources, mitigating and averting impacts on water supply catchments and water bodies, etc. These decisions typically span a time of 40-50 years, meaning climate change impacts will have to be considered. There are different possible approaches that could be applied to come to a decision about management strategies and investments: optimum, precautionary, and robust approaches.

In the optimum approach, attention focusses on the scenario that is considered the most likely. This approach is advised where the relationships between cause and effect are well characterised and understood, the values and priorities are clear, uncertainty is well defined, and there is a clear best answer (Lempert & Collins, 2007). However, with climate change, this is very rarely the case. On the contrary, the deeply uncertain nature of climate change projections means that decisions will have to be taken without having information that will reliably tell us the optimum solution. The precautionary approach as an approach for acting under uncertainty sets out to prevent a future harm that is deemed unacceptable, and prevents actions that might lead to this harm. However, it often leads to extremely restrictive actions without balancing multiple goals (Sunstein, 2005). Therefore, robust approaches have been emerging that build on the maxim of finding solutions that deliver acceptable outcomes under a wide variety of futures. There are several robust decision-making frameworks that vary in their ability to address uncertainty and range from being static to dynamic (Walker et al., 2013). Among these are for example Robust Decision Making, Adaptive Policymaking, Adaptation Tipping Points, or Dynamic Adaptive Policy Pathways (Haasnoot et al., 2013; Walker et al., 2013). Many robust decision-making frameworks provide means to include precautionary aspects while allowing an evaluation of trade-off and a discussion about the level of precaution (Lempert & Collins, 2007).

Hallegatte (2009) distinguishes between several robust strategies:

- No-regret strategies, with measures that have positive impacts even in the absence of climate change, for example reducing water leakage from pipes.
- Reversible strategies, for example infrastructure with cheap upgrade possibilities.

- Safety margins that are very low cost, for example designing infrastructure that can cope with the most pessimistic climate scenario from the start rather than retrofitting.
- Soft strategies, including institutional and financial tools, such as long-term plans, insurance schemes or early warning systems.
- Reducing the decision-making time horizon by choosing investments with a shorter lifespan.
- Considering synergies and conflicts between strategies, for example adaptation options that are energy-intensive would not be compatible with mitigation policies.

Robustness and flexibility are important criteria for decision-making under uncertainty, but especially when the consequences of a particular occurrence are disastrous (for example from contaminated drinking water), decision-makers might rather be guided by precaution, meaning that they will design their response under the presumption that this particular risk will materialize. Practically, this usually takes the form of threshold in an observable parameter and discard any action that threatens to exceed this threshold (Lempert & Collins, 2007). While precautionary approaches could lead to very costly measures for reducing a very small risk, measures using the precautionary approach may at the same time be robust (e.g., low-cost safety margins). Decision-makers will have to balance risk against costs and may well apply several criteria to decide on the best response strategy and actions.

II. Models

Models are invaluable tools to provide a basis for decision-making, but they need to be able to efficiently address the purpose. If the model is overly complex, it may introduce uncertainty, require efforts into data collection and computation that is not truly necessary, and effectively hinder decision-making. On the other hand, models that are too simple may not adequately represent key processes or be unable to address relevant management questions (Fu et al., 2019). There are different types of models, and many different individual models for water quality that could be used in the investigation of negative impacts of climate and land use changes on water supply catchments.

Broadly, two uses of models can be distinguished: improve our understanding of catchments and water bodies, or assess different possible impacts (Fu et al., 2019). With regard to the first aspect of improving our understanding, more specifically models can be used to identify and quantify sources of contaminants or examine the importance of explanatory factors. In the absence of high resolution and high frequency data, models can help to fill some gaps and better understand processes, and highlight areas of uncertainty. Regarding the second aspect, models can be used to extend information to a wider region, predict concentrations or loads based on altered inputs or scenarios, and help to develop the range of different possible outcomes based on different futures and uncertainties (Fu et al., 2019; Weaver et al., 2013). Models can also be used to identify gaps in monitoring and where further monitoring sites may be needed to improve model reliability, which is ultimately designed to improve understanding but could well be part of a strategy for mitigating and adapting to climate change. In this way, models are important to aid decision-making, but they need to be able to deal with the data and capacity that is available, their outputs need to be understandable and communicable to decision-makers, and they need to be transparent in what kind of information they provide, and don’t provide, and the uncertainty attached to it.

Empirical, or statistical, models use observations to predict outcomes, whereas mechanistic, or process-based, models use theoretical understanding to mathematically represent processes and enable predictions under altered conditions. While there is a great number of process-based models available for catchment water quality modelling, only a few are widely used, with some more prone to be used for some regions, such as the suite of models developed under the Integrated Catchment model (INCA) in Europe, or the eWater Source platform in Australia and Asia, while the Soil and Water Assessment Tool (SWAT) is the most widely used model worldwide (Fu et al., 2019).

SWAT is a catchment model that has been used to model streamflow, sediment, temperature, nutrient, pesticide, carbon, and pathogen processes worldwide, using a continuous daily time-step (Arnold et al., 1998; Du et al., 2019; Sadeghi & Arnold, 2002). INCA was originally designed for modelling nutrient patterns in aquatic and terrestrial environments (Wade et al., 2002; Whitehead, Butterfield & Wade, 2009), but further models have been developed based on it, including for carbon and pathogens. While the INCA model family has been more extensively used in Europe, the Hydrological Simulation Program-FORTRAN (HSPF), a catchment process-based, lumped parameter model, is more widespread in the US. SWAT and HSPF break up the catchment into so-called Hydrologic Response Units, which are made up from land use in combination with soil type. In INCA, calculations in the terrestrial compartment are performed in 1 km^2 cells. The models require a digital elevation model, soil data, land use data, precipitation and temperature data, as well as flow and pollutant data for calibration and validation. The models have also been used to evaluate the impacts of land use and management, mitigation approaches, and climate change (Aherne et al., 2008; Desai et al., 2011; Ghaffari et al., 2009; Harmel et al., 2010; Hevesi et al., 2011; Kim et al., 2018; Kiros et al., 2015; Moyer & Hyer, 2003; Oni et al., 2012; Peterson et al., 2011; Whitehead et al., 2016).

Process-based models require a large amount of high frequency data for calibration and validation. These highly parameterised, complex models also often struggle with overparameterization, where the model includes more parameters than are necessary and thus effects and processes cannot be disentangled. To address this, simpler models have been developed with more lumped structures, such as by Birkel et al. (2014) or Dick et al. (2015). Another approach to address these issues are risk-based models that rely on qualitative or semi-quantitative data (such as expert opinion) to represent processes to determine relative risk of pollution or pollution source areas. An example for such a risk-based model is the Sensitive Catchment Integrated Mapping Analysis Platform (SCIMAP). SCIMAP is a risk-based modelling framework with a distinct focus on hydrological connectivity (Reaney et al., 2011). The model tries to identify relative importance of areas contributing to a specific problem and was originally developed for diffuse fine sediment (Reaney et al., 2011), and later also used for nutrient pollution (Milledge et al., 2012) and tested for faecal indicator organisms (SCIMAP-FIO; Porter et al., 2017). While taking the processes into account, risk-based approaches do not explicitly model them and do not attempt to predict concentrations (Oliver et al., 2016), hence delivering no quantitative outputs.

In catchment-scale water quality monitoring, process-based models may struggle when they are required to upscale processes that have been studied at field scale (Oliver et al., 2009). Furthermore, they cannot be applied where data are non-existent or too coarse to provide adequate basis for calibration and validation. In this case, successful predictive models can be achieved with empirical modelling (Helliwell et al, 2007; Kay et al., 2005; McGrane et al., 2014; Rothwell et al., 2010; Schoonover & Lockaby, 2006). Such models rely on adequate data that cover potential relevant explanatory variables and conditions (De Brauwere et al., 2014; Tetzlaff et al., 2012). They are particularly suited as screening tools to understand broad-scale drivers, assess source importance and infer spatial controls (Kay et al., 2010; Monteith et al., 2015).

III. Risk screening

Risk assessments need to be appropriately designed to make sure that the outcome adequately informs the decision-making process. Risk screening as a prior step can help to identify the scope by focusing on the most important aspects, allowing efficient allocation of resources (Gormley et al., 2011). Risk screening could have the purpose to identify

- which risks should be included (and which may be neglected),
- areas or features to prioritise,
- if enough data are available or additional evidence is needed for subsequent assessment,
- which method is most suited to the problem and data, and/or
- where immediate action is required.

Process-based hydrological and water quality models are usually complex and highly parametrised which means they need considerable effort and data to set up. Risk screening using simpler quantitative or qualitative methods can achieve a risk ranking that allows identifying where the effort for more detailed assessment and planning is justified (Dunn et al., 2015; Sample et al., 2016). It can also give insight into the nature and degree of uncertainty and thus give indications of the most suitable adaptation and mitigation strategies.

2.3 Catchment protection and management

Functioning ecosystems have been increasingly recognised as crucial to sustain livelihoods, promote sustainable development, and build resilience against climate change (WWAP, 2018). This has led to the concept of Ecosystem-based approaches to Adaptation to climate change (EbA), which uses natural capital to adapt to climate change impacts by preserving and enhancing ecosystems (Munang et al., 2013). Functioning ecosystems provide a range of ecosystem services, which are “ecological processes or functions having monetary or non-monetary value to individuals or society at large” (IPCC, 2014). Ecosystem services are organised into four categories: supporting services (such as sustaining biodiversity), regulating services (such as carbon sequestration or water purification), provisioning services (such as food or timber production), or cultural services (such as tourism, spiritual appreciation etc.). Water utilities rely on two ecosystem services: water provision and water quality regulation, and managing catchments and ecosystems has become of increasing interest to ensure supply of high water quality, with potential co-benefits.

The concept of Integrated Catchment Management (ICM) recognises the catchment as the appropriate organising unit for the management of natural resources (Xenopoulos et al., 2003) which integrates all environmental, economic, and social issues within the catchment into an overall programme or strategy, in order to derive the greatest possible mix of benefits for the communities and future generations while preserving the natural resources on which they rely (Sharp et al., 2006). It is a people-orientated approach and seeks to engage all stakeholders in the catchment through networks and partnerships. EbA combines ICM with targeted restoration to promote the resilient qualities of complex natural ecosystems. It accepts that the future is intrinsically uncertain and that the most effective strategies to reduce risk are measures to improve system resilience, thereby ensuring that ecosystems can deliver their various services long-term. EbA "uses biodiversity and ecosystem services as part of an overall adaptation strategy to help people and communities adapt to the negative effects of climate change at local, national, regional and global level" (Carluer & De Marsily, 2004). It includes the sustainable management, conservation, and restoration of ecosystems to provide services that help people adapt to both current climate variability, and climate change (CBD, 2009).

Using ICM and EbA to safeguard and improve raw water quality for drinking water offers the advantages of targeting the problem at its source, within the catchment, in a balanced and participatory way that has potentially much wider benefits. These wider benefits include for example enhanced recreational possibilities, increased biodiversity, reduced flooding, or increased carbon storage (Martin-Ortega et al., 2014; Grand-Clement et al., 2013; Everard, 2013). Many water companies or local governments have realised the potential benefits of a catchment approach, and catchment initiatives have been increasingly employed over the past decades as a starting point to improve water quality and reduce treatment effort. It is difficult to reliably estimate quantitative outcomes of initiatives due to the complex and interlinked processes within catchments. However, where projects have been successful, it is possible to deduce benefits of such initiatives and sometimes to also put a monetary value on the gains made.

2.3.1 Catchment-based measures to improve water quality

There are several catchment-based measures that have the potential to improve water quality, impacting on a number of parameters looked at in water quality (Table 2.2). Many of these concern land management, especially in agriculture and forestry, while some involve ecosystem restoration to utilise the functions for water purification these ecosystems offer.

Agricultural impacts can be mitigated through land and farm management, such as conservation tillage, targeted and efficient application of fertilizer and manure, keeping livestock out of sensitive areas, or observing buffer strips and zones, and these measures are well documented and have been summarised for policy or guidance documents (e.g. Newell-Price, 2011).

Afforestation may result in reduced pollutant loads if the new woodland replaces degraded agricultural land, through improved soil water infiltration and storage, reducing surface runoff. However, if the soil condition is structurally and hydraulically poor, further degradation may occur, and a change from healthy agricultural land to forest can increase pollutant loads (van Dijk & Keenan, 2007). Under good management, the reduction of surface runoff provided by forests through build-up of litter and undergrowth and increased surface roughness should reduce the volume and character of sediment, nutrients, or salts reaching the receiving water bodies.

Instead of converting whole areas of land to another form of land use, small strips of land along water courses are used to harvest the potential of unused or differently used land to filter pollutants. Riparian buffer strips can have a variety of forms, starting from unfertilized agricultural zones to a zone with different vegetation, often trees, and a wet zone. They trap nutrients, sediment, or pesticides from upslope areas and function by filtering polluted overland and subsurface flows and protecting banks against erosion (Mander et al., 2005). On top of this, they are seen to have wider benefits such as improving habitats and ecological connectivity, stream shading, carbon sequestration and cultural ecosystem services (Stutter, Chardon & Kronvang, 2012). However, the effectiveness of buffer strips is down to complex hydrological and ecological processes, and depends on many parameters, including characteristics of the strip itself (such as width, slope, vegetation height and type, density, or species composition), hydrology and rainfall, and sediment transport (Verstraeten et al., 2006). The retention efficiency of buffer strips is best when polluted water enters it in short events, i.e., during intensive rainfall or intensive thaws), and when buffer zones consist of different plant communities and oil complexes (Mander et al., 2005). It is necessary to place the strip where runoff from agricultural land enters the stream directly (Verstraeten et al., 2006). Buffer strips also have a finite lifespan to retain phosphorus (Stutter, Chardon & Kronvang, 2012) and can start releasing phosphorus into the stream under some conditions, for example when the input is below a certain threshold (Mander et al., 2005). As buffer strips require some land and have, on a landscape scale, lower effectiveness than measures on the field (Verstraeten et al., 2006), they are probably best employed in a catchment approach when used as a final step in a number of mitigation measures.

Wetlands such as fens, swamps, bogs, or salt marshes, are susceptible to changes in the quantity and quality of water and even small hydrological changes can lead to major changes in plant communities (Acreman et al., 2007). The loss or degradation of wetlands leads to loss of functions, and restoring their natural characteristics and function, or constructing new wetlands, is employed to help regain the services they deliver, but restoring wetlands requires a good understanding of the hydrology underpinning them with questions remaining about water availability at required times, the utility of reconnecting river channels to floodplains, and the impact of restoration on the hydrological functions of wetlands (Acreman et al., 2007). It is also important to consider the context of the wetland in the landscape, as it has been suggested that individual restoration projects that are not taking the whole catchment into account can achieve only limited impact on water quality and biodiversity in the landscape (Richardson et al., 2011). Contrary to the restoration of natural wetlands, constructed wetlands have been tools for water treatment since the 1950s. These are often hydrologically disconnected from the surrounding landscape; however, wetlands that are constructed as part of the landscape have the potential to produce not only a water quality benefit, but also provide carbon sequestration, noise screening, or biodiversity benefits at substantially lower cost than treatment (Doody et al., 2009; Everard & McInnes, 2013).

Damaged and degraded peatlands have a number of environmental impacts, including biodiversity loss, declines in water quality, peat erosion and increasing carbon fluxes. Drained peatlands have been shown to release larger concentrations of nitrogen, carbon, phosphorus, sediments, and metals into discharge waters (Wilson et al., 2011). Habitat restoration projects often try to permit vegetation recovery by raising water levels, involving the blocking of drains. While studies are ambiguous on success and speed of recovery of peatland, they show that there is potential for improvement of water quality alongside other ecosystem services gains such as carbon storage, increased biodiversity, and cultural services (Armstrong et al, 2010; Grand-Clement et al., 2013; Martin-Ortega et al., 2014; Wilson et al., 2011).

Reducing runoff from urban areas can be achieved through measures known collectively as Sustainable Urban Drainage Systems, where water is retained or slowed down before reaching surface waters. Management practices to prevent pollution from urban areas to reach water bodies can be structural, such as rainwater retention and pollutant removal systems (e.g., ponds, infiltration trenches, or grassed swales), or non-structural (e.g., street cleaning, minimization of impervious areas, rainwater reuse facilities, green roofs) (Barbosa et al., 2012).

Table 2.2: Summary of usual effects of catchment-based mitigation measures on water quality for selected parameters, their cost to landowners, and other ecosystem services (+ = positive effect, - = negative effect, 0 = neutral).

Measure	**Water quality parameter effected**						**Cost effect**	**Ecosystem Services effected**				**References**
	N	P	Microbial	Metals & chemicals	Turbidity	Colour		Suppor-ting	Provi-sioning	Regula-ting	Cultural	
Measures to increase fertilizer/manure efficiency	+	+	+		+		+					Newell-Price, 2011; Reichenberger et al., 2007; Gooday et al., 2014; Schoumans et al., 2014; Buckley & Carney, 2013; Kay et al., 2009; Randall et al., 2015; Liu et al., 2013
Buffer strips	+	+	+	+	+		-		-			
Soil and crop management, e.g., conservation tillage, cover crops	+	+	+	+	+	+	-/0[1]					
Livestock management, e.g., optimising diets, reducing time in the field	+	+	+		+	+	-/+[2]					
Optimising infrastructure, e.g., position of feeder, watering troughs, gates	+	+	+		+	+	-/0					
Land use changes in agriculture, e.g., from arable to extensive grazing, woodland	+	+		+	+		-	+/0	-		+/0	

[1]There might be an initial outlay to implement the measures, but they are cost neutral in the long run

[2]Keeping livestock indoors may induce costs as well as necessitate good management to prevent pollution to water courses from the animal housing; optimising animal diet can save costs

SUDs	+	+	+	+	+		n/a			+	+	Herngren et al., 2005; Barbosa et al., 2012
Forest management	+	+		+			n/a	+			+	Van Dijk & Keenan, 2007; Calder, 2007; Nisbet et al., 2011
Vegetated riparian buffer strips	+	+/-[3]	+	+	+	+	n/a	+		+	+	Tang et al., 2012; Broadmeadow & Nisbet, 2004; Stutter, Chardon & Kronvang, 2012; Mander et al., 2005; Verstraeten et al., 2006; Anbumozhi et al., 2005
Wetland restoration	+	+	+	+	+		-/+[4]	+	0/-[5]	+	+	Acreman et al., 2007; Fisher & Acreman, 2004; Richardson et al., 2011; Wilkinson et al., 2014; Grand-Clement et al., 2013; Martin-Ortega et al., 2014; Moxey & Moran, 2014; Holden, 2005; Wallage et al., 2006; Armstrong et al., 2010

[3]Riparian buffer strips need to be managed for P retention and can release P if they have reached their retention capacity or when P input is reduced

[4]Although there are costs associated with restoration, research suggest that benefits may outweigh the restoration costs

[5]The effect on provisioning services will depend on what use the land was put to before restoration

2.3.2 Examples from the drinking water sector

Taking a catchment approach to improve or stabilize raw water quality is increasingly recognised in the drinking water industry to secure water supply in the long-term and save costs in water treatment. The number of catchment projects in the drinking water sector has increased and while it is difficult to assess their effectiveness e.g., in terms of how much costs are being saved in comparison to traditional approaches, or in terms of long-term benefits, it is possible to draw general conclusions from the outcomes and lessons learned.

A key example for a catchment approach, and the economic benefits associated with it, is the New York City Watershed Agricultural Program. The drinking water in New York comes from the Catskill, Delaware and Croton systems. Based on the quality of the water supply, New York City has a temporary exemption, currently running until 2027, to use unfiltered water from the mostly forested Catskill-Delaware system that provides 450 million litres of water daily to the population of New York City, making up 88% of the supply (NYC DEP, 2022). In the 1980s, the water quality began to decline due to pressures from intensified farming and forestry, and holiday home development. In order to preserve the area and to ensure that the water coming out of these catchments for the city stayed of high quality, the New York City Watershed Memorandum of Agreement was signed in 1997, setting out a programme for land acquisition by New York City as well as regulations and provisions for activities within the watershed. They address mainly wastewater treatment, septic systems, and storm water pollution (Willett & Porter, 2001). The measures are aimed at conserving the natural forest and reducing or eliminating commercial activities that have deteriorating effect on land and water. In addition, agricultural activities are controlled through the Watershed Agricultural Council, who implement the Whole Farm Planning. In this voluntary approach, individual farms and their impact on water quality are assessed and a plan to reduce the impact is developed. The Watershed Agricultural Council is farmer-run and financed by New York City and achieved a high participatory rate of 93% within five years (Appleton, 2002). Turbidity and faecal coliform standards are being continuously met at a reservoir at the endpoint of the system, and research at another reservoir indicated that phosphorus loads decreased by 50% and 17% for dissolved and total phosphorus respectively (Kousky, 2015). In 1990, the city estimated that putting in a filtering system would cost US$ 4 to 8 billion, while today it is estimated to cost US$ 8 to 12 billion, with annual running costs of US$350 million. This would be double the costs of water rates in the city, and substantially more than costs of US$ 1.5 billion for the measures under the Memorandum up to 2010 (Kousky, 2015).

Benefits to farmers have been named as increased fertilizer efficiency, saving of farmer time, and increased herd health. Participating in the programme also means they are exempt from New York City's regulation, and it helps them to meet federal regulations, while preserving their autonomy (Kousky, 2015).

Several private water suppliers in England have started catchment projects, dealing with similar but slightly different issues and approaches. United Utilities in the Northwest of England faced the problem of water colouring due to peatland degradation as well as to pollution from agricultural activity (United Utilities, 2022). Their Sustainable Catchment Management Programme (SCaMP) began in 2005 with the aim to improve water quality while simultaneously benefit wildlife. The first phase of the project ran until 2010 and realised moorland restoration, woodland management, farm infrastructure improvements and watercourse protection. SCaMP 2 ran from 2010 to 2015 and enlarged the area of intervention and focused on changing farming practices. SCaMP 3 then ran from 2015 to 2020 and focused on high priority areas in terms of increasing issues around water quality and potential for successful intervention and established drinking water safeguard zones. Measures carried out under SCaMP included restoration of blanket bog by blocking drainage ditches and gullies, restoration of eroded and exposed peat, restoration of hay meadows, new woodlands, scrub planting, restoration of heather moorland, improving livestock housing, new waste management facilities, livestock fencing, and assistance to farmers to enter Higher Level Stewardship schemes. The scheme also includes a monitoring programme. Overall, the effects are seen as beneficial with significant effects on water quality, e.g., colour production is stable or slightly decreasing over the monitored period for most sites (United Utilities, 2022). Estimates for the value created through the restoration efforts have been put at £152 per hectare and year for unmanaged moorland habitat, and £180 per hectare and year for new woodland (Morris & Holstead, 2013).

South West Waters entered into a partnership with the Devon Wildlife Trust, the Cornwall Wildlife Trust, the Westcountry Rivers Trust and the Exmoor National Park Authority to implement the Upstream Thinking project. The project started in 2010 and focused on two main elements: changing farming practices through advice and grants to farmers, and peatland restoration. Grants for farmers can be used to improve slurry storage, livestock fencing, alternative water sources, or pesticide management including new equipment. Within the programme, over 2780 ha of peatland were restored, and research carried out by the University of Exeter showed that the restored bogs release a third less water during storm events in shallower peats and two thirds in deep peat, and only about one third of dissolved organic carbon compared to pre-restoration (SWW, n. d.). South West Water envisaged a spending of £9.1 million, but estimate the benefit:cost ratio at 65:1 over 30 years (Morris & Holstead, 2013).

Yorkshire Water focuses on restoration of peatland, as upland, peatland dominated catchments are the main drinking water catchments. This includes blocking of gullies and grips, seeding with nurse crops such as grasses and covering with heather brash, and has been carried out on 3250 ha of degraded peat. Yorkshire Water has also been part of ancient woodland restoration projects. As a result, habitat has improved for a range of bird, fish, reptile, plant, and dragonfly species, as well as for the red kite which have been released and produced 430 young birds in the area (Morris & Holstead, 2013).

Wessex Water work with farmers and land managers through catchment advisors and financial assistance for farmers. They estimate that this approach costs them one sixth of a more traditional engineering approach (OFWAT, n. d.).

The mineral water brand Vittel belongs to Nestlé Waters and its water stems from a spring in the French Vosges mountains. Towards the end of the 1980s, the water quality deteriorated due to nitrate input from agricultural practices in the Vittel catchment. This posed a threat to the company as it might have lost the right to market the water as "natural mineral water" and suitable for infant feeding. Vittel in consequence increased its ownership of land in the catchment to 45%, mainly by buying land from retiring farmers. However, buying more land was not possible and not practical, as the French law made it necessary to keep the land in agricultural use and the right to choose farming practices would be with any tenant of Vittel. The company therefore started to negotiate with farmers and to form individual contracts that obliged farmers to employ specific agricultural practices in return for financial compensation, financing of equipment, free tenancy of the Vittel owned land, and free technical advice. The contracts ran over a duration of 18 to 30 years (Depres et al., 2008). The situation was initially difficult as there was a lack of trust between the company and the farmers that manifested itself for example in attempts to value the change of farming practices. This was mainly overcome by Vittel contracting a research team to advise on the kind of measures to be taken to achieve the desired outcome. The research team played a mediating role, gathering data about each farm and developed feasible solutions together with the farmers, integrating their concerns (Depres et al., 2008). Vittel achieved to contract with 92% of the farmers, covering 96% of the targeted area. The main mitigation measures included the elimination of corn crops on their farms, a ban on pesticides, compost of all animal waste and fertilization through composted manure, limitation of livestock density and farmyard management (Depres et al., 2008). Although the whole process, including purchase of land, designing the contract process, compensation to farmers and monitoring of compliance, cost Vittel more than €24 million over seven years, this is seen as lower than the potential damage would have been if Vittel had lost the right to market the water from this source as natural mineral water (Depres et al., 2008). Nestlé Waters also adopted this approach subsequently in other companies, so a learning benefit was generated.

These projects had the explicit goal of achieving better raw water quality for drinking water purposes and save costs on treatment, thereby making the supply of clean water meeting the statutory requirements cheaper and more secure. They show positive outcomes not only for water quality but also for a range of other ecosystem services and stakeholders (Table 2.3). For some of the projects, economic benefits have been observed or estimated, such as the anticipated a benefit to cost ratio of 65:1 from the Upstream Thinking project through savings on treatment costs, and the New York City example also demonstrates significant savings in treatment costs. Stakeholders other than water companies also often benefit, either by saving through the implemented practices or by being compensated for a reduction in income. Other economic benefits can also arise from enhanced ecosystem services such as fisheries, tourism, or recreation. Apart from cost savings and new economic opportunities, the projects often have wider, and non-economic, benefits. These include for example increased biodiversity or enhanced carbon storage, increased landscape aesthetics or reduced flooding.

Table 2.3: Overview of described projects specifically designed for improving raw water for drinking water purposes.

Project	**Measures taken**			**Named benefits**	
	Agricultural practices/ Farm management	*Afforestation/ forest management*	*Habitat restoration*	*To water company*	*To stakeholders*
NYC (USA)	✓	✓		Avoidance of estimated costs of filtering system: US$ 4 billion + and running costs of US$ 350million/a, versus spent US$ 1.5 billion	Tailored best practices plans for farms, leading e.g., to savings in e.g., fertilizer costs, reduced time spent Creation of recreational areas
SCaMP (UK)	✓	✓	✓	Reduced and stabilized levels of DOC in raw water	Increased suitability of habitats for biodiversity development (improving the quality of SSSI) Possible positive impact on flood attenuation
Upstream Thinking (UK)	✓		✓	Anticipated benefit to cost ratio 65:1	£20,000 cost savings on average per farmer
Yorkshire Water Strategy (UK)		✓	✓		Improved SSSI sites and enhanced biodiversity
Wessex Water Catchment Management (UK)	✓			Estimated cost reduction of 83%	
Vittel (France)	✓			Avoidance of financial losses and reputational damage by retaining the right to market water as ‘natural mineral water’	Enlargement of farms through Vittel land Compensation payments Free technical advice

2.3.3 Catchment management and climate change mitigation and adaptation

As the reality of climate change has become increasingly clearer over the past years, pledges to reduce greenhouse gas emissions and moving towards “Net Zero” have increased on international, national, and local level (Deutch, 2020), as have initiatives to mitigate and adapt to climate change. Measures are often distinguished as either engineered (‘grey’) or nature-based (‘green’) solutions (Seddon, Daniels et al., 2020). Nature-based solutions (NbS) work with natural systems to address societal challenges such as mitigating and adapting to climate change, and include actions that protect, restore, and manage ecosystems (Chausson et al., 2020). These measures are likely to have a range of other effects on ecosystem services, among them water quality regulation (Seddon, Chausson et al., 2020). While many water utilities as well as policy makers have recognised the potential for catchment-based approaches to increase water quality and have started initiatives as described above, practical examples of using a catchment approach under the umbrella of EbA or NbS to specifically increase ecosystem resilience to climate change for water quality purposes are rare, or not well documented or accessible, and its effectiveness is difficult to assess due to the future dimension those projects naturally incorporate.

Generally, the evidence base for the effectiveness of EbA is weak and scattered, with much of it grey literature or anecdotal evidence (Doswald et al., 2014; Reid et al., 2018). Doswald et al. (2014) and Chausson et al. (2020) used systematic mapping protocols to catalogue available evidence of the effectiveness of EbA or NbS. Both reviews included studies that varied widely in their primary focus and the observed and reported effects (climate impacts, GHG mitigation, economic, social, and ecological effects, multiple benefits, and comparison to alternative approaches). They found that most studies showed positive effects towards reducing climate change impacts, and more reported multiple benefits rather than trade-offs, although there may be a bias due to non-reporting of negative outcomes (Doswald et al, 2014; Chausson et al., 2020). In a more specific review on the wider benefits of natural flood management, Iacob et al. (2014) also reported that the majority of included studies found net positive results. Where the comparison was made, most studies concluded nature-based solutions to be as effective, or more, than alternative approaches. However, the authors also pointed out that most studies failed to make such as comparison or to examine broader societal, ecological, and economic effects (Chausson et al., 2020). This may be because these effects are hard to quantify (Doswald et al., 2014; Reid et al., 2018).

Water quality regulation seems to be among the ecosystem services that are ambiguously affected depending on the measure taken, with positive and negative effects taking place, while most negative effects were connected to water availability (Iacob et al., 2014; Chausson et al., 2020). This highlights that measures with a primary focus on water quality need to be implemented carefully but carry the potential for multiple benefits.

2.4 Drinking water in Scotland

There are more than 125,000 km of river and over 25,500 lakes, or lochs, in Scotland (Critchlow-Watton et al., 2014). These vary greatly in size and natural conditions and hence have variations in water quality. Although water is overall abundant in Scotland, multiple uses compete over water resources. While most rivers and lochs are in “good” condition under the Water Framework Directive (WFD; see 2.4.1), anthropogenic influences degrade water sources, including for drinking water purposes. Many of the lochs that have degraded water quality are affected by land management practices or acidification, and rivers are affected by agriculture, hydropower schemes and urbanisation (Critchlow-Watton et al., 2014).

2.4.1 Regulatory framework

While Scotland as part of the UK is no longer a member of the European Union, acts and regulations that were created to comply with European legislation are still in force. The drinking water quality directive of the European Union (Directive (EU) 2020/2184 of the European Parliament and of the Council of 16 December 2020 on the quality of water intended for human consumption (recast)) sets minimum standards for drinking water quality that need to be implemented in the member states. The directive obliges all member states to take all necessary actions to ensure that water intended for human consumption (water used for drinking, cooking, food preparation and other domestic purposes as well as water used in any food production) is ‘wholesome and clean’, meaning ‘free from any micro-organisms and parasites and from any substances which, in numbers or concentrations, constitute a potential danger to human health' and meeting the standards set for microbial and chemical parameters in Annex I of the directive. In Scotland, the original drinking water quality directive was implemented through the Public Water Supplies (Scotland) Regulations 2014, and the Water Intended for Human Consumption (Private Supplies) (Scotland) Regulations 2017 for larger private supplies. Smaller private supplies (serving only domestic premises and less than 50 persons in total) are regulated through the Private Water Supplies (Scotland) Regulations 2006. The Cryptosporidium (Scotland) Directions 2003 specifies how the supplier must respond to the risk from the *Cryptosporidium* parasite as well as sampling requirements.

Further legal documents that constitute the legal framework for drinking water in Scotland are the Water (Scotland) Act 1980 and the Water Industry (Scotland) Act 2002. The Water (Scotland) Act 1980 established the duty of the water authorities to supply wholesome water. The Water Industry (Scotland) Act 2002 created the Water Industry Commission, the Drinking Water Quality Regulator (DWQR) and Scottish Water. Scottish Water is the public water supplier and replaced the previous Water and Sewerage Authorities. The Water Industry Commission for Scotland is the economic regulator for the Scottish water industry, and as such sets charges to customers, facilitates competition and monitors and reports on economic performance. The DWQR is responsible for ensuring that the duties of the public supplier are complied with and supervises the enforcement by local authorities over private suppliers. To this end, the DWQR is given powers to obtain information and power of entry, inspection, and enforcement, as well as emergency powers to force the supplier to carry out work to make the water safe for human consumption.

Other actors involved in drinking water regulation are the Scottish Environment Protection Agency (SEPA) as being responsible for environmental protection. SEPA is the regulator for Scottish Water regarding any of the environmental aspects, including for example water abstraction and wastewater treatment and discharge. Next to the economic regulator, customers are also represented through the customer forum and consumer futures, through which customer views and interests are identified and voiced. The Scottish Public Services Ombudsman is responsible for investing complaints about public services, including Scottish Water (Figure 2.7).

Figure 2.7: Actors involved in the regulation of the Scottish public drinking water supply (reproduced from Scottish Water, 2022).

Apart from direct legislation for drinking water, raw water quality (the water going into the treatment works), is influenced by water quality legislation for other purposes. Several pieces of legislation made in the EU directly aim at improving overall water quality, others set quality standards for specific areas of water use. The former includes the Water Framework Directive, which is an overarching piece of complex legislation aiming at improving water quality for different uses (e.g., aquatic ecology, valuable habitats, drinking water and bathing water), using an integrated approach with the catchment, or river basin, as the basic unit. Surface water bodies are required to reach “good ecological status” as well as “good chemical status” as set out in the directive, with more stringent requirements for specific protection zones designated for other uses (such as drinking water and bathing). Groundwater bodies should be protected from anthropogenic influences by a prohibition of direct discharges and through monitoring to enable reversal of any deterioration. Within the river basins, the status of water bodies is first determined before the effectiveness of measures under other existing legislation is assessed, and further actions decided if deemed necessary to achieve the objective of good status. The analyses as well as measures are laid out in detailed river basin management plans, which are developed including information and consultation of the public. The WFD is implemented in Scotland through the Water Environment and Water Services (Scotland) Act (WEWS) 2003, supplemented by the regulations for river basin management planning (The Water Environment (River Basin Management Planning: Further Provision) (Scotland) Regulations 2013, and The Cross-Border River Basin Districts (Scotland) Directions 2014) and the Water Environment (Controlled Activities) (Scotland) Regulations 2011. WEWS establishes the process to create River Basin Management Plans, putting SEPA in charge of producing and implementing the plans. Scotland has two major river basins, Scotland and the Solway-Tweed (including parts of Northern England), which are divided into eight and two sub-basins respectively. The status of the water bodies identified in these catchments is determined, standards are set for the concentration of pollutants, the level of flows in rivers, from lochs, and the physical structure to determine the status, which can be high, good, moderate, poor, or bad. The current standards are laid out in the Scotland River Basin District (Standards) Directions 2014 and the Solway Tweed River Basin District (Standards) (Scotland) Directions 2014.

Related legislation aiming to reduce certain types of pollution are the nitrates directive (Council Directive 91/676/EEC of 12 December 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources), the priority substances directive (Directive 2013/39/EU of the European Parliament and of the Council of 12 August 2013 amending Directives 2000/60/EC and 2008/105/EC as regards priority substances in the field of water policy) and the urban waste water treatment directive (Council Directive of 21 May 1991 concerning urban waste water treatment (91/271/EEC)). The aim of the nitrate directive is to reduce and prevent nitrogen pollution of waters from agricultural sources. To this end, member states are obliged to identify waters that are affected by pollution, and to designate as nitrate vulnerable zones (NVZ) such areas that contribute to the pollution of these waters. Member states must further establish a code of good agricultural practice to be voluntarily implemented by farmers, and action programmes for the nitrate vulnerable zones.

In Scotland, the Action Programme for Nitrate Vulnerable Zones (Scotland) Regulations 2008 apply to farms in NVZs and set out rules for the management of fertilizer and manure, e.g., each farm has to have a fertilizer and manure management plan, rules about storing manure or the application of nitrogen fertilizer. There is also a code of good agricultural practice (Prevention of Environmental Pollution from Agricultural Activity), setting out measures to reduce pollution and actions that are mandatory for farmers in NVZs, actions that are required to receive Single Farm Payment (now replaced by the schemes under the Scottish Rural Development Programme) and actions that are voluntary. Apart from trying to reduce nitrates through the nitrate directive, the EU targets a list of substances called priority substances through the priority substances directive, which sets environmental quality standards for 33 priority substances. Priority substances include for example hazardous and toxic compounds or metals. Special measures are to be taken by member states to reduce pollution from these priority substances. The directive was implemented in Scottish law through amendments to WEWS. The Council Directive of 21 May 1991 concerning urban wastewater treatment (91/271/EEC) specifies minimum standards for the treatment of municipal wastewater, with the objective to protect the environment from adverse effect through these waste waters. The Urban Waste Water Treatment (Scotland) Regulations 1994 implement this directive in Scotland.

The bathing waters directive (Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 concerning the management of bathing water quality) and the shellfish directive (Directive 2006/113/EC of the European Parliament and of the council of 12 December 2006 on the quality required of shellfish waters) set standards for specifically designated water bodies. The bathing water directive regulates the monitoring and classification of bathing waters, the management of bathing water quality, and the provision of information on bathing water quality to the public. The directive is implemented in Scotland through the Bathing Waters (Scotland) Regulations 2008, which lays out the duties of SEPA with regard to monitoring and safeguarding the water quality in bathing waters. The shellfish directive is designed to protect aquatic habitat of bivalve and gastropod molluscs (e.g., oysters, mussels, cockles, scallops, clams). Member states are to designate shellfish waters which need to comply with the standards set for water quality. The directive is implemented in Scotland by the Surface Waters (Shellfish) (Classification) (Scotland) Regulations 1997, together with the Surface Waters (Shellfish) (Classification) (Scotland) Direction 2012. The direction sets out the latest waters classified as shellfish waters. Efforts to improve the water quality in bathing and shellfish waters to comply with the directives will also benefit raw water quality for drinking water, and vice versa; this can for example be observed by the fact that changes in the catchment e.g. through the priority catchments rural diffuse pollution programme (see 2.4.5) led to improvements in bathing water quality. Similarly, investments made by Scottish Water in water and drainage infrastructure under the Quality & Standards programme have shown a positive effect on bathing water quality (SEPA, 2015).

2.4.2 Drinking water supply and quality

Presently in Scotland, 2.57 million households and over 150 000 businesses are provided with water from the public supply network. Scottish Water provides 1.53 billion litres of water daily and operates 237 water treatment works (Scottish Water, 2021). Sources of water supply are reservoirs, lakes, rivers, or groundwater. Apart from the type of source, the characteristics of the catchment influence the quality of the water entering the treatment works. Water of higher quality, as often found from springs and boreholes, may need only simple filtering followed by disinfection, whereas lowland surface water sources will normally require some more extensive treatment as it contains more pollutants. The type of treatment is determined by the raw water quality data (Scottish Water, 2020). Water quality of the public supplier is tested at the water treatment works, within the distribution system and at consumer’s taps. Testing at customer’s taps is for 51 parameters and compliance is generally very high; in 2020, 99.95% of samples tested at the consumer’s tap adhered to standards (DWQR, 2021a).

Although most people in Scotland are served by the public supplier, around 180,000 people (3.3% of the population) rely on private water supplies for their drinking water (DWQR, 2021b). These vary in size and can serve from one household to thousands of people, and the quality of water from private supplies can also vary considerably. Sources of private water supplies are rivers and streams, private reservoirs, wells, boreholes, and springs.

Private water supplies must be sampled by the local authorities either every year (for bigger supplies with 50 or more people or more than 10 m^3 of water) or within 28 days upon request by the owner or user (for smaller supplies). The bigger, or "regulated" private supplies must also be risk assessed every five years. In 2020, only 38% of the regulated supplies were sampled, and of these, 9% did not meet the standard. Of the sampled smaller supplies, 13.3% failed. *E. coli* was detected in 12.3% of the tested regulated supply samples (DWQR, 2021b).

2.4.3 Raw water quality

Water quality is influenced by the pathways water takes through landscapes and hence characteristics including soil, rock and vegetation. Ecologically, the water that corresponds to the naturally occurring ecosystem would constitute "good" water quality – normally this means a progression of increasing trophic status and hence changing ecosystems further downstream (Ferrier et al., 2001). This may however not correspond to "good" water quality defined for drinking water through standards.

I. Natural conditions

The retreat of the ice formations after the last glacial period heavily impacted on the topography of Scotland. The highlands of Scotland show many over-steep slopes, over-deepened valleys, and complex suits of glacial and periglacial drift deposits (Soulsby et al., 2002). The east coast by contrast is mainly characterized by gentle hills. Scotland has a maritime climate with major air circulation from West to East, so that the altitudes of the highlands provide a barrier to the westerly airflow from the Atlantic, leading to large precipitation in the West of Scotland (over 4000 mm/year in some of the highest western areas) to less excessive rainfall in the East (as low as 600 mm/year). Precipitation occurs throughout the year with peaks in winter. In high areas, up to 30% of precipitation occurs as snow. Actual evaporation rates are typically between 350 and 400 mm/year, and evapotranspiration can be high in the summer months, depleting soil water stores and reducing river flows (Johnson & Thompson, 2002). Temperature differences are most extreme in the Cairngorm Mountains.

In geological terms, Scotland can be divided into five distinct parts that show a Southwest to Northeast trend and that are separated by faults (the Moine Thrust, the great Glen Fault, the Highland Boundary Fault and the Southern Upland fault), many of which are still recognisable in the landscape (Figure 2.8). Starting from the Northwest, the first three are commonly grouped together as the Highlands which are dominated by igneous and metamorphosed igneous and sedimentary rocks. Further to the Southeast are the Midland Valleys with mainly non-metamorphosed sediments and lastly the Southern Uplands with predominantly weakly metamorphosed sedimentary rocks (Langan & Soulsby, 2001; Gillespie et al., 2013). Water can move through the pores of sedimentary rock or through fractures, whereas in igneous and metamorphosed rock water can only flow through connected open fractures (Gillespie et al., 2013).

Solutes in water are derived from precipitation or dry fallout from the atmosphere, or from the passage of water trough biomass, soils, and rock (Trudgill, 1986). The topography, together with the geological and climatic conditions of Scotland provide the basis for the distribution of soils and for hydrological regimes, which underpin the water quality characteristics of Scottish surface waters. As soils act as a reservoir for rainfall and a channel for transferring rainfall to surface waters, as well as a buffer against atmospheric input and a supplier of nutrients and medium for plant growth, soils are a crucial influence on water quality. Soils in Scotland are mostly highly organic (Figure 2.9) and acidic (Figure 2.10), reflecting the underlying acidic geology (Langan & Soulsby, 2001).

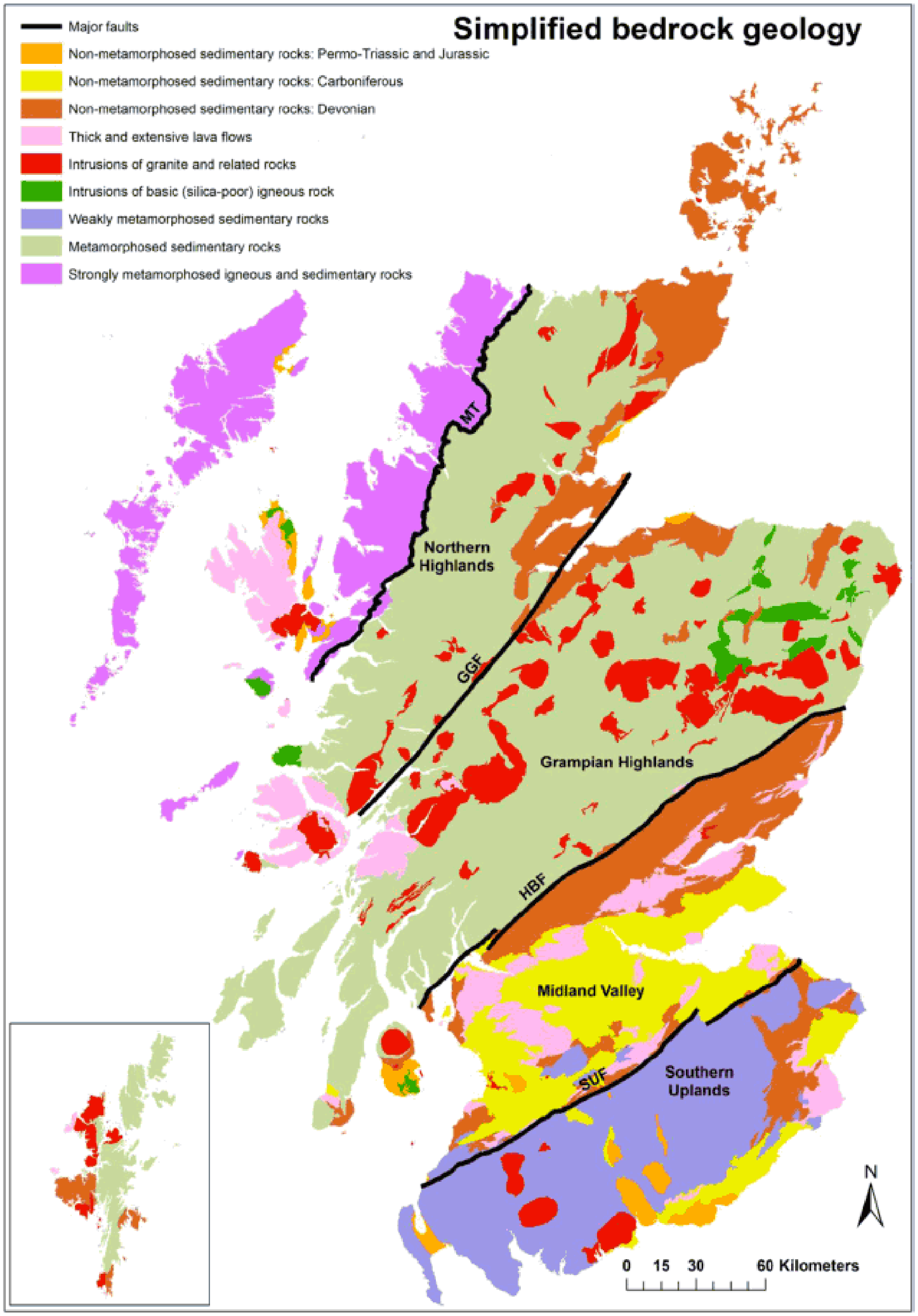


Figure 2.8: Scotland's bedrock geology (reproduced under the Open Government Licence v3.0 (http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/) from Gillespie et al., 2013).

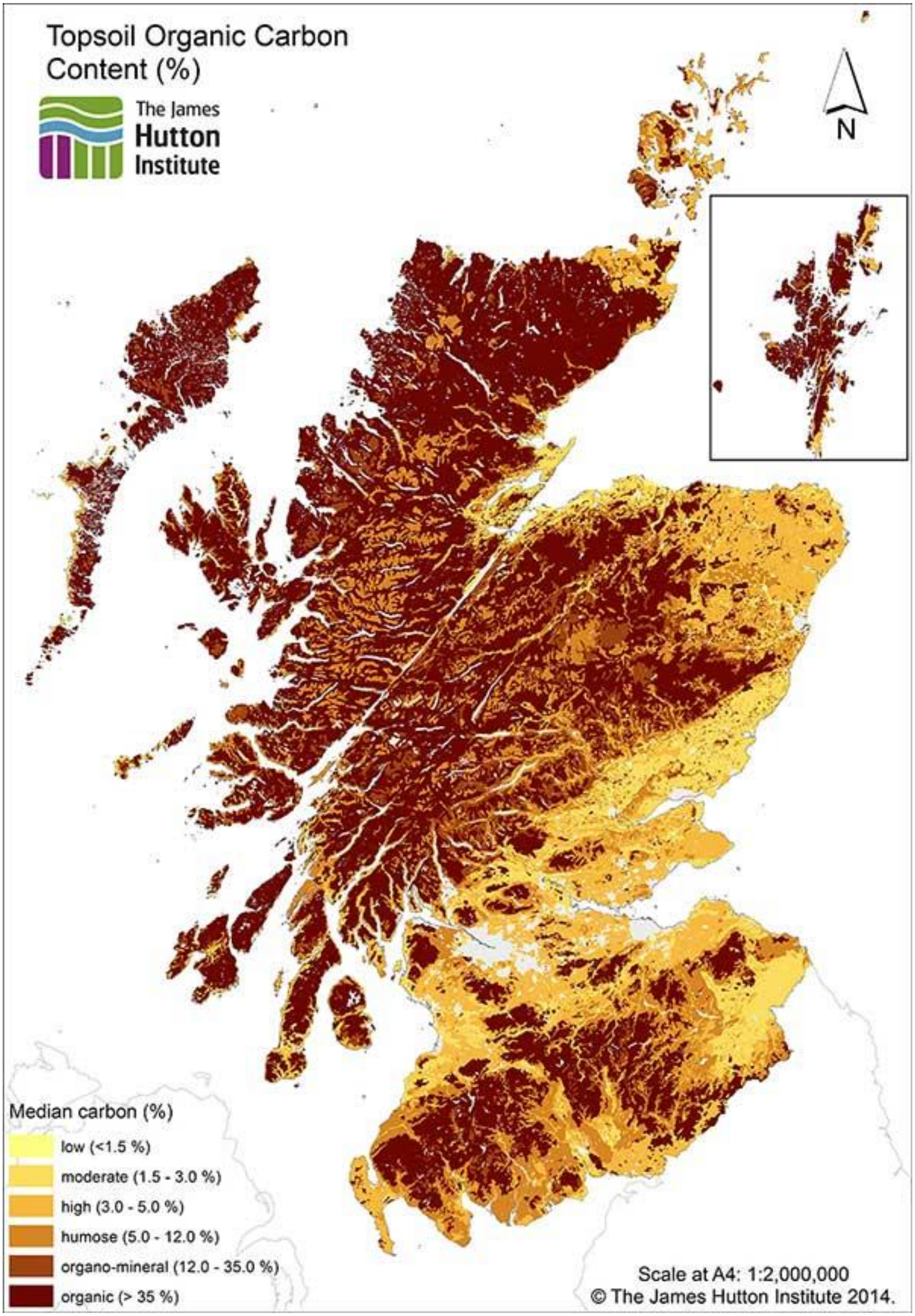


Figure 2.9: Organic carbon content in Scotland's topsoil. Reproduced from Lilly et al., 2012.

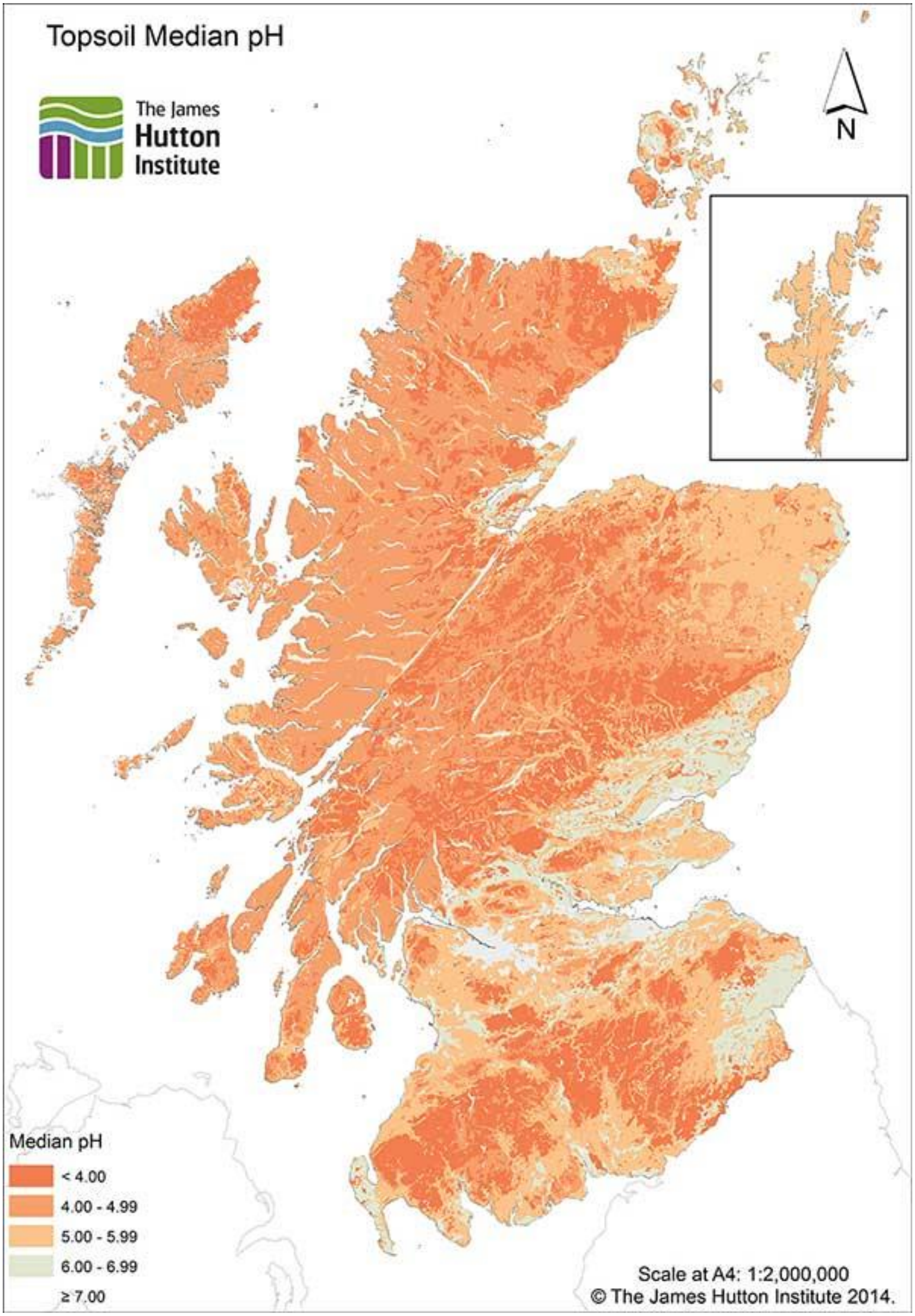


Figure 2.10: Median pH of Scotland's topsoil. Reproduced from http://ukso.org/static-maps/soils-of-scotland.html [03/12/2022].

The most common types of soil are peat, gley soils, brown forest soils, and podzols. These four types of soils account for almost 90% of Scotland's soils (Scotlandguides, 2020). On exposed mountain tops, soils are usually very shallow as soil formation is slow. These soils are very susceptible to pressures and are unlikely to recover naturally when eroded or polluted. Peats are a deep layer of organic matter when vegetation decays slowly and accumulates due to waterlogged conditions. Peat soils hold a substantial amount of all of Scotland's soil carbon. Gley soils are also poorly drained but do not accumulate organic matter. Podzols are well drained but develop on acid conditions where nutrients get washed out, in contrast to brown forest soils which are well drained and fertile, meaning they are favorable for agriculture. Where these occur are the principal areas in Scotland where arable agriculture is possible.

Soil characteristics, along with climatic and topographic conditions, form the basis for assessing viable land uses for regions, called land capability assessment. In Scotland, the land capability assessment for agriculture distinguishes seven major classes (Bibby et al., 1991), ranging from land class (LC) 1 as capable of producing high yields of a wide variety of crops, to LC 7, with extremely severe climate restriction leading to very limited use for agriculture. Generally, LCs 1-3.1 are defined as 'prime', although land classed as 3.2 is also capable of sustaining the production of crop (Table 2.4). Scotland has only few areas of prime land, mainly along the East coast, and along the Central belt (Figure 2.11).

Table 2.4: Description of LC classes for Scotland, adapted from Brown et al. (2008).

Class	Category	Climate limitations	Land use
1	Prime	Non or very minor	Very wide range of crops with consistently high yields
2	Prime	Minor	Wide range of crops, except those harvested in winter
3.1	Prime	Moderate	Moderate range of crops, with good yields for some (cereals and grass) and moderate yields for others (potatoes, field beans, other vegetables)
3.2	Non-prime	Moderate	Moderate range of crops, with average production, but potentially high yields of barley, oats, and grass
4.1	Non-prime	Moderate-severe	Narrow range of crops, especially grass, due to high yields but harvesting may be restricted
4.1	Non-prime	Moderate-severe	Narrow range of crops, especially grass, due to high yields but harvesting may be severely restricted
5	Non-prime	Severe	Improved grassland, with mechanical intervention possible to allow seeding, rotavation or ploughing
6	Non-prime	Very severe	Rough grazing pasture only
7	Non-prime	Extremely severe	Very limited agricultural value

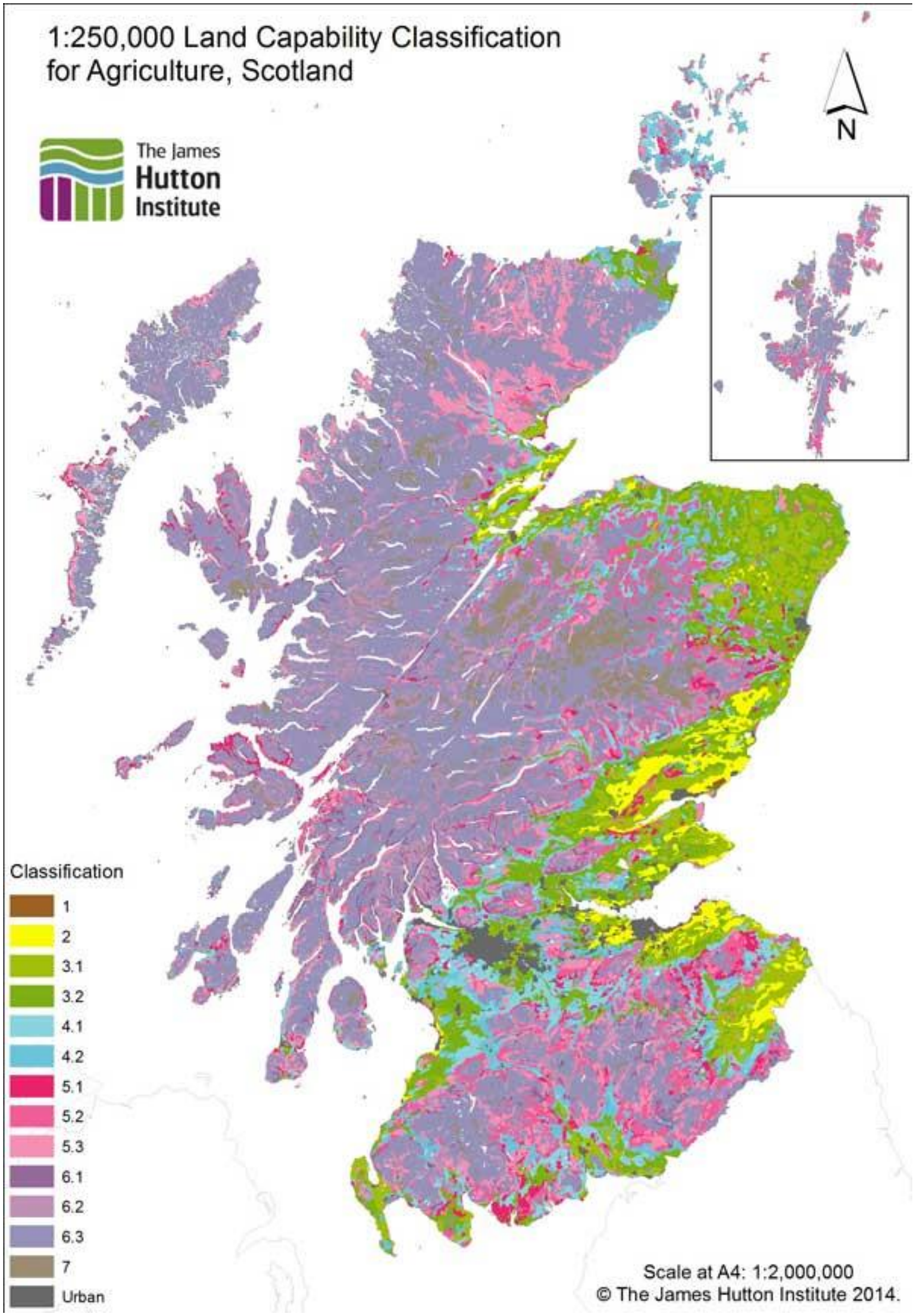


Figure 2.11: Land capability classification for agriculture of Scotland. Reproduced from http://ukso.org/static-maps/soils-of-scotland.html [last accessed 03/12/2022].

Lowland and upland areas present a marked difference in hydrology, induced by the differences in topography and climate. Upland areas with steep slopes and high precipitation, and mainly covered with peat or podzolic soils that have limited storage ability, are usually highly responsive to rainfall events and show rapid runoff responses. Baseflows are normally low, although groundwater aquifers can exert a strong influence on stream flows in some areas. Snowmelt in spring can contribute to major flood events, next to intense summer rainfall events or prolonged rainfall on already saturated ground. Lowland areas with less marked reliefs and sedimentary underlying rocks or extensive drift deposits, have less flashy runoff regimes and a stronger contribution of groundwater to stream flow (Langan & Soulsby, 2001).

The upland areas of Scotland provide the headwater sources for the major river systems and so are key to quantity and quality issues essential to sustainable management of water resources. Maintaining high water quality of the upland water resources is hence essential to the sustainable management of water resources. The flashy nature of upland waters means that water quality can be very variable over time, mainly because of the different pathways the water takes during storm events and during low flow periods, because of dilution or concentration effects in storm events, and because of the importance of groundwater sources to the baseflow. Generally, water from upland areas is oligotrophic and nutrient concentration is low. River systems that drain the main areas of granite are acid sensitive and have a low alkalinity (Soulsby et al., 2002). Average alkalinity, pH and base ion concentrations are highest in summer baseflows and lowest in winter storm events, but there are marked variations in seasons between the years e.g., through wet summers and freezing conditions in winter when winter baseflows occur. Atmospherically derived chloride and sulphates tend to be highest in concentration in the winter months. Nitrate also tends to peak in winter reflecting the low biotic uptake. Many parts of the Scottish highlands are acid sensitive and high in precipitation, leading to high levels of atmospheric deposition that outstrip the buffering capacities of catchments (Soulsby et al., 2002), although there have been signs of recovery over more recent years (Battarbee et al., 2014). Although many upland areas are N-limited and hence nitrogen tends to be sequestered by the ecosystems, there are indications that nitrate losses to surface waters are occurring (Chapman et al., 2001). While this is the general picture, it is important to remember that individual catchment hydrology is complex – some catchments for example are much more groundwater dominated, which means they deviate from this flashy runoff regime rule, while in others metamorphosed calcareous rocks are present, so that alkalinity is much higher.

Groundwater quality varies in Scotland, depending on the age of the water. Groundwater ranges from young, weakly mineralized water to older and moderately mineralized water. Some waters derive from rainfall that predates the industrial and agricultural revolutions and show near pristine quality, others are modern in origin and rich in nutrients and organic chemicals (Robins, 2002). Groundwater recharge potential is highest in the west of Scotland where rainfall is high, however recharge is inhibited by steep topography, widespread till cover and inadequate transmissive properties of older basement rocks. Most shallow groundwaters are actively recharged and relatively young. Discharge as baseflow to surface waters can influence the quality of low stream flows, e.g., by buffering acidic water. Most Scottish groundwater resources are shallow and unconfined, so they are generally vulnerable to point source and diffuse surface pollution. The most common contaminant in Scottish groundwater is nitrate, corresponding to the areal extend of agricultural land uses (MacDonald et al., 2017).

II. Anthropogenic pressures

Land use in Scotland has changed throughout history, the major change being the loss of woodland which was replaced by moorland, grassland, and pasture (Edwards & Ralston, 1997). Land use in Scotland is characterised by a split between upland and lowland areas, with lowland areas being predominantly used for arable cultivations or improved grassland. Upland areas are mainly maintained as heather moorland, reflecting the acidic soils which are favourable for the main heather species *Calluna vulgaris*. These landscapes are very important from a touristic perspective, as being often perceived as typically Scottish. In wetter conditions, heather moorland is replaced by peatland communities, and along the river valleys by forest and woodland, especially on steeper slopes where the soil is better drained. Forests are predominantly monoculture plantations with exotic species, although the use of indigenous species, and of afforestation with semi-natural or natural woodland is rising. Grasslands tend to be unimproved at more elevated areas and increasingly improved moving to lower elevations. Urban areas are concentrated around four major hotspots

(Glasgow, Edinburgh, Dundee, and Aberdeen). Population density itself is low compared to other European countries.

As upland areas play a crucial role as drinking water sources, the management of upland areas is also a key factor in drinking water quality. Heather moorland is mainly managed for sheep and grouse/deer shooting. Large numbers of deer maintain the moorland character and prevent tree cover. High grazing pressures by deer also often lead to compacted tracks which cause erosion problems and sediment delivery into streams. Deer grazing in the riparian area of streams causes concerns over microbiological pollution (Soulsby et al., 2002). For grouse shooting, heather moorland burning is often practiced to encourage a mosaic pattern. When natural features such as streams are used to restrict fires, riparian areas are burned which increases erosion. Together with moorland drainage, burning changes the hydrology of the land by reducing the runoff response time and vegetation removal and thus induces erosion (Bragg, 2002), leading to increased sediment and total organic carbon input (causing colouration of water). However, it needs to be pointed out that the effect of burning on dissolved organic carbon in surface water is somewhat unclear. Studies looking at the effect of burning regimes either through laboratory experiments, plot studies or catchment scale studies are contradictory, some identifying burning as the main driver for colour and others seeing no increase in DOC through burning (Holden et al., 2012). The effect of burning can often not be clearly disentangled from the effect of vegetation cover, or the stream water data cannot be reconciled with the available data from peatland soils (Holden et al., 2012).

Recreational activities in the Scottish Highlands are increasing, including hill walking, climbing, skiing, mountain biking, and motor sports. These activities impact on water quality in several ways. Firstly, the fact that activities usually peak at a certain time can challenge rural treatment plants, especially in summer months when low flows reduce the dilution ability (Soulsby et al., 2002). Secondly, to reach remote areas, road maintenance is required, including road salts in winter with subsequent runoff. This also includes other pollutants such as hydrocarbons and heavy metals. Lastly, in areas with few sanitation facilities, there is potential for pathogens from humans to enter surface waters (Soulsby et al., 2002).

Forests in Scotland are mainly commercially used and can in places cover the whole of some upland catchments. Bad management of forests can contribute to water quality deterioration through drainage and harvest and maintenance operations and increasing acidity, nutrient leaching, and sediment loads (Forestry Commission, 2017). Native, unmanaged woodland however is thought to be beneficial to water quality, although these have shown higher levels of pollution with faecal coliforms (Schoonover & Lockaby, 2006).

Most agricultural pollution occurs in the productive arable areas of eastern Scotland. About a third of the nitrogen input in catchments are exported to surface waters (Poor & McDonnell, 2007). Nitrogen is imported into a catchment through agricultural use of fertilizers and manure, and through sewage emissions and atmospheric deposition. Most areas in the uplands however are used for sheep grazing, which is associated with increased levels of erosion and fine sediment delivery to streams (Soulsby et al., 2002). Sheep dips (pesticide solution into which sheep are immersed to control external parasites) can also have an impact if they are discharged or leak into surface waters, or when draining off the sheep (Hooda et al., 2000). It can be observed that except for acidification, water quality is directly related to population density and the intensity of human activity (Gilvear et al., 2002).

2.4.4 Effects of future changes

The annual mean temperature in the UK has increased by 0.9% for 1991-2020 compared to 1961-1990, with all months and seasons seeing an increase, and annual average rainfall has increased by 6% (Kendon et al., 2021). The largest increases have been seen across Scotland with 2008-2017 being on average 11% wetter than 1961-1990 (Met Office, 2019). Temperatures are projected to further increase in the UK under the high emission scenario, with 0.7°C to 4.2°C in winter and 0.9°C to 5.4°C in summer. Changes in UK averages for precipitation are projected at -1% to +35% for winter and -47% to +2% for summer (Met Office, 2019). Overall, winters are projected to become warmer and wetter and summers hotter and drier, with more frequent droughts (Gosling, 2014). Snow cover and its duration are projected to decrease in Scotland (ASC, 2016).

Reduced summer rainfall in combination with higher summer temperatures are likely to lead to reduced runoff and low flows, and more and more severe droughts, whereas an increase in winter rainfall can lead to higher runoff and flows, increased catchment wetness and an increase of flooding and soil erosion. Rising temperatures overall will lead to an extension of the growing season and rising water temperatures. All of these will affect vegetation and species, the landscape and society. Shifts in the spatial range of species can already be observed, for example cold-loving species in Scotland that have lost habitat at the margins of their southern range. The reduced snow cover is affecting composition of montane vegetation with a shift towards more homogenous, less diverse communities (ASC, 2016).

Changes in temperature and precipitation change the concentration of substances in surface water through altered physical and chemical processes. Higher temperatures probably lead to increases in dissolved substances and decreases in dissolved gases (Delpla et al., 2009), while changes in flow also influence substance concentration. Low flows in summer could be especially problematic as they lead to lower dilution potential. The pattern of precipitation also influences solute and sediment transport, with droughts and higher precipitation in winter inducing more erosion and nutrient leaching (Panagos et al., 2017; Poggio et al., 2018). Peatlands, which are especially important for drinking water supply, are very sensitive to changes in soil moisture regimes and warmer and drier conditions could put pressure on already stressed peatland habitats and induce increased release of carbon, leading to water colouration (ASC, 2016).

Effects of climate change on human activities and land use will also influence water quality. It is envisaged that increases in temperature will make some areas in Scotland, especially those currently marginal for cultivation, more suitable for arable agriculture, while areas currently being used for crop growth could become more prone to drought conditions (Brown et al., 2010). The potential for expansion of prime agricultural land is estimated at 20-40% by the 2050s for Scotland, while 40-50% of prime land could have moderate to severe drought risk (ASC, 2016). Grassland productivity will benefit from warmer temperatures particularly in the marginal upland areas, while it may decline in the drier areas in eastern Scotland. These changes will likely result in changes to water quality at catchment scale and in standing waters through increased input of organic and inorganic fertilisers and chemicals leading to increases in nitrogen, phosphorus, and pesticides (Pakeman et al., 2018). Arable cropping is also associated with increases in soil erosion and suspended sediment in water (ASC, 2016). Pasture is associated with higher pathogen contamination and contamination with suspended solids, phosphorus, ammonia, nitrate, and organic carbon (Pakeman et al., 2018).

Land use changes, which are a main driver for water quality alterations, will respond directly not only to climate change but also prompted by government policies and commitments. There are for example commitments to raise forest cover, enhance and protect biodiversity, or increase energy generation from renewable energy in Scotland. The extent to which the implementation of these targets contributes to water quality improvements or deterioration will mainly depend on how sensitively this is carried out with regard to sustainability and to water sources.

2.4.5 Catchment initiatives to improve water quality

The WFD has introduced integrated planning, ensuring that all uses of and pressure on water resources are considered and actions to reduce pressures to achieve good status are implemented. Working towards good status in all water bodies is beneficial to increasing water quality also in drinking water catchments, bathing water and shellfish waters. As one of the biggest impacts to water quality in Scotland is rural diffuse pollution, SEPA has a Rural Diffuse Pollution Plan in which Diffuse Pollution Priority Catchments play a key role (SEPA, n. d.). In two cycles, 54 priority catchments have been identified that fail to meet environmental standards. The work in these catchments starts by characterising the catchment and collecting evidence of pollution sources and pathways (e.g., cultivation too close to water bodies, slurry and manure spreading, erosion caused by livestock etc), raising awareness, and engaging one to one with land managers (e.g., through workshop, evening meetings, and site visits). It also helps to identify if landowners comply with the Diffuse Pollution General Binding Rules (DP GBR), which were established under the Controlled Activities Regulations. The DP GBRs set out a range of measures (e.g., with regard to fertilizer and manure application, livestock management, land cultivation, pesticide application or operating sheep dips) to be adopted by farmers to mitigate diffuse pollution from their activities.

An initiative that tries to promote a catchment approach specifically to improve drinking water quality is the Drinking Water Protection Scheme by Scottish Water. Many catchments that supply drinking water do not meet the required standard and some expensive treatment is hence necessary. The incentive scheme is aimed at land managers to implement catchment-based mitigation measures to reduce pollution to water bodies that go beyond what they are legally required to do (e.g., by the rules for NVZs or the DP GBR). These incentives consist of financial assistance to implement identified measures. The scheme is only available in specific areas identified by Scottish Water, based on catchments where water quality issues are experienced. Land managers within these areas are eligible to apply for financial assistance to implement suitable mitigation measures identified for each of the catchments. These include Nutrient Management Plans, pesticide control, stock fencing, livestock watering, field management, measures to reduce surface flow, and peatland restoration (Scottish Water, 2017).

The measures selected under this scheme already indicate areas of concern for public drinking water sources. Using data from the monitoring programme of the public supplier (Chapter 3), raw water quality is more strategically examined and analysed to understand drivers, pressures, and intrinsic catchment vulnerabilities as a basis to understand impacts of future changes (Chapter 4).

3. Methodology and data

This research builds on the relationships between water quality and catchment properties to understand how changes in hazards, exposure and vulnerability may impact water quality in future. It also aims to look at differences in water quality profiles between catchments, and how this is linked to different catchment characteristics. The analysis therefore uses data that can broadly be distinguished as water quality data and as catchment data, and combines these two in various different analysis steps.

This section first discusses approaches to modelling (3.1). It then describes the two data sets, water quality data (3.2) and catchment characteristics data (3.3), including the origin, steps to prepare it for the chosen approach to the analysis, and a description of some summary statistics. Finally, how these two datasets are used in the subsequent analyses is briefly outlined in 3.4.

3.1 Methodological approach

Water quality data were provided by Scottish Water from their routine monitoring programme. Depending on the water quality parameter, samples are taken from once a week to once every three months, and as they are not taken on exact dates, the data points are not equally spaced in time. Samples are taken at the treatment work rather than in the catchment and are thus not accompanied by data that would help to describe wider environmental conditions such as flow or temperature. It was checked if flow data could be obtained from nearby streamflow gauges, however as catchments tend to be small headwater sources for the majority there were few in sufficient proximity that would have created reliable data, so any form of flow-weighting was discounted. It was also checked if additional data to the data for the chosen indicators could be obtained, either to include in the analysis or to compare against, for example from monitoring by SEPA or from the research projects of the James Hutton Institute (JHI). Again, as Scottish Water catchments are very small, for the vast majority there is no meaningful overlap between the SEPA or JHI catchments and this could not be used for the broader analysis.

Therefore, while data from a large number of catchments were included in the dataset, the dataset for each catchments tended to be small, and flow data were not available. The data were therefore unsuitable for process-based models. However, the number of parameters covered provided a good basis for an attempt at catchment profiling, while the good spatial coverage of the dataset invited to focus on spatial patterns in water quality and draw conclusions from those. These approaches are in line with the identified need for a more strategic approach to risk assessment and management on a programme level, that considers water supply at the system level. Risk screening was identified as especially suitable to act as a frame, as well as a first step, for integrating climate and land use change impacts into risk assessment as set out in Figure 1.1. Empirical modelling approaches were considered the most appropriate to achieve the aims of the research, answer the research questions and draw on the strength of the available dataset, and the data were accordingly prepared for use in empirical analysis.

3.2 Water quality data

Scottish Water takes samples at treatment works from taps, both for the “raw” water (all water going into the treatment process; this could be water mixed from several sources, or mixed with already treated water to aid the treatment process), and for some catchments also for “source” water (water from one single catchment before treatment). For the purpose of this analysis, it is essential that the water quality data are untreated and reflect just one catchment, so the analysis focused exclusively on source water. Some catchments for which source data have been available have nevertheless been excluded. These are catchments:

1) at the beginning or in the middle of cascade systems, as piping between the reservoirs of one system is common – water quality data from catchments at the end of a cascade have been included but catchments have been merged to include the whole catchment area of the system (8 cases).

2) where two or more catchments serve one treatment work and piping between catchments is suspected.

3) where there are issues with the catchment boundaries – this was the case for groundwater sources, as catchments are defined by Scottish Water by drawing circle of 1 km radius

around the intake. This does not reflect the actual catchment and groundwater sources were therefore excluded.

4) with very few data (less than 15 samples for one parameter).

From the original total of 398 supply catchments (Figure 3.1), this left 154 catchments for which water quality data were analysed (Figure 3.2). Only 8 parameters are sampled for all: aluminium, colour, pH, iron, manganese, presumptive coliforms, presumptive *E. coli*, and turbidity. Data was originally made available for the period 2011-2016. Sampling regimes vary per catchment with parameters being sampled from every three months to every week. For most catchments, samples are taken once per month. Several catchments have a higher frequency of sampling however, for example weekly. This means that the sample sizes vary between catchments, but range between 18 and 238 samples per indicator, with usually approximately 100 samples for aluminium, colour, manganese, iron, and turbidity, and approximately 40 for bacteria and pH.

To arrive at comparable data for each catchment, summary statistics were calculated for each catchment. Parameters included are the median (as distributions are highly skewed, medians are the robust measure reflecting the middle points of the data better than means), minimum and maximum values, the 5th and 95th percentiles (to get a better value for high and low values that are not as much affected by extreme outliers), and the 1st and 3rd quartiles (or 25th and 75th percentiles).

Additional data were later provided for colour for the years 2017 and 2018, and for total organic carbon (TOC) for 127 catchments for 2013-2016. These data were used for specific parts of the analysis (colour data for 2018 in 5.3.2, and TOC data in 5.2.1, 5.2.3, and 5.2.4).

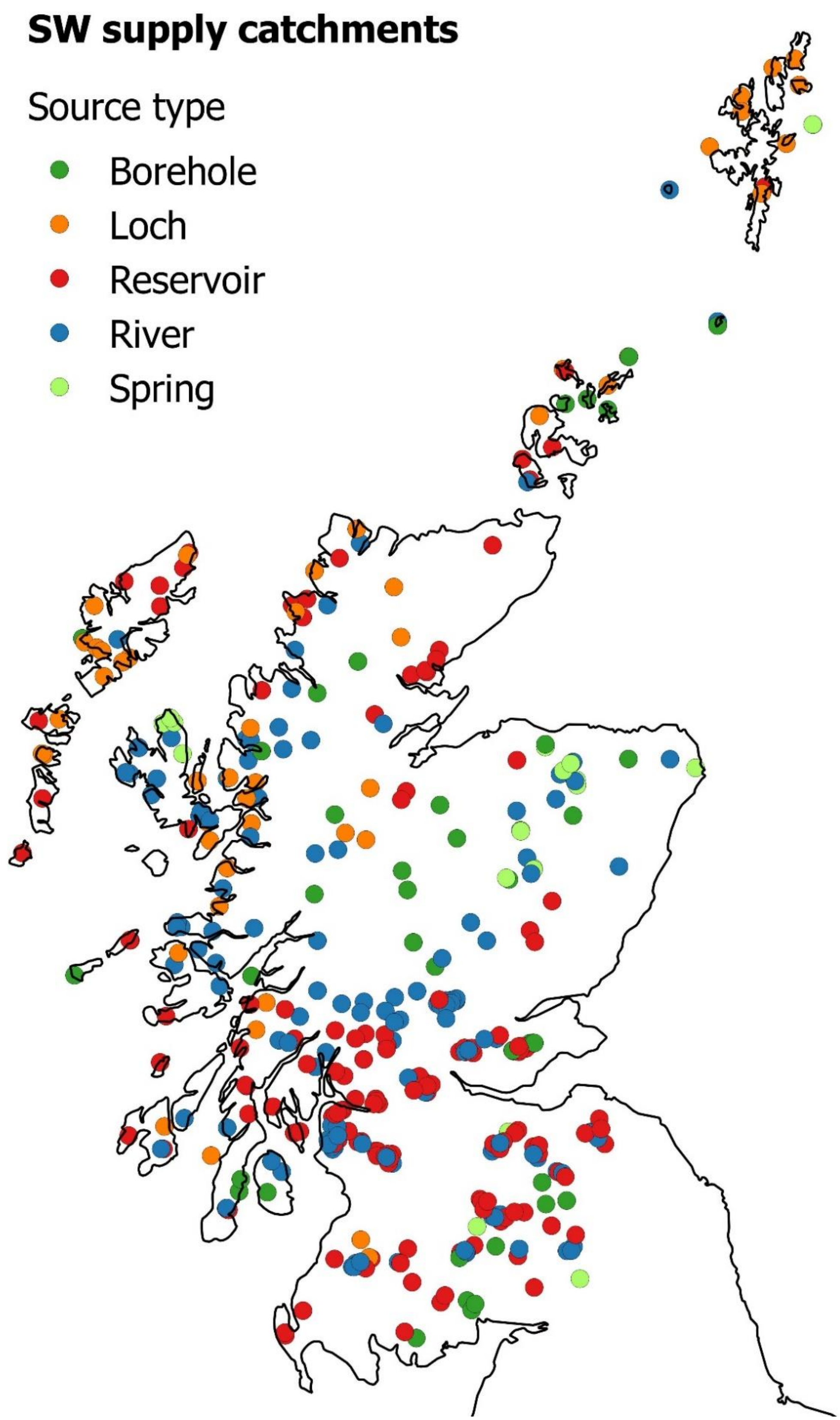


Figure 3.1: Locations of all 398 catchments of water bodies used by Scottish Water as supply sources, by type of source.

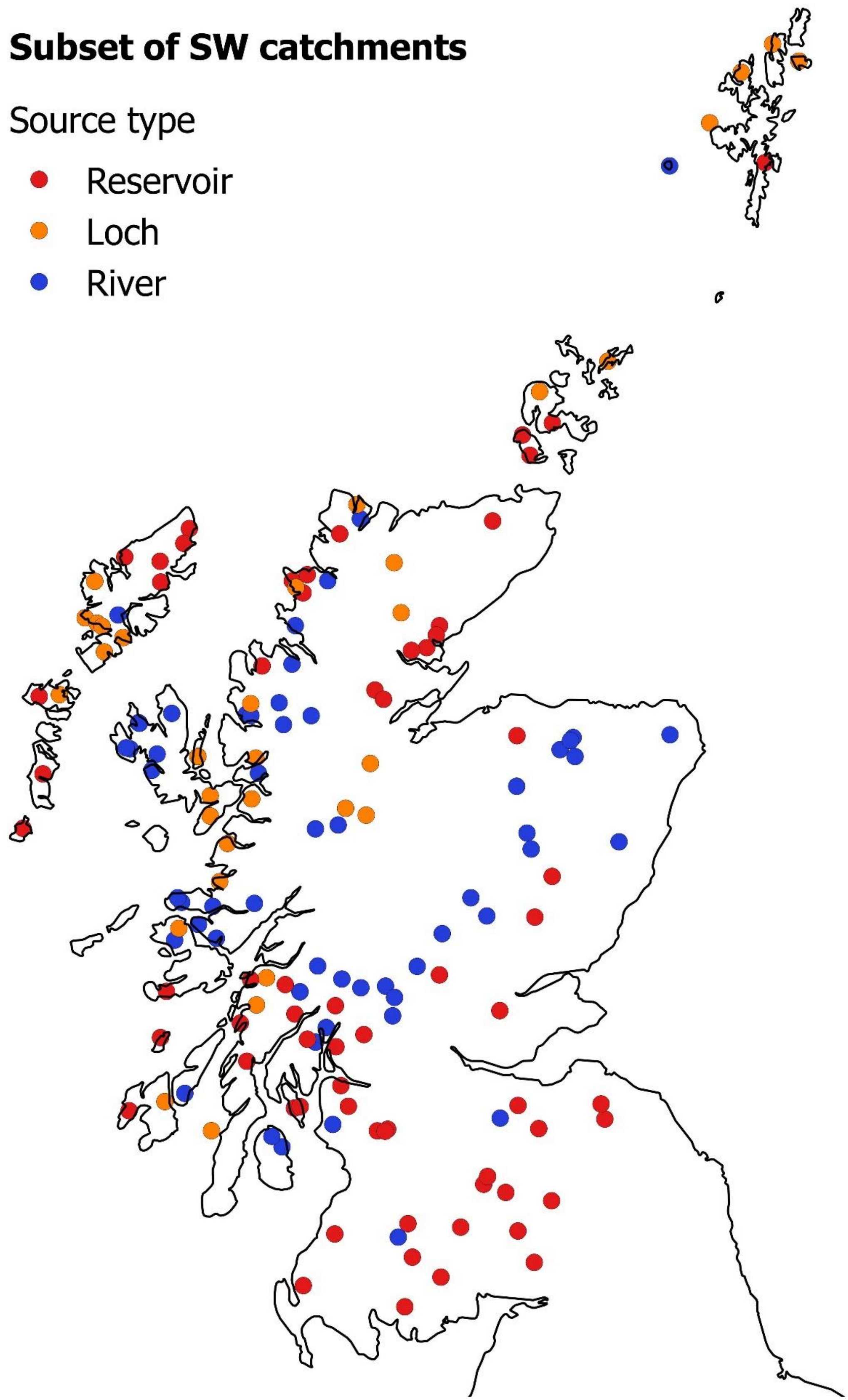


Figure 3.2: Locations of the 154 catchments of water bodies used by Scottish Water as supply sources that were included in the analysis (referred to as subset), by type of source.

3.2.1 Validation

In the initial passing and cleansing of the data for analysis it was observed that in many catchments, and for all indicators except for pH, there were a few exceptionally high values compared to the majority of lower concentrations. Some outliers also seemed unfeasibly high suggestive of errors, so the data needed to be checked for integrity and representativeness.

Data points that are of concern in the planned analysis are extreme outliers (usually very high values; in case of pH also very low values) as they will strongly influence some descriptive statistics and may affect some statistical analysis. If these points are errors, they need to be removed. However, genuine extremely high values are of interest in the analysis as they can reflect a change in conditions (e.g., a flush out event) or contamination event, and they are of concern with regard to water treatment. These data points need to be kept. It is assumed that erroneous data points are the result of wrong entries into the database, rather than analytical errors or contamination.

Bacterial data (coliforms, *E. coli*) can peak to extreme highs e.g., if contamination occurs. Sources cannot definitively be traced back to catchment perturbations as they may be unrelated to weather conditions, instead may arise from unpredictable events such as faecal contamination from livestock ingress into the channel network or uncontrolled spillages from farming operations. Validation of these points is therefore difficult to impossible. *E. coli* is a form of coliform bacteria, so as a basic check for *E. coli* data, any *E. coli* values that were higher than the corresponding coliform value were deemed erroneous and adjusted to the coliform value *(4 data points)*.

For aluminium, iron, manganese, colour, turbidity, and pH, it was assumed that errors usually occur in only one parameter at a time and without any hydrological explanation such as a prolonged drought period or heavy rainfall event. Therefore, data points that were more than four times the standard deviation removed from the means of the parameter in the catchment were checked in several steps, illustrated in Figure 3.3.

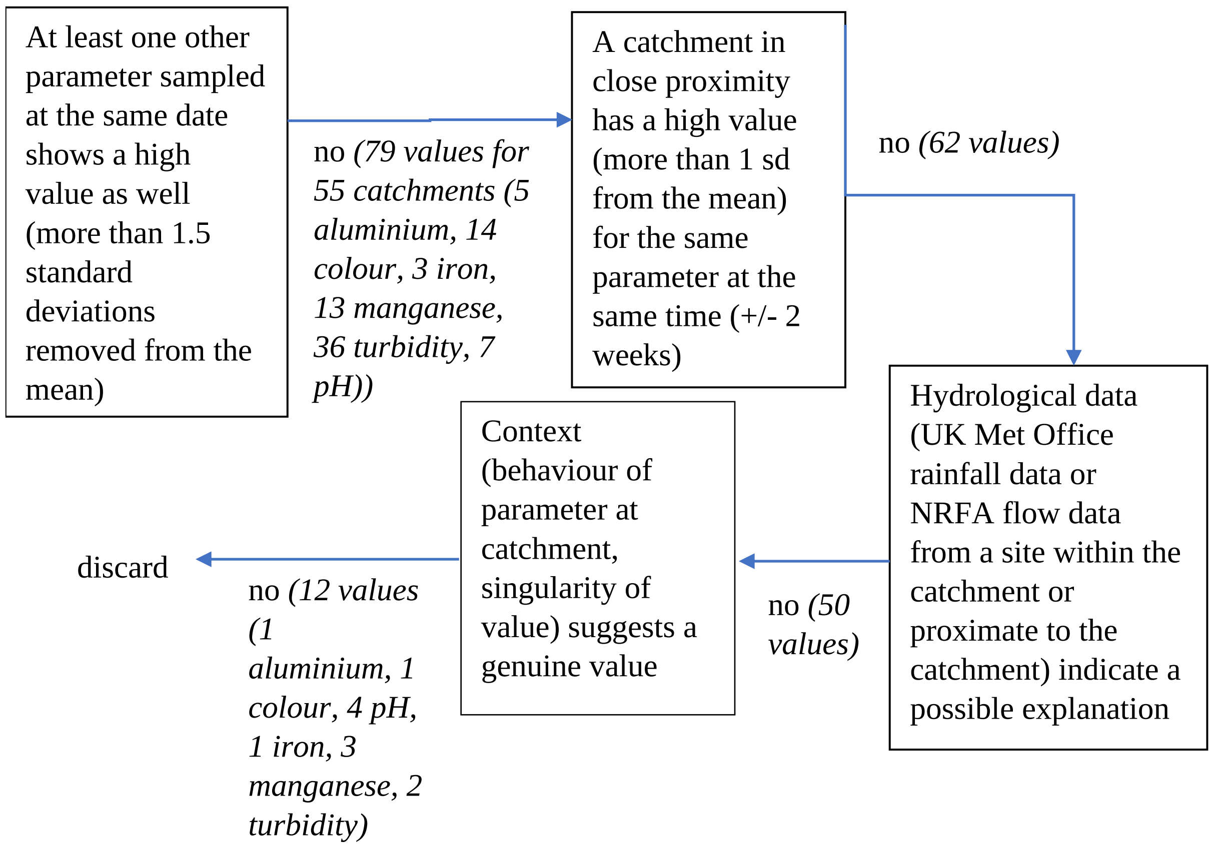


Figure 3.3: Validation process for data points that were more than 4 times the standard deviation removed from the means of the respective parameter in the respective catchment. Water quality data undergoing this process were part of the initially provided data by Scottish Water for aluminium, colour, iron, manganese, pH, and turbidity from 2011 – 2016.

As a result of this procedure, 12 values in total were removed from the dataset for the purpose of further analysis.

3.2.2 Indicators

Eight commonly sampled parameters were examined: aluminium, colour, pH, iron, manganese, coliforms, *E. coli*, and turbidity. As the purpose was to relate these water quality indicators to catchments characteristics, it is crucial to understand what drives their presence and concentrations in surface waters.

I. Aluminium

Aluminium solubility increases with increasing acidity at low pH and with increasing alkalinity at high pH (Neal et al., 2011). There is a wide array of aluminium complexes in natural waters, plus urban or industrial contamination occurs with aluminium oxides/hydroxide. Aluminium concentrations are generally highest under high flow conditions when soil water input is highest, and it often correlates with iron and DOC, as these are also mobilised within the soil. Reversal of acidification has led to decreasing aluminium concentrations, although this may be counteracted by colloids formation (Neal et al., 2011). As forested catchments are generally more acidic and reversal of acidification is slower, higher aluminium concentrations are often found coming from these catchments (Battarbee et al., 2014).

II. Iron

In Scotland, high iron and manganese concentrations are generally associated with acidic pH and organic, poorly drained upland peat soils and high flow conditions. Some catchments observe highest concentrations with the first autumn storms and successively declining concentrations, due to iron and manganese accumulating in the soil over the drier summer months (Abesser et al., 2006). In surface waters, iron readily binds to natural organic matter, and with decreasing acidification, iron concentrations have been increasing (Neal et al., 2008).

III. Manganese

Manganese is mobilised under reducing conditions, but lost from water in well-oxygenated environments where oxidation occurs. Concentration in groundwater is higher than in surface waters (Homoncik et al., 2010). Its release from the catchment is significantly influenced by pH, temperature, and organic matter (Chen et al., 2015). In a study in a forested catchment in Wales, Rowland et al. (2012), found that manganese reached surface water bodies from rainfall and surface soil leaching and concentrations are highest at the beginning of high flows before a dilution effect sets in, and that concentrations increase for several years after felling. Manganese concentrations can also be influenced by in-lake processes and may override any catchment controls (Graham et al., 2012).

IV. Colour

Colour arises from suspended particulate matter and solutes including iron, manganese, and DOC. In Scotland, colour is usually induced by DOC, originating from organic material in the soils and is therefore usually highest in catchment draining peat or other organic rich soils (Dawson et al., 2008). DOC production and release is also influenced by soil wetness, drying and wetting cycles, and temperature (Freeman et al., 2001; Clark et al., 2009). Large increases in flow are often associated with high loadings in colour, iron, and manganese (Abesser et al, 2006). There have been increasing trends for DOC over the past decades in Scotland and the Northern hemisphere, possibly due to acidification reversal (Monteith et al., 2007). DOC can also be produced or removed within the water body (Chapman & Palmer, 2016).

V. pH

The pH of environmental waters results from the balance of acids and bases in the water, with carbon dioxide being particularly influential in regulating pH (Boyd, 2015). PH is thus influenced by the surrounding rocks and soils, but also precipitation (e.g., acid rain) and biological processes (photosynthesis, respiration, decomposition). Carbonate and bicarbonate add to the alkalinity of water, meaning the capacity to buffer acid influences, so waters with low alkalinity show higher fluctuations in pH (Boyd, 2015). Waters with high alkalinity are often associated with limestone.

VI. Turbidity

Turbidity describes the cloudiness or haziness of water due to suspended solids. Solids in water originate from dissolution and suspension of minerals and organic matter from soils and geological formations, or aquatic organisms and their remains. Sources of suspended particles in surface water are erosion of catchment soils by rain and overland flow, erosion of stream and lake beds and banks, and resuspension of sediment (Boyd, 2015). Disturbance of soils (through forestry, agriculture, mining etc.) and runoff from urbanized areas can lead to an increased amount of sediment and particles reaching water bodies (Siakeu et al., 2004).

VII. Coliform

Total coliforms include bacteria found in the soil, vegetation, water, human and animal waste. Coliform bacteria are often used as a general indication of the sanitary condition of the water, with waters containing a significant number of coliforms considered a health hazard due to the possibility of disease (Boyd, 2015).

VIII. E. coli

E. coli is a faecal coliform originating from the faeces of warm-blooded animals, and is a widely used indicator for faecal contamination of water (Odonkor & Ampofo, 2013). Sources for *E. coli* in surface water are usually livestock, wild animals, pollution from sewage or slurry and manure application (Rotariu et al., 2012; Vinten et al., 2004; Tetzlaff et al., 2012).

3.2.3 Water quality description

Looking at median concentrations for catchments, it can be seen that similarly to the individual samples per catchments, most catchment medians are within the lower concentrations with few catchments branching up to very high median concentrations, except for pH as this is on a log-scale (Table 3.1 & Figure 3.4). For some parameters (colour, coliform bacteria, and *E. coli*), most catchments show median concentrations that are higher than the Scottish drinking water standard (The Public Water Supplies (Scotland) Regulations 2014), emphasising a widespread need for treatment. Catchments falling outside the required range for pH (6.5-9.5) were more acidic, while no catchments have medians that are too high.

Table 3.1: Distribution of median concentrations per water quality indicator for the 154 Scottish Water catchments included in the analysis.

	Minimum	1st quartile	Median	Mean	3rd quartile	Maximum	Drinking water standard
Aluminium (µg Al/l)	9.00	42.25	64.75	76.26	92.88	404.00	200
Colour (mg/l Pt/Co)	2.00	24.00	33.75	38.95	49.62	167.50	20
Iron (µg Fe/l)	7.0	108.1	179.5	267.4	361.8	1465.0	200
Manganese (µg Mn/l)	1.00	5.50	13.00	22.12	27.00	405.50	50
pH	5.5	6.7	7.1	7.064	7.4	8.4	6.5-9.5
Turbidity (NTU)	0.20	0.45	0.60	1.03	1.00	6.10	1 (leaving the treatment works)
Coliforms (CFU in 100ml)	9.0	88.0	190.0	279.2	285.0	4950.0	0
E. coli (CFU in 100ml)	0.00	1.00	4.25	17.62	10.00	580.00	0

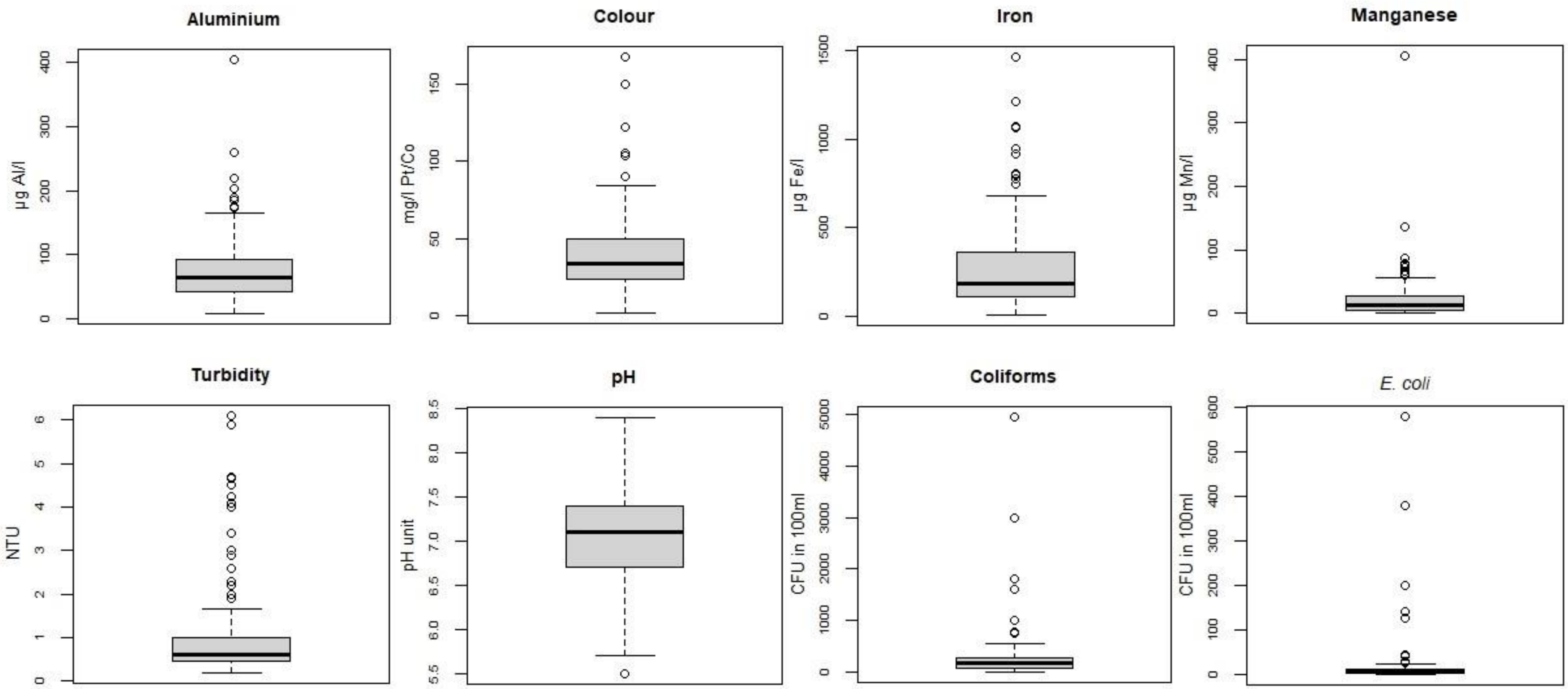


Figure 3.4 Boxplots showing the distribution of catchment median concentrations per water quality indicator, including the 154 catchments of the subset.

Some of the water quality parameters looked at are expected to show correlations, and to influence each other. Catchments that have high concentrations of *E. coli* should also show high coliform values. Conditions that increase aluminium concentrations also influence the release of DOC, so catchments that show high colour values will probably also show higher concentrations for aluminium. Iron is also influenced by DOC and by similar mechanisms that influence DOC release, so it is expected to see a link here (Halliday et al., 2012). Furthermore, iron and manganese also produce colour, so a relationship between colour, iron and manganese should be observable. The pH influences the solubility of metals, e.g., aluminium and iron are more soluble in acidic waters. Iron and manganese also affect the turbidity of water. Spearman's rank correlation testing was carried out in R to see the strength of linear relationships between catchment medians (pH was omitted as no linear relationship is expected, Table 3.2).

Table 3.2: Results for Spearman's rank correlation tests on catchment median concentrations per water quality indicator (orange for strong relationships (rho >0.6), yellow for moderate relationships (rho>0.4), green for weak relationships (rho<0.4) that are significant (p<0.05)).

	Colour	Iron	Manganese	Coliform	*E. coli*	Turbidity
Aluminium	rho = 0.54 p = 7xe^{-13}	rho = 0.55 p = 2xe^{-13}	rho = 0.42 p = 5xe^{-08}	rho = 0.17 p = 0.033	rho = 0.24 p = 0.002	rho = 0.51 p = 2xe^{-11}
Colour		rho = 0.7 p = 2xe^{-16}	rho = 0.5 p = 3xe^{-11}	rho = 0.12 p = 0.13	rho = 0.06 p = 0.47	rho = 0.52 p = 4xe^{-12}
Iron			rho = 0.71 p = 2xe^{-16}	rho = 0.28 p = 0.0005	rho = 0.38 p = 1xe^{-06}	rho = 0.74 p = 2xe^{-16}
Manganese				rho = -0.03 p = 0.72	rho = 0.27 p = 0.0008	p = 2xe-16 rho = 0.8
Coliform					rho = 0.56 p = 4xe^{-14}	rho = 0.09 p = 0.245
E. coli						rho = 0.37 p = 2xe^{-16}

Strong positive relationships between catchment median concentrations could be observed between iron and colour, manganese, and turbidity, and between manganese and turbidity. Moderate positive relationships were found between aluminium and colour, iron, manganese, and turbidity; between colour and manganese and turbidity; and between coliform bacteria and *E. coli*. Positive weak, but significant correlations were found between aluminium and coliform bacteria and *E. coli*, between iron and coliform bacteria and *E. coli*, between manganese and *E. coli*, and between turbidity and *E. coli*.

3.3 Catchment data

Catchment data were needed to understand the relationship between natural or man-made characteristics of the catchment and water quality. In order to do so, it is therefore necessary that the data represented factors that might have an influence on water quality and differentiated enough to be able to explain variations. To be able to choose which variables should be included, in what detail, and from which source to derive them, it was therefore necessary to first identify which characteristics influence the water quality and how, and how these characteristics could be represented (Table 3.3).

Table 3.3: Overview of catchment characteristics under consideration for inclusion in the dataset intended for use in the empirical modelling of this research.

Characteristic	**Influence**	**Descriptors**
Kind of source	Rivers, lochs, and reservoirs have different dynamics that influence water quality. In general, lakes have a longer residence time for water and a greater mixing of water sources from the catchment than rivers, and seasonal stagnation and cycling patterns. Reservoirs, depending on their design, can resemble either rivers or lakes, or a combination, with regard to their water chemistry.*	Categorisation Loch, River, Impounding Reservoir
Area	Bigger and more diverse catchments might be more resilient as they may have more capacity to buffer influences. However, the bigger catchments might be the ones on lower altitudes with more anthropogenic impacts that are associated with higher pressures and subsequently lower water quality. The stability in water quality should be reflected in the range of concentrations in an indicator, however this does not entirely reflect fluctuations as some catchments might have isolated incidences of e.g., high bacteria load while otherwise remaining quite stable. The water quality data will also not reflect complete ranges as sampling is not frequent enough to ensure capturing all high and low concentration events.	Catchment size
Location/ Continentality	The location influences other factors such as climate, soils etc. Rainfall shows a strong East-West trend in Scotland, and it would be expected to see some differentiation here. Continentality of the catchment would influence the temperature, especially the day and night differences and seasonal differences in temperature, as well as the amount of rainfall received, with oceanic catchments receiving more rain than continental ones.	British National Grid (BNG) coordinates, Distance to sea, Latitude, Annual temperature range, Conrad's formula
Topography	The topography is important as it determines how the water flows through the landscape, which has implications for water quality. For example, water that runs off the land quickly can pick up particles more easily. Important parameters for topography are therefore the steepness of the slopes. This will be dependent on land cover too, however. A high percentage of South and Southwest facing aspects could mean that more rain is falling on the ground and with higher impact, as this is the predominant wind facing direction, which would potentially result in more soil disturbance and higher runoff. Elevation in itself is not directly related to water quality, but is used to derive further topographic characteristics of the catchments, such as relief ratio (also giving an indication	Slope classes (little, medium, steep), Elevation (mean, max, min), Relief ratio, Hypsometric integral, Elongation ratio, Aspects facing South and

	of steepness), and the hypsometric integral (giving an estimate of the distribution of elevation, high values indicating a large proportion of the catchment at high elevation). Elongation ratio could be used to get an idea of the shape of the catchment, also indicating how close most areas of the catchment are to the surface water, indicating how far water would have to travel.	Southwest,
Soil	The soil influences water quality by controlling how water is stored and flows through it, and by minerals and ions that are in the soil and washed out by water. Soils that have low absorbing potential will yield quicker runoff and a flashier regime, which brings a higher potential of carrying sediment and particles into the water source. This would be reflected in all parameters. The HOST** class is a number that is derived from a variety of parameters that influence the hydrological response of the soil. The baseflow index allows an estimate of how much the water is influenced by groundwater as opposed to runoff from land, which will influence composition of the water quality. Standard percentage runoff will give an indication how much water runs off as overland and subsurface flow, again influencing the speed of the runoff and the quality of the water when it reaches the water body. A crucial factor in water quality for Scotland is organic carbon (causing colouration), which is strongly influenced by organic soil (mainly peat).	HOST class, average baseflow index (BFI), dominant BFI, average standard percentage runoff (SPR), dominant SPR, percentage peaty soils, percentage eroded peat
Bedrock	The type of bedrock will determine the way water flows but also influence the chemical properties of the water and the pH. With regard to the former, the permeability of the rock is important, with regard to the latter, the type of rock (e.g., siliceous, calcareous, or organic). Manganese, iron, and aluminium will also be influenced by the natural environment.	Bedrock classed into 4 categories: igneous and metamorphic, sandstone, limestone, other sedimentary
Land cover and vegetation	Vegetation influences the way water runs off the land in various ways: interception of the rainfall (raindrops don't hit the soil directly), roots allow water to infiltrate more easily, and facilitation of aggregation, again facilitating water infiltration. Vegetation also influences the nutrient cycling in the soil (and atmosphere), hence influencing nutrient content in the soil that can be washed out. Arable land has inputs of nutrients and other agrochemicals that can reach the water bodies, and faecal pathogens through slurry application. Improved grassland also has inputs of nutrients as well as microbiological pollutants through grazing animals. Woodlands can be	Land cover classes

	beneficial for water quality, e.g., through slowing water down. Natural woodlands are generally considered beneficial while planted, commercial forests can have water quality impairments if managed badly. Semi-natural habitats are considered to have less nutrient enrichment and a better ability for water purification. Urban land has a high percentage of impervious surface that causes water to run off quickly, picking up pollutants on the way (such as heavy metals, nutrients, microbial pollutants).	
Livestock	Land use alone is not necessarily related to water quality, but land management may play a big role. One factor for example would be the number of livestock on the field.	Number of sheep/cattle
Wildlife	Wildlife can lead to contamination with bacteria, sediment, and colour (higher surface runoff from moorland).	Numbers or density of deer/wildfowl
Conservation area	Land that has a conservation designation might be land that is functionally less compromised and that is managed in a "water-friendly" way. This is not necessarily so, but it might still be worthwhile to see if it works as an indicator for water quality.	Area under a conservation designation
Rainfall	Rainfall has a multitude of effects, directly and indirectly. For example, higher rainfall leads to higher runoff, influencing water quality. Depending on conditions in the catchment prior to rainfall, e.g., the wetness of the catchment, rainfall could lead to a wash out into the water, or to dilution, or a combination. Seasonal effects are likely to show.	Total rainfall, rainfall days
Temperature	Temperature could influence land cover and use. Air temperature will also influence water and soil temperature, in turn potentially influencing water quality. Temperature could also be an indicator of continentality of the catchment, which in turn will mean variation for a number of factors (temperature ranges, rainfall, wind speeds, slopes, land use etc.)	Annual average temperature, monthly means
Combination	Many factors will have varying influences, depending on other factors. Rainfall for example would be expected to act differently regarding water quality in a steep catchment with low absorbing soils and short vegetation than in a catchment with high absorbing soils, tall vegetation, and little slopes.	
Unidentified	There will be pressures that influence the water quality such as e.g., groundwater influence, land management, land compression, artificial drainage, etc. which have not been used to describe the catchments and can therefore not be identified in the analysis. Some of these pressures could be associated with other characteristics or spatial associations. This could then give an indication if other data to describe these pressures should be obtained.	

* Hayes *et al.* (2017) suggest several differences between reservoirs and lakes that also influence water quality, e.g., that catchments are generally larger for reservoirs than lakes. This seems to be indicated for this dataset – the median catchment size for reservoirs is greater than for lakes (4.29 km^2 and 1.82 km^2 respectively), although the biggest catchments are found within the loch group. It is also suggested that reservoirs would be located lower in the catchment. Other suggested differences such as that reservoirs usually have a more elongated shape, are shallower and have a higher temperature, cannot be tested as there is no data.

**Hydrology of soil types (HOST) classes have been developed considering the distribution of soils and the hydrological response of catchments using conceptual models of the processes in the soil and substrate, resulting in 29 classes (Boorman et al., 1995).

3.3.1 Data sources and preparation

Shapefiles for catchments were provided by Scottish Water. Some variables could be calculated directly from the shapefiles using GIS, others were calculated using additional data obtained from various sources (Table 3.4). To be able to see if the catchments represent the “standard” for Scotland or if they over- or underrepresent some characteristics, means and percentages for Scotland as a whole were calculated.

Many of the identified characteristics will influence each other. For example, it will be important what kind of vegetation is on what kind of soil. The variables as created are not able to cover this but trying to reflect this in this kind of data would create a large number of variables which would become impractical for empirical analysis. The proximity of the different characteristics such as geology, soils and land cover to the water body could also play an important role that is not reflected in the variables. It might be possible to include variables for a buffer zone (of e.g., 200 metres around the water body). Obtaining a reliable map of water courses however proved unfeasible so these data were not created.

Land management has a big impact on water quality especially for arable, forest and urban land covers. This could include the timing of fertilizer and slurry application on fields, the timing of livestock on fields, access of livestock to water courses, felling operations in forests etc. These considerations were not practicable to assess at the national scale.

Table 3.4: Variables describing catchment characteristics that were included for empirical modelling in subsequent analysis, and how they were derived, including sources and processing steps.

Characteristic	**Variable**	**Data & source**	**Data preparation (ArcGIS)**	**Comment**
Source	*Type* - River - Lake (loch) - Impounding reservoir	Scottish Water catchments shapefile		As the water bodies under observation here differ greatly (from small upland tributaries and lochs to major rivers (such as the Tay or the Ugie) and lochs (such as Loch Ness), categorising them as lochs, rivers and impounding reservoirs will probably not yield clear grouping. It would be expected that lochs and rivers e.g., from a similar upland region will show a more similar picture with regard to water quality than if compared to other rivers and lochs.
Topography	*Slope classes:* - *little slope = 0-3°* - *moderate slope = 4-16°* - *steep slope >16* Percentage of catchment area *Aspect* Percentage facing south or southwest (158°-247°)	OS Terrain DEM (50m) downloaded from Digimap (https://digimap.edina.ac.uk/)	DEM tiles put together using the 'Mosaic' tool Calculated using the 'Slope' tool, raster then converted to shapefile, percentage for each catchment calculated using 'Tabulate Intersection' Calculated using the 'Aspect' tool, raster reclassified into 0 (0-157 and 248-360 degrees) and 1 (158-247 degrees, South and Southwest), then converted into shapefile, percentage for each catchment calculated using 'Tabulate Intersection'	There will be some collinearity between parameters here, however it might be interesting to see if the last three parameters can be used in the statistical analyses and if they might be used instead of slope and aspect (as they might be more easily derived). Classing slopes might mask an effect.

	Elevation - *Minimum* - *Maximum* - *Mean* *Relief ratio* *Hypsometric integral* Area under the hypsometric curve; found to be mathematically very close to the elevation relief ratio (Pike & Wilson 1971, Singh et al., 2008) *Elongation ratio* Ratio of the diameter of a circle with the same area as that of the basin to the maximum basin length. The value approaches 1.0 as the shape of a drainage basin approaches to a circle. The ratio varies from 0.6 to 1.0, typical values are close to 1.0 for regions of very low relief and are between 0.6 and 0.8 for regions of strong relief and		Derived through 'Zonal Statistics as Table' (or 'Zonal Statistics' in QGIS) (Maximum elevation – Mean elevation) / Maximum length of river basin Maximum length of river basin: Calculated by putting a rectangle (by width) around the catchment using the 'Minimum Bounding Geometry' tool (add geometry characteristics) (Mean elevation – minimum elevation) / (maximum elevation – mean elevation)	

	steep ground slope (Kumar, 2011).			
Area	*Size* in km^2	Scottish Water catchments shapefile		
Continentality	*Distance to nearest sea* *Annual temperature range* *Conrad's formula* Hyper-oceanic when CCI in between -20 to 20, oceanic/ maritime when CCI in between 20 and 50; sub-continental when CCI in between 50 and 60; continental when CCI in between 60 to 80 and as extreme/hyper-continental climate when CCI in between 80 and 120 (Gadiwala et al., 2013)	Polyline file of the Scottish coastline (http://www.naturalearthdata.com/downloads/10m-physical-vectors/10m-coastline/)	Derived using the 'Generate Near Table' tool, using the coastline as 'Near feature' Taken as the difference between the maximum and the minimum mean monthly temperature (see below). (1.7*Annual temperature range/sin (latitude+10))-14 (Conrad, 1946; Snow, 2005). Latitude derived through converting the XY coordinates of the centroids of the catchment polygons on http://www.gridreferencefinder.com/batchConvert/batchConvert.php	As continentality influences water quality indirectly through influence on rainfall and temperature, both of which are included as direct parameters, this might lead to collinearity. It might be interesting to see if this could be used as proxy in absence of specific rainfall and temperature data.
Geology	*Bedrock classes* *- Igneous and metamorphic* *- sandstone* *- limestone* *- other sedimentary* Percentage of catchment area	BGS Geology 625k (https://www.bgs.ac.uk/datasets/bgs-geology-625k-digmapgb/)	Converted shapefile into raster using the overall class, then reclassified raster into smaller classes (as the 4 above). Then converted into shapefile and used 'Tabulate Intersection' to derive percentages for each catchment.	There is only little percentage of limestone and sandstone in Scottish Water catchments so effects may not be discernible.
Soils	*HOST class group* 4 groups according to the SPR	1:250,000 Soil Map (National	'Tabulate Intersection' on the class	BFI and SPR are already included in HOST class, so these variables

	that determines the class: - *HOST1 very well drained* - *HOST2 medium well drained* - *HOST3 medium poorly drained* - *HOST4 very poorly drained* Percentage of catchment area *Peat* Percentage of the catchment area with soils with a peaty main component *Eroded peat* Percentage of catchment area with eroded peat *BFI* Measure of the proportion of the river runoff that derives from stored sources; the more permeable the rock, superficial deposits, and soils in a catchment, the higher the baseflow and the more sustained the river's flow during periods of dry weather (CEH, n. d.; Gustard et al., 1992) - *value that dominates (highest percentage) in the catchment*	Soil Map) (https://soils.environment.gov.scot/)	Selected all polygons from the soil shapefile that had a peat main component and saved as new shapefile, then 'Tabulate Intersection' for deriving the percentage per catchment Selected all polygons that had Eroded Peat included in description and saved as new shapefile, then 'Tabulate Intersection' for deriving the percentage per catchment Determined the percentage of different values per catchment by using 'Tabulate Intersection'	are dependent. HOST class might be all that is needed with regard to the hydrological behaviour. There are a large number of HOST classes, so they have to be summarized which may mask effects. The same is true for either using an average for BFI or SPR or the dominant value. HOST class does not capture the aspect of what soil contributes directly to water quality parameters. This is captured with regard to organic carbon through peat component or topsoil organic carbon. *Topsoil organic carbon content data were used when it became available and replaced the peat/eroded peat data.*

	- *average value* *SPR* Shows a soil's hydrological response. Low values indicate freely draining soils, higher values indicate soils with impeded drainage (Helliwell et al. 2007) - *value that dominates (highest percentage) in the catchment* - *average value* *Topsoil organic carbon content* Average over the catchment	 Map of topsoil organic carbon concentration (https://soils.environment.gov.scot/)	 Determined the percentage of different values per catchment using 'Tabulate Intersection'	
Land cover and use	*Land cover group* - *arable* - *improved grassland* - *coniferous woodland* - *deciduous woodland* - *mixed woodland* - *semi-natural* - *urban* - *water* Percentage of catchment area *Conservation designation* Percentage of catchment area under designation: - Local Nature Reserve - National Nature reserve - SSSI	CEH Land cover map (shapefile) 2007/2015, downloaded from Digimap Individual shapefiles for conservation designations; Scottish Natural Heritage (now	Converted shapefile into raster, reclassified into the above groups, then converted to shapefile and percentage of cover per catchment determined using the 'Tabulate Intersection' tool Shapefiles merged into one shapefile, then used 'Tabulate Intersection' to derive percentage per catchment	There is the possibility of misclassifications. Summarising classes might mask effects. It might be necessary to distinguish between planted and natural woodland. Natural woodlands might be identifiable through the ancient woodland inventory and the semi-natural woodland inventory from SNH. The semi-natural land cover class comprises a lot of different habitats that might have very different impacts on water quality. It might make sense to separate out a few classes that are usually beneficial, such as heathland,

	- SAC - Ramsar Wetlands - World Heritage Sites - Biosphere Reserves - Country Parks - SPA *Mean number of deer* *Average number of cattle* *Average number of sheep* *Number of septic tanks*	NatureScot; https://cagmap.snh.gov.uk/natural-spaces/category.jsp?code=pa) Deer count density polygons (https://cagmap.snh.gov.uk/natural-spaces/category.jsp?code=hs) Scottish Government Agricultural Statistics; provided for 2013-2016, per parish SEPA Septic tank register (provided by Scottish Water)	 Mean number identified using 'Zonal Statistics as Table' Average number per km^2 calculated per parish, then percentage of area with average number per catchment calculated using 'Tabulate Intersection' and catchment average calculated.	wetlands etc. The influence of a conservation designation might not be strong enough to show, and conservation status alone is not necessarily a measure of the natural integrity of the area. *Land cover was updated to the 2015 data when the dataset became available and subsequently used. The category "mixed woodland" was no longer included.*
Land capability	*LC classes* *- 1-3.1 (Prime land)* *- 3.2-5 (Livestock)* Percentage of catchment area	Land capability for Agriculture national cover map (1:250,000) (https://soils.environment.gov.scot/)		The difference between classes 3.1 and 3.2 is not big although 3.2 is not officially classed as prime land, so it might be worthwhile to see if the differentiation in LC 1-3 and LC 4-5 is better.

Rainfall	*Monthly mean rainfall in mm* catchment average, long-term average over time series 1981-2010 *Mean number of days per month with rainfall >10 mm* average over the years 2007-2011 *Summer effective rainfall (SER)* average over period 1981-2000	5km grid data from Met Office/CEDA Archive (https://catalogue.ceda.ac.uk/uuid/d715d2ac53f14f21acc6952a0278817d) Provided by Dr. Iain Brown	Reclassified into 100m raster, then Zonal Statistics in QGIS For each year from 2007-2011: reclassified into 100m raster, then using Zonal Statistics in QGIS, and averaging over the five years	Using the data per months rather than averaging out over the whole year could make the models more accurate, and it will allow to explore the change within the scenarios per month or season. However, water data are mostly not sufficient for averaging months, so this might not work. *SER data were obtained and used later (from 5).*
Temperature	*Annual mean temperature* catchment average, long-term average over time series 1981-2010 *Annual accumulated temperature (AAT) above 5.5°C* average over period 1981-2000	5km grid data from Met Office (see above) Provided by Dr. Iain Brown	Reclassified into 100m raster, then Zonal Statistics in QGIS	*AAT data were obtained and used later (from 5).*
Projections	*SER* Average over period 2041-2060 *AAT* Average over period 2041-2060 *Land capability* Percentage of classes for period 2041-2060	Provided by Dr. Iain Brown	Based on UK Climate Projections 2018 See Brown et al. (2008)	
Scotland	Boundary	Shapefile downloaded	Scotland selected and exported in ArcGIS	

		from http://www.naturalearthdata.com/downloads/50m-cultural-vectors/		

3.3.2 Catchment description

The shapefile that included catchment outlines for water sources used by Scottish Water included 398 active catchments as raw water sources for drinking water. These consist of direct river outtakes (141), natural lochs (49), impounding reservoirs (146), springs (21) and boreholes (41). The catchments are defined by the point of outtake, following hydrological principles for the surface water sources, but borehole and spring catchments were created by drawing a circle with a radius of 1km around the outtake point. Catchments are distributed all over Scotland, as water provision is by its nature decentralized as it is in most cases uneconomical to transport water over larger distances (Figure 3.1).

The catchments have an average size of just below 40 km^2, however this mean is being distorted by a few very large catchments, and the median size is 3.14 km^2. The biggest catchment, with 4991 km^2, is the River Tay. There are two more catchments above 1000 km^2, Loch Ness (1782 km^2) and Cairnton, River Dee (1383 km^2). Only another 13 catchments are above 100 km^2 in size. Of the rest, the majority are smaller than 5 km^2 (Figure 3.5).

Most catchments are at a mean elevation of between 250 – 300 m (Figure 3.6). The maximum elevation can be found in Cairnton, River Dee with 1309m. Most catchments are dominated by moderate slopes, with only little steep slopes (Figure 3.7). However, some catchments also have high rates of steep slopes (e.g., Allt An Fhuarain with almost 95% of steep slopes and no slopes below 3 degrees). This can also be observed when looking at relief ratio and elevation relief ratio, where values show distributions to be expected for Scotland with a few high outliers (Figure 3.8).

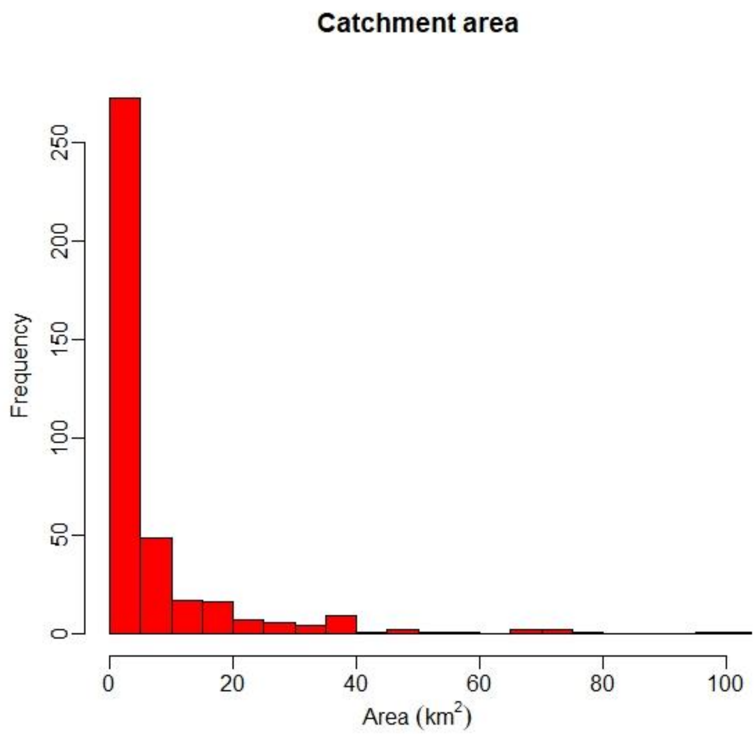


Figure 3.5: Histogram of catchment areas with catchments up to 100 km^2 included (n=382).

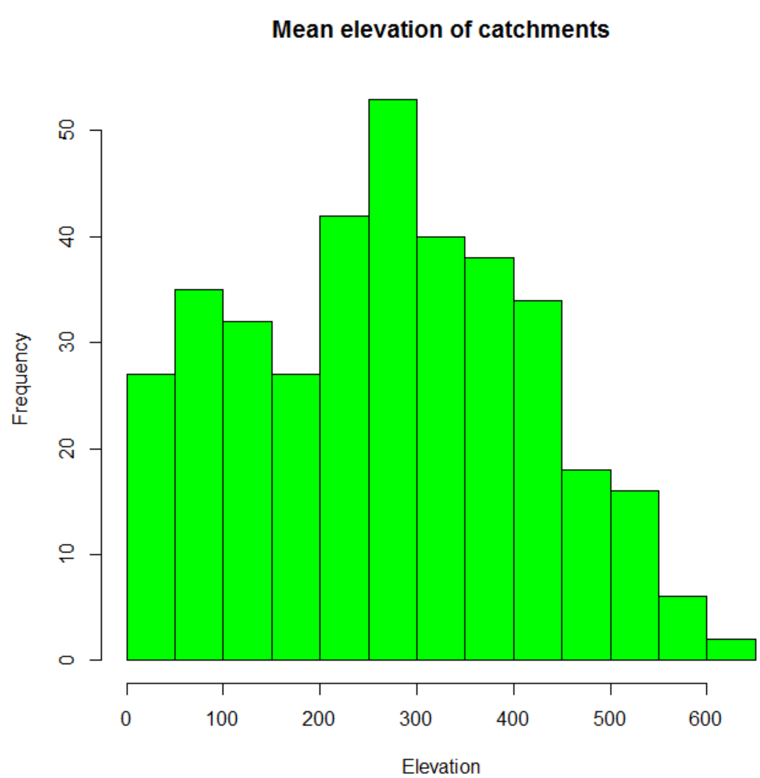


Figure 3.6: Histogram of catchment mean elevation (m AOD) for Scottish Water catchments (n=398).

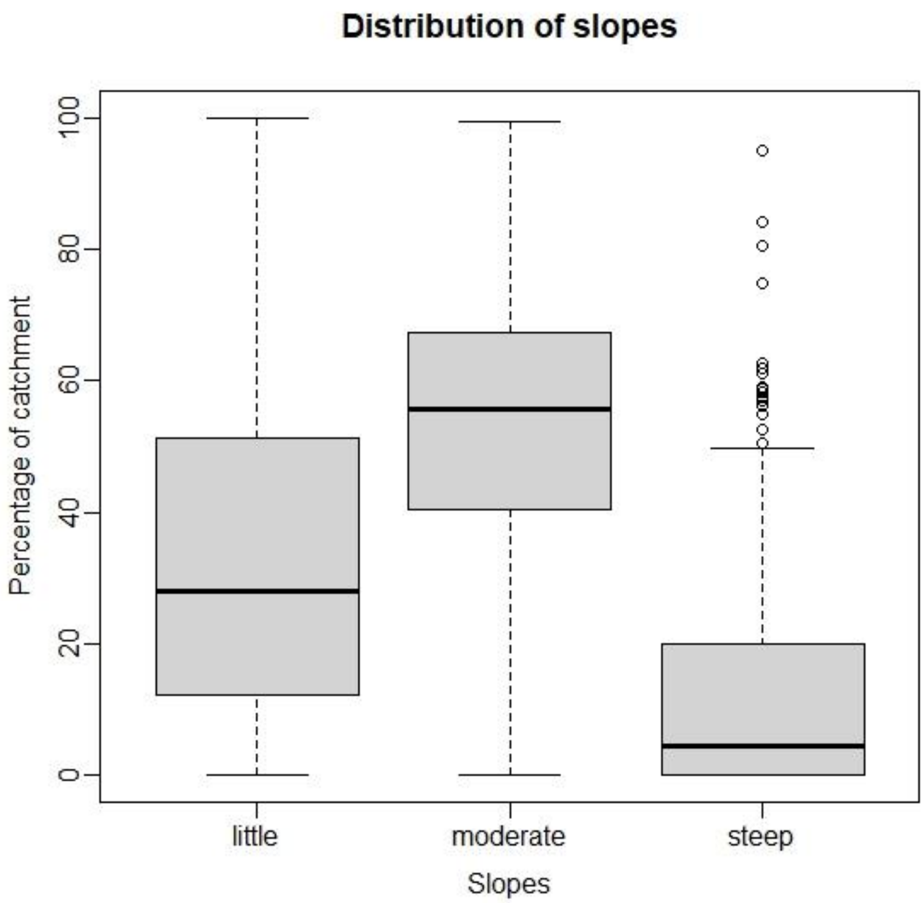


Figure 3.7: Boxplots showing the distribution of slopes within the Scottish Water catchments (n=398). Little = below 3 degrees, moderate = 4 to 15 degrees, steep = above 16 degrees

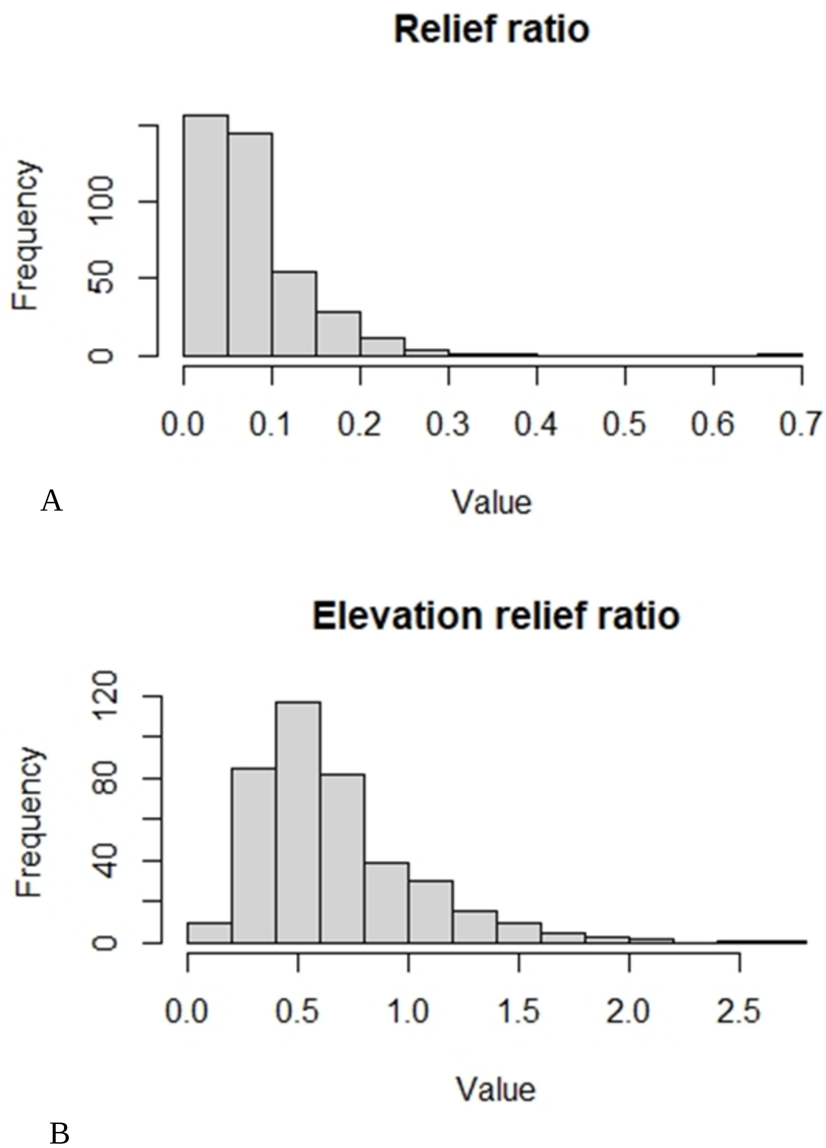


Figure 3.8: Histograms of A. relief ratio values and B. elevation relief ratio values for SW catchments (n=398).

As wind mostly comes from a South and Southwest direction, areas exposed to this direction tend to have a higher exposure to rainfall. It is therefore possible that the amount of slopes facing this direction could make a difference to the hydrological behaviour of the catchment and hence the water quality. Looking at how much of the catchment faces a South or Southwest aspect, we see that most catchments have around 25% of South and Southwest facing aspects (Figure 3.9).

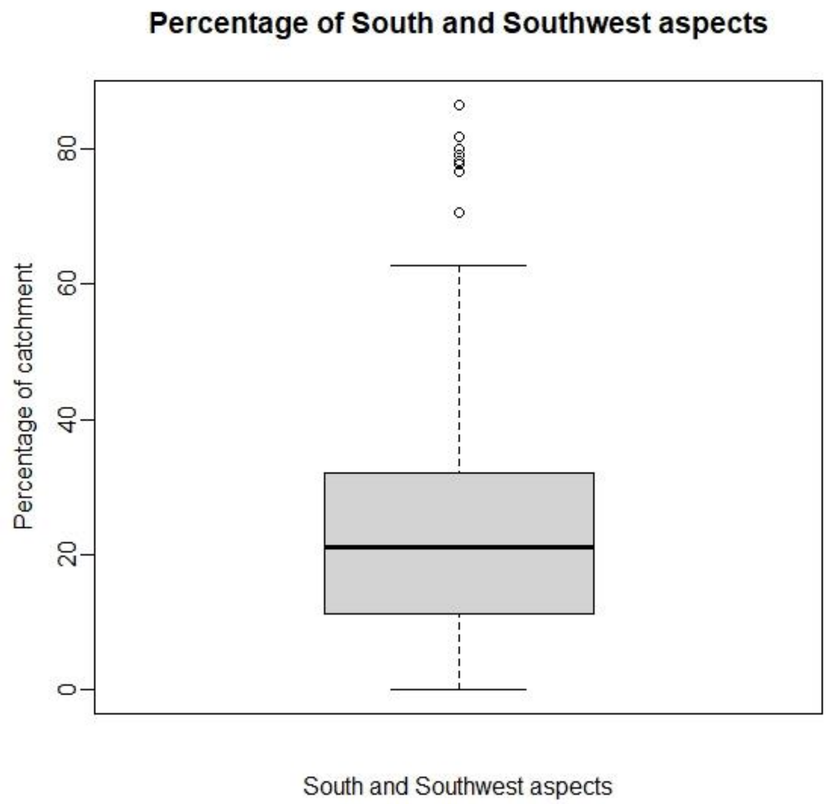


Figure 3.9: Boxplots showing the distribution of percentage of South and Southwest facing aspects in Scottish Water catchments (n=398).

Looking at Scotland as a whole, elevation ranges from sea level to 1347 metres, with an average of 241m. South and Southwest aspects make out 23%, and little slopes occur 17% of the time, moderate slopes 65% and steep slopes 18%. The catchments therefore seem to represent Scottish topography relatively well.

In terms of geology, most catchments are highly dominated by igneous and metamorphic bedrock. Some catchments are also high in sandstone, but there are only few catchments that have a high percentage of limestone (Loch Borralie is an exception with 100% limestone) or other types of sedimentary rocks (Figure 3.10). This represents the Scottish bedrock distribution very well.

HOST classes give an indication of the hydrological behaviour of the soil. The distribution of HOST classes in Scottish Water catchments indicates that most catchments are dominated by soils that have only limited permeability and storage capacity, so that the soil gets saturated from rainfall and water runs off quickly (Figure 3.11).

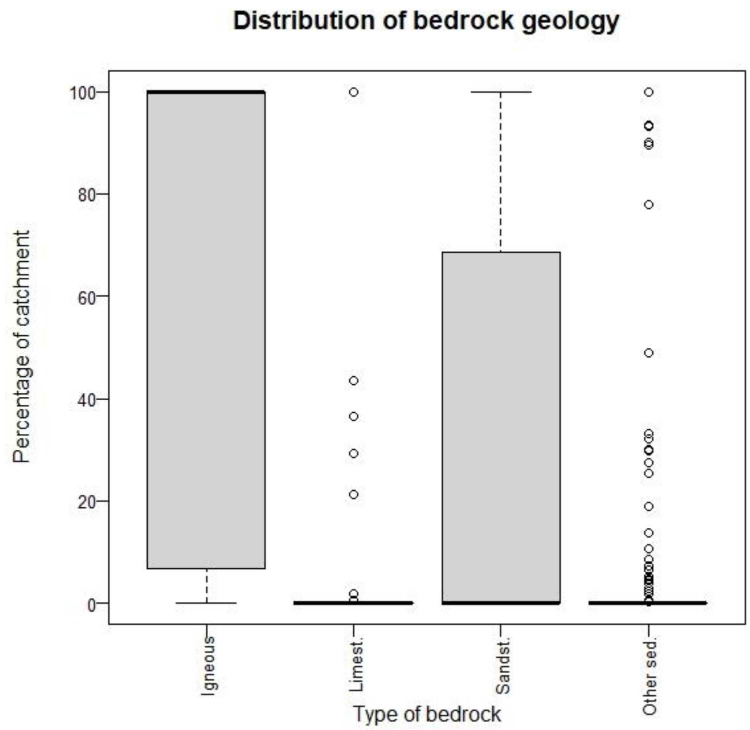


Figure 3.10: Boxplots showing the distribution of percentage of bedrock geology types for Scottish Water catchments (n=398).

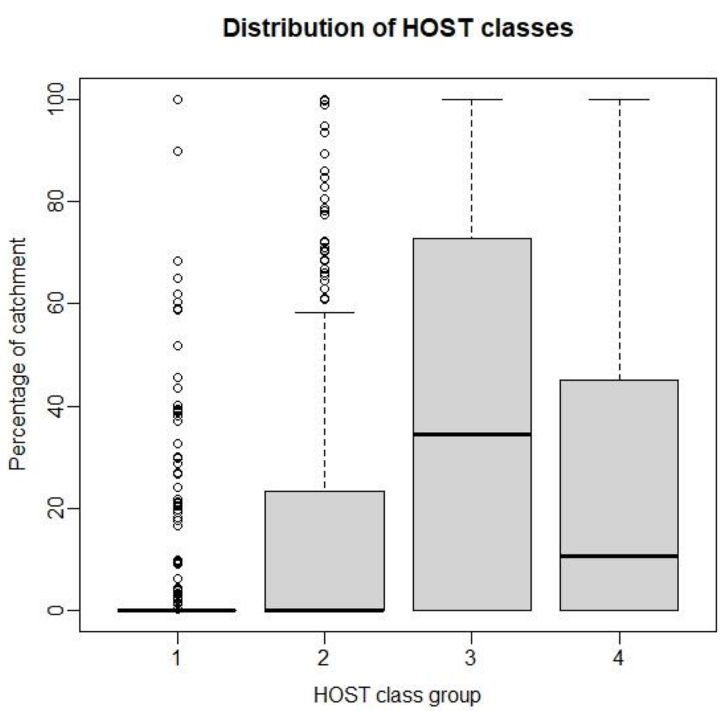


Figure 3.11: Boxplots showing the distribution of percentage of HOST class of soils in Scottish Water catchments (n=398; 1 = HOST classes 1 - 5, 11; 2 = HOST classes 9, 10, 14, 16, 17; 3 = HOST classes 6 - 8, 15, 18, 21, 24, 25; 4 = HOST classes 12, 19, 20, 22, 23, 26 – 29).

Many catchments have a high percentage of peaty soils, a few also exhibit a high amount of eroded peat soils (Figure 3.12). Ten catchments are covered by more than 25% of eroded peat, Gossa Water leading this list with almost 70%.

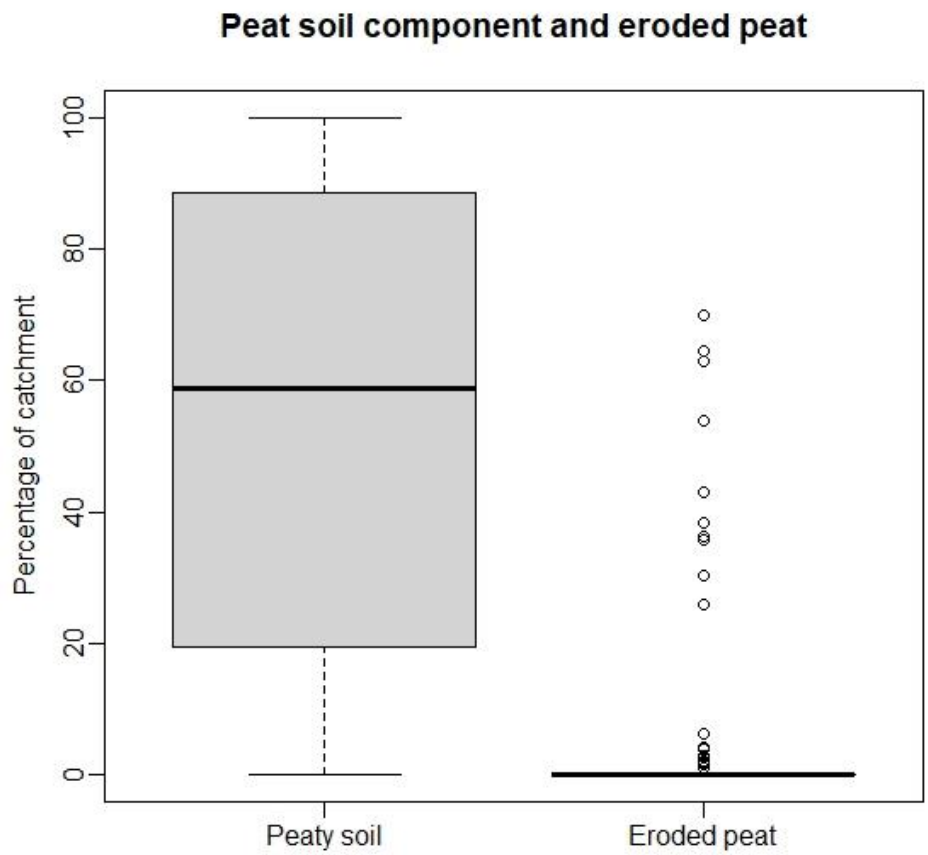


Figure 3.12: Boxplots showing the distribution of percentage of peaty soils and eroded peat in Scottish Water catchments (n=398).

When it comes to land cover and land use, most catchments are highly dominated by semi-natural or natural habitat, followed by coniferous forest (Figure 3.13). Many catchments also have a high percentage of water cover, which is to be expected especially for lochs and reservoirs. Most catchments have low urban areas (below 10%), except for the Larchfield borewell, which has over 54% urban area. With an average of 67% semi-natural habitat, this is slightly above Scotland as a whole (55%). Arable areas are underrepresented, with a mean of 2.7%, compared with 8.9% for Scotland as a whole.

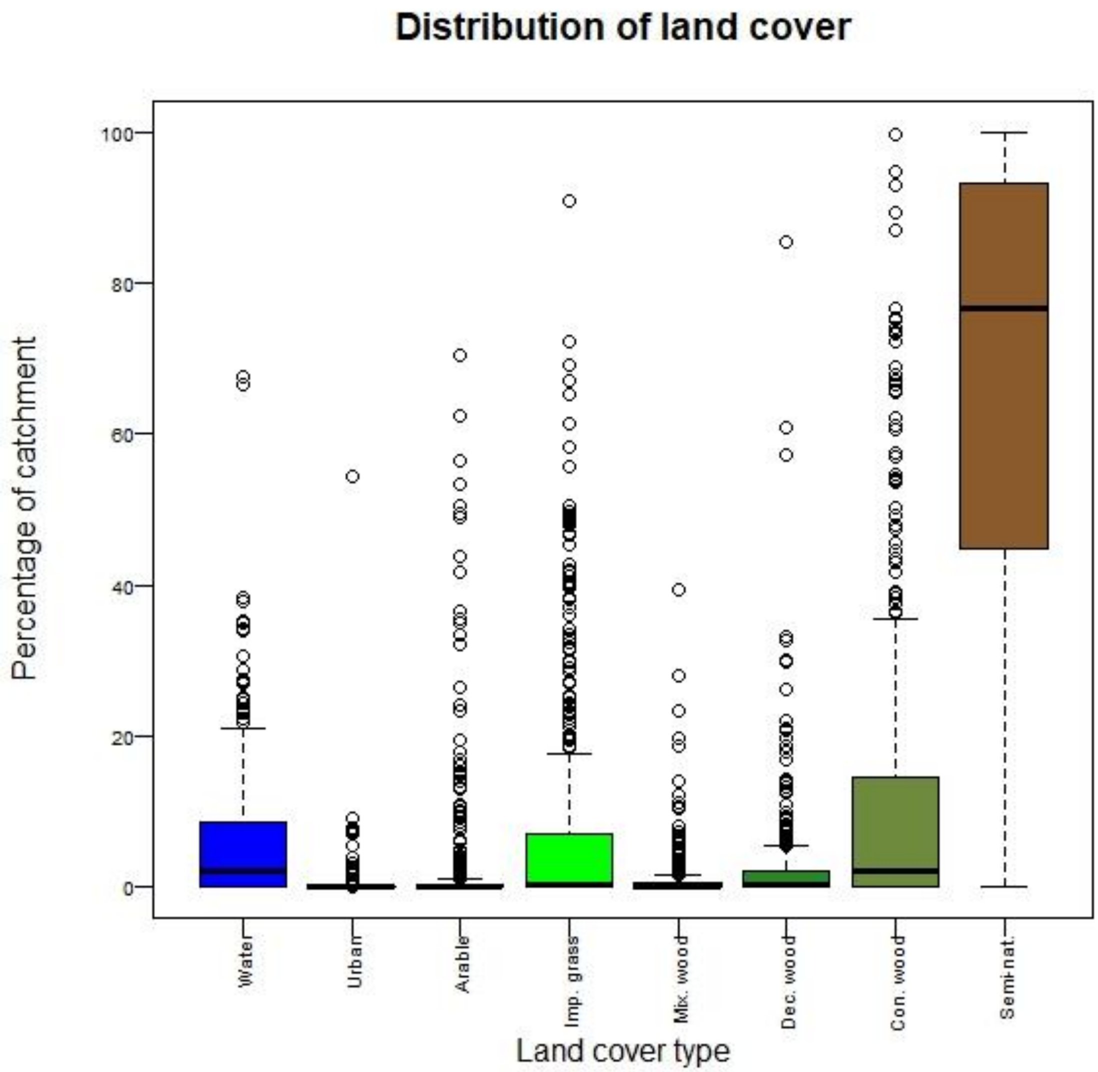


Figure 3.13: Boxplots showing the distribution of catchment percentages of land cover (2007) in Scottish Water catchments (n=398; Water = % area covered by water, Urban = % urban area cover, Arable = % area covered by arable agriculture, Imp. grass = % area covered by improved grassland, Mix. wood = % area covered by mixed woodland, Dec. wood = % area with deciduous woodland, Con. wood = % area with coniferous woodland, Semi-nat. = % all other area).

There are 80 catchments where more than 75% of the area is under a form of conservation designation. 250 catchments have no or only small areas (less than 5%) with a conservation designation.

The mean annual temperature in the catchments ranges between 4.5 and 9.1 degrees Celsius, with most catchments having a mean temperature between 6.5 and 7 degrees (Figure 3.14).

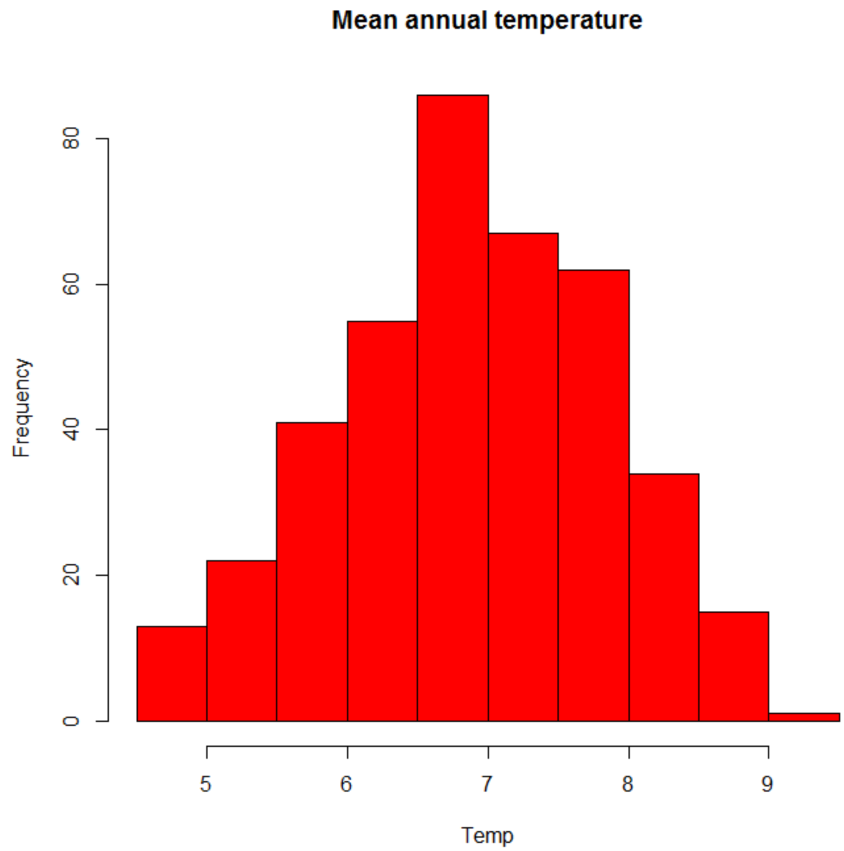


Figure 3.14: Histogram of mean annual temperatures in °C (long-term average 1981-2010) of Scottish Water catchments (n=398).

The mean monthly total rainfall is 135 mm (Figure 3.15A), although a few catchments have significantly higher mean monthly total rainfall, e.g., Loch Sloy has a mean monthly total rainfall of above 300 mm. Most catchments have on average 4 days of rainfall above 10 mm per month (Figure 3.15B).

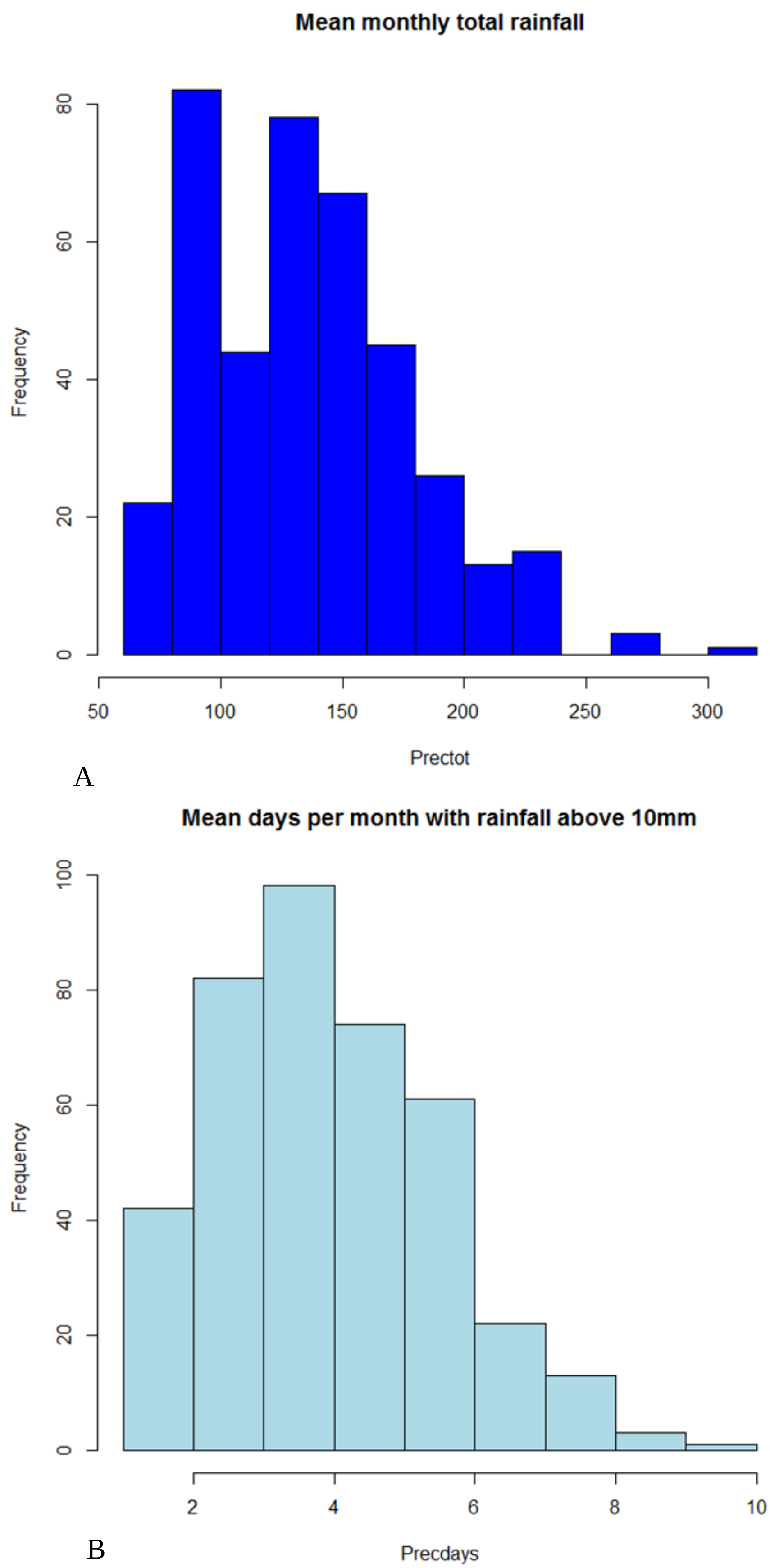


Figure 3.15: Histograms of A. mean monthly total rainfall in mm (long-term average 1981-2010) and B. average number of days with rainfall above 10 mm per month (average 2007-2011) in SW catchments (n=398).

There is a west-east gradient for rainfall in Scotland due to the predominant west wind – this trend seems to be true for the catchments as well (Figure 3.16).

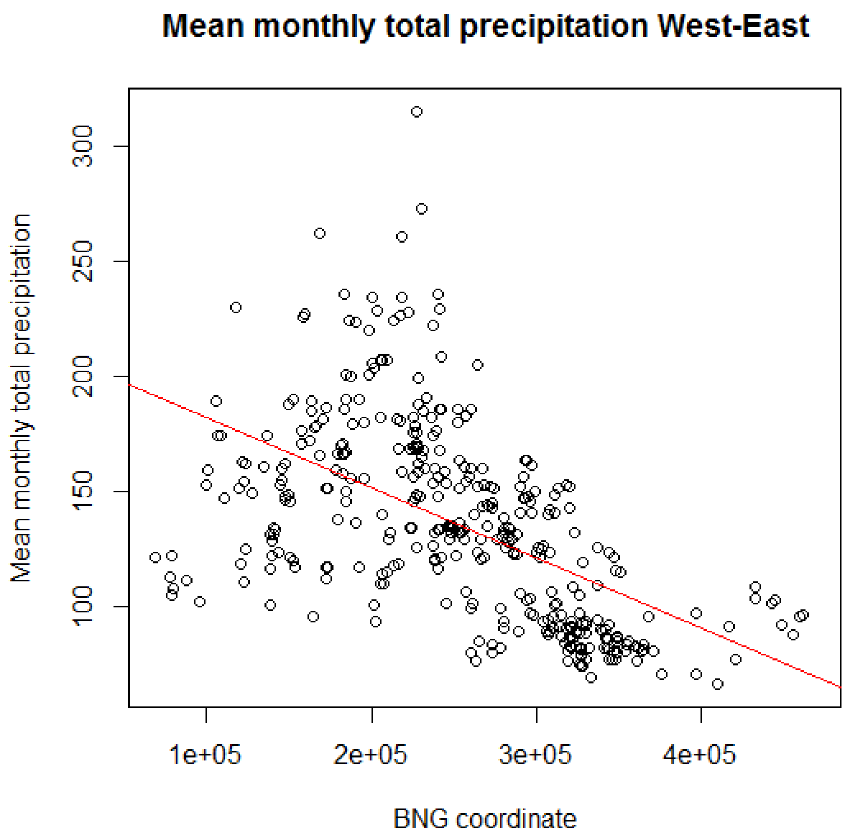


Figure 3.16: Scatterplot with regression line showing the mean monthly total precipitation by west-east coordinate (British National Grid) of the catchment.

In summary, due to catchments being distributed all over Scotland, they reflect the different natural and anthropogenic conditions in Scotland, apart from slight overrepresentation of semi-natural land cover and underrepresentation of arable areas.

3.3.3 Subset of catchments with water data

To analyse the relationships between the identified catchment characteristics and water quality, a subset of 154 catchments was chosen for which water data in sufficient sample density and size were available (see 3.1). The subset included 64 reservoirs, 32 lochs, and 58 river intakes. Ideally, this subset of catchments should represent the diversity found in the total set of Scottish Water catchments. Wilcoxon tests were run to test if any variables displayed significant differences in medians between the groups (included in the subset and not included in the subset), and boxplots were created to visually compare the groups.

When it comes to area, the subset covers the range of the overall set, but the mean is higher (84.6km^2), which indicates that a larger number of smaller catchments has been removed. However, as most catchments are small, this is to be expected. The median is 3.77 km^2, which is only a little higher than the median of the group not included in the subset (3.14 km^2). The elevation also still covers the full range, and the mean is similar with 257 m. Regarding slope, the subset includes a slightly higher proportion of catchments with steeper slopes and a slightly lower proportion of catchments with more gentle slopes. However, both relief ratio and elevation relief ratio show no significant difference in medians. The median percentages in south and southwest facing aspects differ slightly with the subgroup not containing quite the range as the complete group (Figure 3.17).

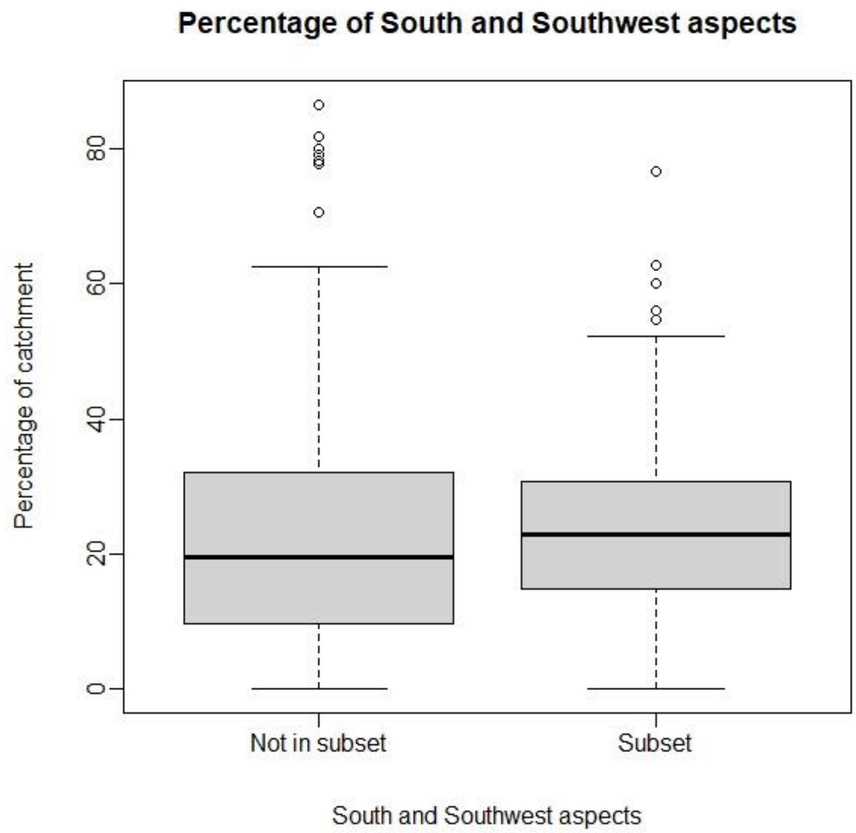


Figure 3.17: Boxplots showing the distribution of percentage of South and Southwest facing aspects in the catchments included in the subset for analysis (n=154) and the catchments not included in the subset (n=244).

For the bedrock geology, the subset of catchments still covers the range of different distributions found in the full range of catchments (Figure 3.18). With regard to the soil, the subset loses the catchments with high percentages of high absorbing soils (Figure 3.19). While the Wilcoxon test gives statistically significant differences for the median for percentages of peaty soils and eroded peat, it is observable that the subset still covers the complete range (Figure 3.20).

Regarding land cover, semi-natural and natural habitat is still more dominant in the subset, while catchments with higher percentages of improved grassland and arable areas seem to have been excluded (Figure 3.21).

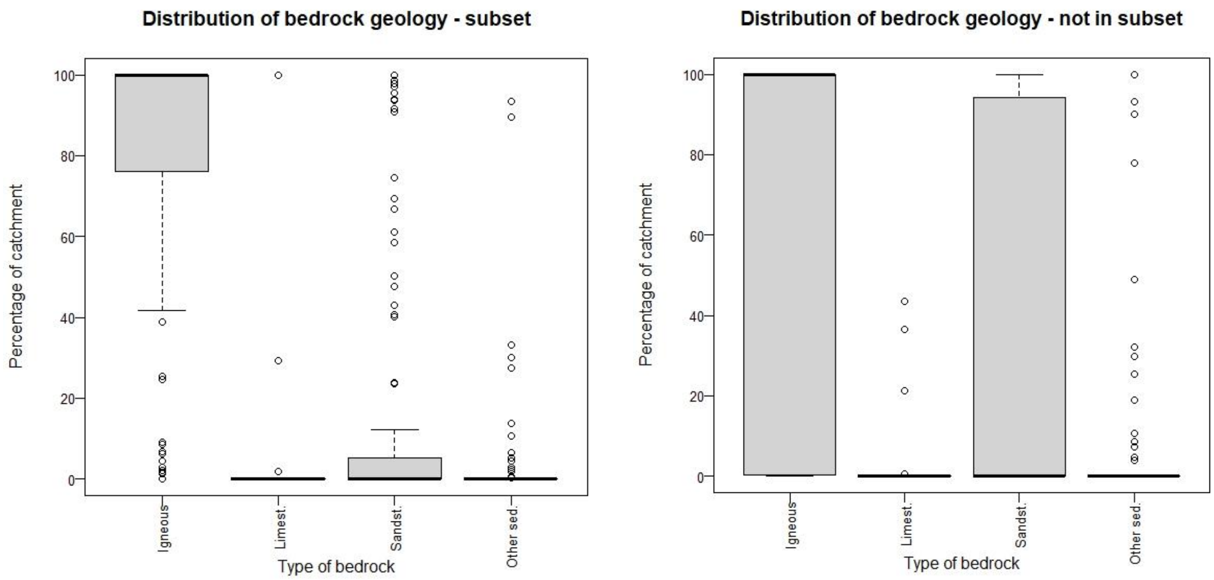


Figure 3.18: Comparison of distribution of percentages for different types of bedrock in the catchments. A. For catchments included in the subset for analysis (n=154). B. For catchments not included in the subset (n=244).

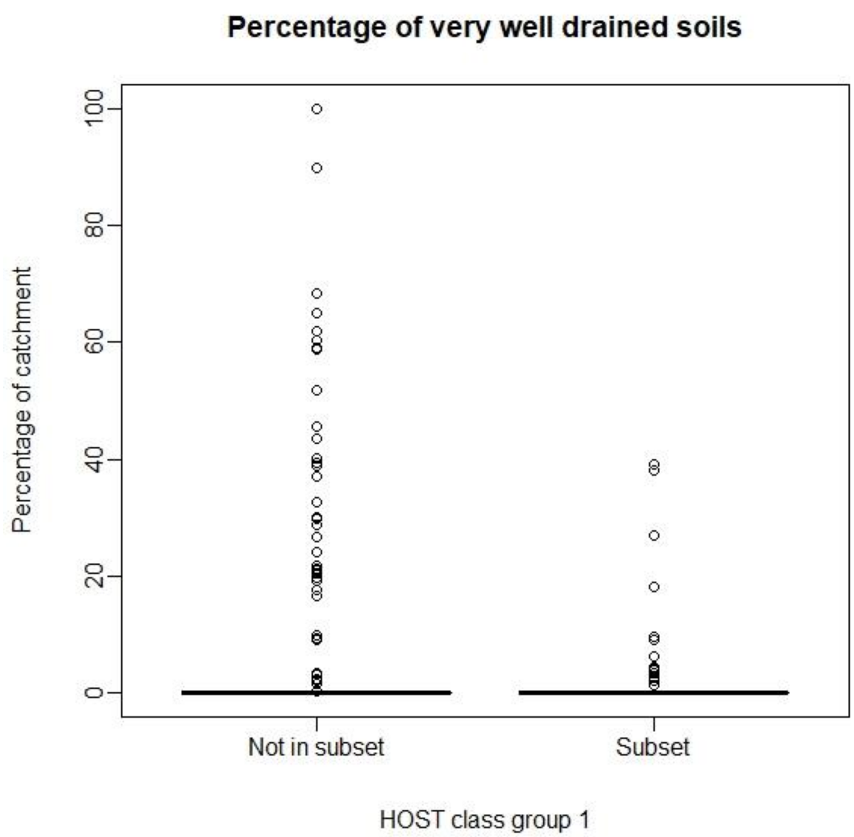


Figure 3.19: Boxplots showing the distribution of percentage for very well drained soils in the catchments included in the subset for analysis (n=154) and the catchments not included in the subset (n=244).

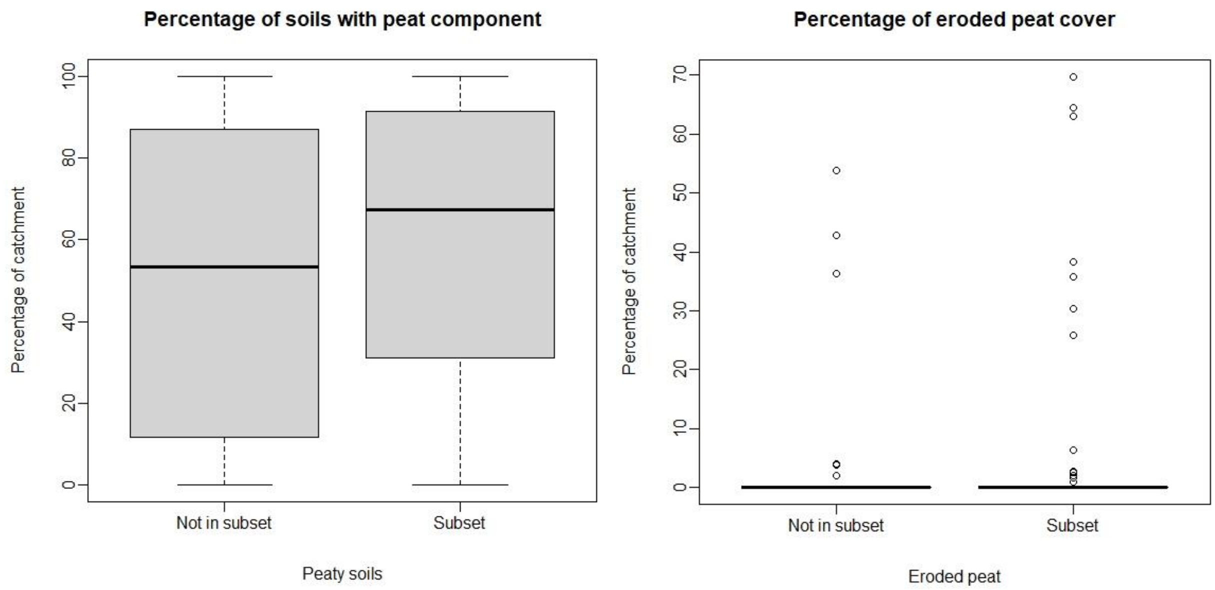


Figure 3.20: Boxplots showing the distribution of percentage of catchment area with peaty main component in the soils (A.) and eroded peat (B.) in the catchments included in the subset for analysis (n=154) and the catchments not included in the subset (n=244).

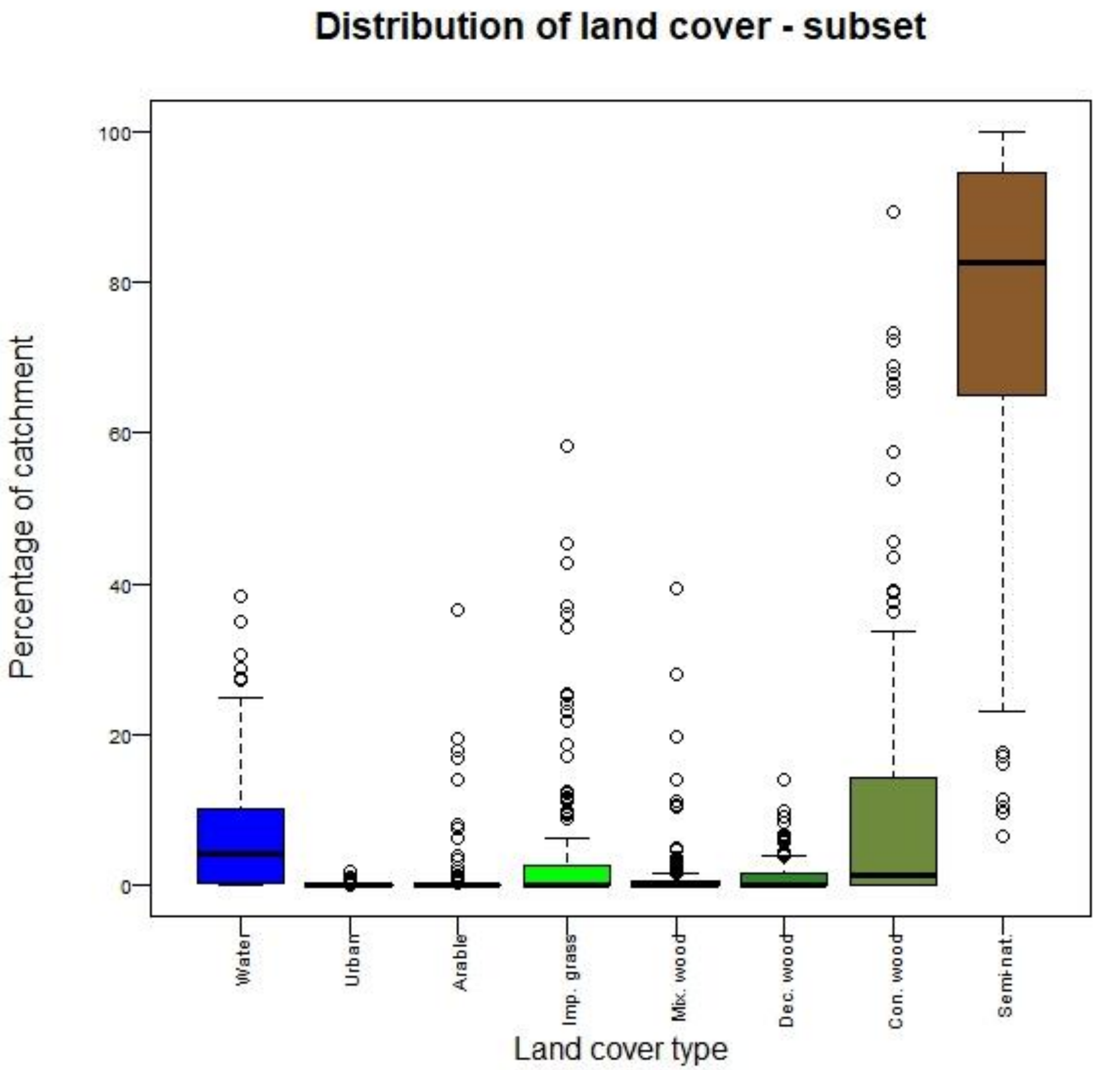


Figure 3.21: Distributions of the percentages of different land covers in the catchments of the subset included in the analysis (n=154; Water = % area covered by water, Urban = % area with urban land cover, Arable = % area with arable agriculture, Imp. grass = % area with improved grassland cover, Mix. wood = % area with mixed woodland cover, Dec. wood = % area with deciduous forest cover, Con. wood = % area with coniferous forest, Semi-nat. = % all other land cover).

For the climate, the mean annual temperature for the subset of catchments is slightly higher (7°C vs 6.8°C), however the full range is still covered (Figure 3.22A). While the distribution of mean total rainfall per month is comparable (Figure 3.22B), looking at the mean number of days per month with rainfall above 10 mm, the majority of catchments has now on average 5-6 days per month with this amount of rainfall (Figure 3.22C). Again, the subset includes catchments that cover the full range of conditions.

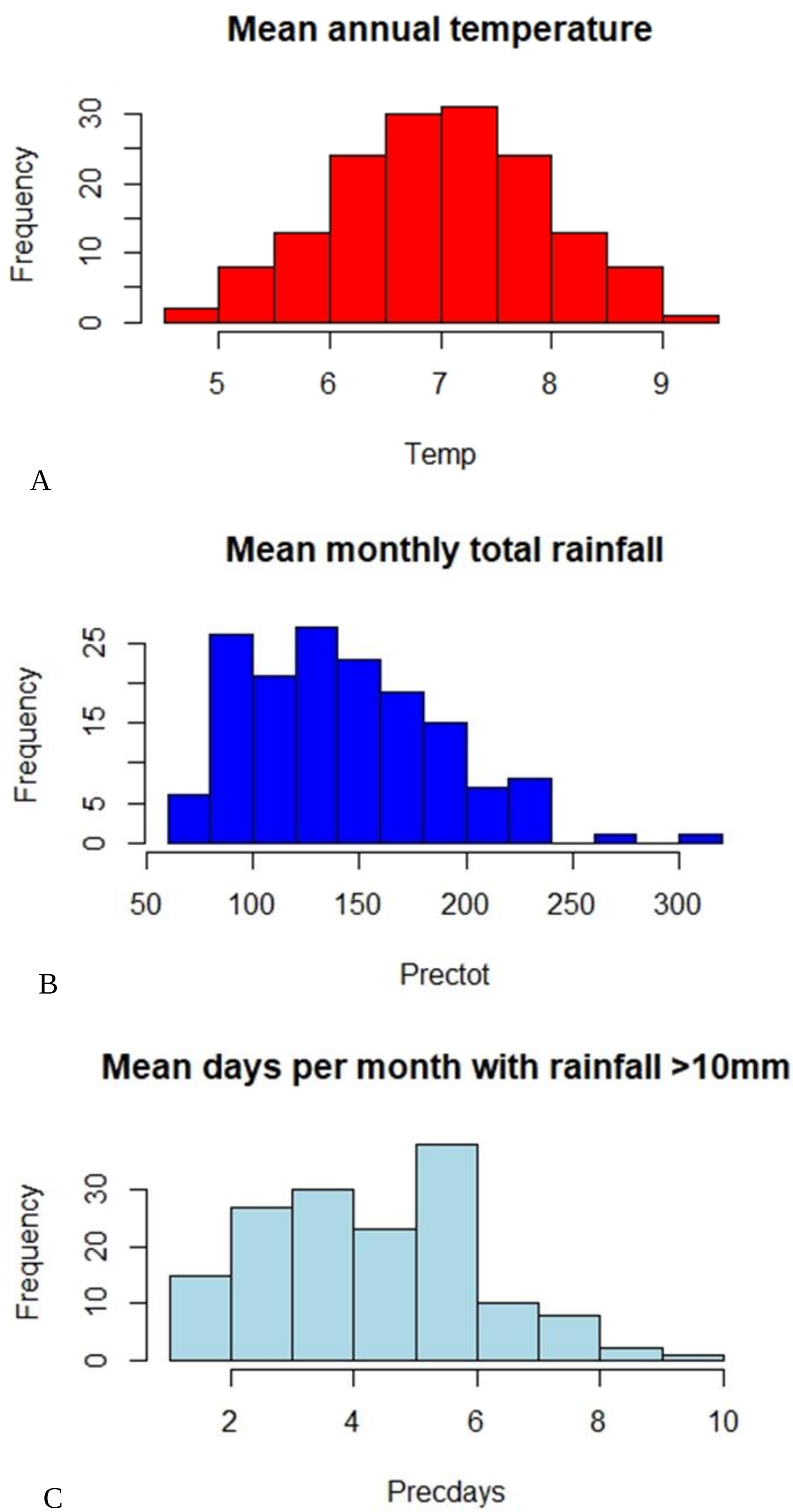


Figure 3.22: Histogram of A. mean annual temperature in °C, B. mean monthly total rainfall in mm, and C. mean number of days per month with more than 10 mm of rainfall of the subset of catchments (n=154).

In summary, some of the variables are slightly differently distributed in the subset of catchments, but most of the variability of conditions is still represented, with the exception of the higher percentages of agricultural land uses in the catchment. The subset is however still considered to be a very good representation.

3.4 Utilisation of data sets in this research

The dataset with the variables describing catchments characteristics originally comprised 40 variables (A.1): Type of source, Slope gentle, Slope moderate, Slope steep, Aspect, Mean elevation, Minimum elevation, maximum elevation, Relief ratio, Elevation relief ratio, Elongation ratio, Area, Distance to sea, Annual temperature range, Conrad's formula, Igneous and metamorphic bedrock, Limestone bedrock, Sandstone bedrock, Other sedimentary bedrock, HOST1, HOST2, HOST3, HOST4, Peat, Eroded peat, Dominant BFI, Dominant SPR, Semi-natural cover, Coniferous cover, Deciduous cover, Mixed woodland cover, Arable cover, Improved grassland cover, Urban area, Water, Conservation designation, Number of deer, Mean monthly rainfall, Mean number of days with more than 10 mm of rainfall, and Mean annual temperature. Some of these variables described the same or similar characteristics and in subsequent analysis, a choice was made which variable to use rather than use all of them. General difficulties with regard to using this dataset in empirical analysis originated from many of the variables, while having numerical values, being in fact categorical as they represented different classes (e.g., bedrock geology, soils, land cover). This also meant that these variables, if all included, would always add up to 100%, although this could be avoided by including only some of these. Many variables were also highly skewed and zero-inflated. This needed to be considered when choosing the statistical method.

Throughout the research, some data became available that were then included in the analysis (see 5 & 6), but were only prepared for the subset of the catchments rather than for the complete set of catchments. These later variables included Topsoil Organic carbon content, Mean number of sheep, Mean number of cattle, Number of septic tanks, Land capability – prime land, Land capability – classes 3.2-5, SER, and AAT. Updated land cover data also became available, which did not include the category "mixed woodland".

As an alternative to “Semi-natural”, the land cover variable “heathland” was generated as a subcategory of “Semi-natural”.

The catchment data of the catchment subset were combined with the water quality dataset comprising summary statistics for each catchment, with concentration median, mean, maximum, minimum, standard deviation, 5th, 25th, 75th and 95th percentile values for each of the eight indicators (A.2). For statistical analysis, only some of these variables were chosen. Additional median concentration values for TOC were later derived for 127 catchments from the catchment subset where data were made available.

Some analyses used individual sample values rather than summary statistics. These included values of colour, TOC, iron, manganese, and turbidity (A.3), colour (A.4), and TOC and *E. coli* in conjunction with climate data (A.5 & A.6). The climate data were daily total rainfall and mean daily temperature data for the period 2010-2016, downloaded as 5 km grid data from the Met Office.

For the risk screening and creation of the risk maps (5.3.6 & 6.3.4), projections in SER, AAT and land capability for a future period 2041-2060 were used. AAT is a metric for the length and intensity of the growing season, and can also be seen to reflect the duration and intensity of bacterial organic matter production. AAT was based upon accumulated degree-days above the standard 5.5ºC threshold for grass. SER is the excess of rainfall over potential evapotranspiration (PET) and is a metric for potential surface runoff, accumulated daily for April-September, with PET calculated by the FAO Penman-Monteith method (Brown, 2017). These metrics and their projections were created and provided by Dr. Iain Brown. For both metrics, the mean value was calculated for the relevant period on a 100 m grid and georeferenced against water intake sources. For the projections in AAT and SER, the UKCP18 perturbed physics ensemble (PPE) was used, with each of the 12 ensemble members representing variants run under different parameterizations, using the 12 km regional model HadRM3 nested within the global model HadGEM3-GC3.05. The models are driven by future GHG concentrations defined by the RCP8.5 scenario (Lowe et al. 2018). RCP8.5 is a high-end scenario, but the climate changes are less influenced by divergence between different emissions scenarios as compared to climate sensitivity (as represented by different climate model parameterizations). Use of RCP8.5 together with HadGEM3/HadRM3, which has relatively high climate sensitivity compared to other climate models, therefore provides an upper end estimate of climate change risk (Hanlon et al., 2021). The UKCP18 PPE model data were further downscaled to match the grid scale of the baseline period climatology using a simple delta-shift (Brown, 2017).

Land capability classification for Scotland has used a similar bioclimate classification in combination with other biophysical variables to define land use potential (Brown et al., 2008). It uses soil moisture deficit rather than SER, but also uses AAT, as in Scotland, temperature and precipitation are the predominant natural constraints on agriculture because many areas are too cool or wet. For the land capability projections, projections in climate parameters derived as described above were used to develop a future bioclimate and land capability scenario. The projections for land capability represent a ‘reasonable worst case’ scenario assuming that irrigation and drainage infrastructure continues to be available so that land use can be optimised according to the capability class for that location.

4. Catchment profiling: Catchment typologies and relationships between catchment characteristics and water quality

Assessment 1.1 looks into the relationships between catchment characteristics and water quality, the possibility for catchment profiling, and current concerns over raw water quality in Scotland. The first step in the analysis to arrive at a risk screening that enables a step-by-step integration of climate change impacts into the existing raw water risk assessments and management procedures therefore considers all identified eight water quality indicators to more broadly reflect on water quality, and investigates if the data allow a degree of profiling, meaning identifying types of catchments with similar water quality characteristics stemming from similar catchment conditions. This builds capacity to understand catchment-water quality connections, to identify major concerns and area of concerns, and to prioritize areas or indicators to investigate further. It therefore particularly relates to stage 1 in Figure 1.1, by identifying how different water quality issues relate to different pressures, connecting this to particular conditions that make catchments potentially vulnerable to these pressures, and look at spatial distribution to understand patterns of exposure and vulnerability that may indicate how temporal changes will affect different types of catchments.

This chapter first identifies possible approaches to investigate the data with these objectives in mind (4.1), describes methods chosen (4.2), presents results (4.3) and discusses them with regard to their meaning for risk factors for water quality, risk management and assessment (4.4), and how they inform further steps (4.5).

4.1 Background

There are numerous examples of a wide range of multivariate statistics used to understand patterns in and controls over surface or ground water quality. It is also common to combine catchment characteristics data with water quality data to enable predictions for water quality e.g., for transferring findings to data-sparse catchments. Especially for the former aspect of recognising patterns and commonalities, methods that allow a form of grouping or reduction of complexity are favoured, while forms of linear models or regression is often used for the latter aspect of predicting water quality outcomes. Several methods are often combined to gain a more comprehensive understanding.

Methods like Principal Component Analysis (PCA) and Factor Analysis (FA) allow identifying patterns in the correlations between different variables. These methods assume that there are latent factors (which cannot be or haven't been measured directly but are described by several other variables) underlying the observed data (FA), or try to identify composites of the variables (PCA). They seek to reduce complexity in the data and have the advantage that results could be used in further statistical modelling that benefits from a reduced set of variables (Field et al., 2012). Especially PCA is a widely used approach to analyse water quality variability. For example, Bengraïne & Marhaba (2003) used PCA on a water quality data set with samples of 20 water quality indicators from 12 sites in the Pessaic River in New Jersey, in order to understand differences in physical and chemical characteristics and associate this with natural and anthropogenic influences. Ferrier et al. (2001) used PCA alongside correlation analysis to identify different types of catchments and their associated water quality in Scotland, identifying three types of catchments ('urban', 'arable' and 'improved'). There are numerous further examples of the use of PCA, alone or in combination, in water quality research around the world (e.g., Arora & Keshari, 2021; Brontowiyono et al., 2022; Kazi et al., 2009; Kowalkowski et al., 2006; Li et al., 2009; Ouyang, 2005; Parinet et al., 2004; Perona et al., 1999; Selle et al., 2013).

Redundancy analysis (RDA) is related to PCA by combining it with linear regression in order to examine the relationships between two data sets of which one corresponds to response variables, the other to explanatory variables. This is also a method that has found application in research on water quality especially for understanding the impact of catchment characteristics and landscape patterns (e.g., Ding et al. 2016; Ou & Wang, 2011; Shen et al., 2015; Shi et al., 2017).

While PCA, FA and RDA can identify underlying commonalities in water quality and be interpreted to relate these to environmental factors, cluster analysis (CA) directly attempts groupings/classifications based on similarities across multiple variables. CA should result in high homogeneity within a cluster and high heterogeneity between clusters. There are different clustering techniques; popular methods in water quality research are hierarchical clustering (e.g., Kazi et al., 2009; Shen et al., 2011; Singh, Malik et al., 2004; Vadde et al., 2018), or clustering with the k-means algorithm (e.g., Chang et al., 2011; Haggarty et al., 2012; Mandel et al., 2015; Shareef et al., 2014).

In contrast to CA, which analyses the data to find groups with maximum differences between them, in discriminant analysis (DA), membership of the groups is already known (pre-defined) and the analysis aims at determining the variables that best differentiate between the groups. DA has been successfully applied in water quality research to reduce complexity in data sets and provide guidance about how to make monitoring more efficient (Bhat & Pandit, 2014; Giao & Nhien, 2021; Singh, Malik et al., 2004; Varol, 2020).

More direct methods to relate water quality to catchment characteristics, especially if the focus is on one or a few particular water quality indicators rather than a range, includes forms of regression. With multiple linear regression, variables describing catchment characteristics can be used to predict a chosen metric for water quality indicators (e.g., means; Davies & Neal, 2004; Helliwell et al, 2007; Monteith et al., 2015; Rothwell et al., 2010). Logistic regression can be used to predict the likelihood of falling into a category (such as “good” water quality status under the WFD) due to catchment characteristics (Donohue et al., 2006).

Geographically weighted regression embeds location data into the regression parameter to account for local variation in relationships between independent and dependent variables, and for spatial autocorrelation. There are several examples where it has been applied in a water quality context (Chen et al., 2016; Pratt & Chang, 2012; Tu & Xia, 2008; Tu, 2011).

Regression trees offer some advantages over linear regression to predict an outcome in a dependent variable from several independent variables by being non-parametric, therefore able to deal with non-linear and complex relationships, and its visualisation is also easy and intuitive to interpret (Breiman et al., 1984). Regression trees have been extended by techniques of multi-target regression, making them able to predict several variables with one model. Multi-target predictive clustering trees (MTPCTs) have been shown to be effective in building predictive models (Demšar et al., 2006; Struyf & Džeroski, 2006) and have been applied to water quality assessment (Nikoloski et al., 2021).

When developing models to predict water quality outcomes from catchment characteristics, the exact choice of independent variables depends on the dependent variable and knowledge about processes that influence its concentrations, as well as data availability and quality. However, models typically include some representation of topography, land cover and/or use, soils, and climate. For example, typical variables include slopes; percentage cover of forests, wetlands, urban areas, or agricultural areas; baseflow index or standard percentage runoff; annual average rainfall; mean or maximum air temperature; soil organic carbon pool; or altitude (Cool et al., 2014; Helliwell et al., 2007; Monteith et al., 2015; Pratt & Chang, 2012; Rothwell et al., 2010).

4.2 Methods

Surface water quality depends on the kind and size of pollution sources in the catchment, how and when they are connected to the surface water, and how pollutants are processed in transit and within the water body. Understanding different catchment properties and sensitivities, and how specific pressures act on these, helps to anticipate how changing pressures might lead to impacts in different catchments. A typology of catchments would allow transferring findings and results from one catchment to others of similar profile. The overall aim of this first analysis of the raw water data was therefore to identify and understand patterns, to find a general typology of catchments according to water quality profiles, and to relate these to catchment properties.

As there are known to be connections between the eight indicators of water quality (3.2.3), it was first investigated if complexity can be reduced and if this variability in water quality can be related to specific pressures. PCA was used to this purpose, and catchment characteristics were superimposed to the PCA solution as supplementary variables to better relate the identified optimal linear combinations (principal components) to catchment sensitivities. RDA was additionally run to examine the influence of catchment characteristics on the variability in the water quality data more directly. Cluster analysis served the purpose to directly establish if water quality data allowed a typology of catchments and if this typology reflects findings from the PCA and RDA. The analysis of clusters also allowed to explore geographic patterns of catchment sensitivity and risk factors. To test relationships between catchment characteristics and specific water quality variables, MTPCT were run for each water quality indicator. Multi-target regression was chosen to allow developing a model that simultaneously predicts the median and the 95th percentile of the water quality indicator to reflect both baseline and extremes.

Data preparation and statistical analysis were done in R. Median values were used in all applied methods with the addition of the 95th percentile in the MTPCTs. For the PCA (see B.1 for R code) and RDA (B.2), water quality data except pH values were log transformed to bring them closer to a normal distribution. Included catchment characteristics were: Area in km^2, Relief ratio, Percentage of gentle slopes, Percentage of steep slopes, Percentage of limestone, Percentage of sandstone, BFI, SPR, Topsoil organic carbon content, Percentage of arable area, Percentage of improved grassland, Percentage of urban area, Percentage of deciduous forest, Percentage of heathland, Number of septic tanks, Average number of sheep in the parish, Average number of cattle in the parish, Mean annual temperature, Mean monthly total rainfall, and Mean days with >10 mm of rain per month. These catchment characteristics were projected onto the biplot space of the first two PCs. Their biplot coordinates were derived from correlation with the PCs (Graffelman & Aluja-Banet, 2003).

Backward variable elimination was done during the RDA, and non-significant catchment characteristics were removed from the model. To reduce collinearity, Percentage of steep slopes and Percentage of gentle slopes were removed manually, as it was chosen to represent topography through Relief ratio. To assess possible issues with multicollinearity, variance inflation factors (VIF) were derived.

As a method for the cluster analysis, the partitioning around medoids (PAM; Kaufman & Rousseeuw, 1990) clustering algorithm was used (B.3). PAM is a partitional clustering technique that is similar to k-means algorithm but using actual data points as cluster centres. Its advantage is that it minimises the influence of outlying observations. The standardized (z-transformed) catchment medians were used, and their dissimilarity measured using Euclidean distances. As the number k of clusters needs to be pre-determined, the clustering was run several times with varying numbers of clusters and the quality of the clustering was assessed with the average silhouette width (Kodinariya & Makwana, 2013). The clustering structure with the best silhouette appearance and width was chosen for further analysis. The significance of the overall differences between the clusters were tested with Kruskal-Wallis tests and pairwise differences with Wilcoxon tests.

For the eight MTPCTs, the software package Clus (Struyf et al., 2011) was used. Performance of the MTPCTs was assessed through the root mean square error (RMSE) and R^2 values from training and testing using 10-fold cross-validation.

4.3 Results

4.3.1 PCA

Eighty percent of the data variance are explained by the first three principal components (PCs) of the PCA (46% PC1, 20% PC2, and 14% PC3). The biplot of the first two PCs (Figure 4.1) shows that catchments are relatively evenly scattered around the origin. The first PC is mostly associated with median concentrations of turbidity, aluminium, iron, manganese, and colour, whereas the second PC associated with pH, coliform and *E. coli* (Table 4.1). The projection of the catchment variables show that Percentage of steep slopes and Relief ratio negatively correlate with PC1, whereas Percentage of gentle slopes positively correlate. Average number of cattle and sheep also positively correlate, as do improved grassland, arable areas and urban areas, and heathland cover negatively correlates. The two precipitation variables also negatively correlate. For PC2, catchment variables showing strong positive associations are the land cover variables except heathland, the number of septic tanks, as well as Percentage of limestone and BFI. Topsoil organic carbon content negatively correlates.

This means that catchments in the upper part of the biplot are associated with higher median concentrations of *E. coli* and coliform bacteria, as well as showing higher percentages of improved grassland, arable and urban areas. Catchments in the right side of the biplot show increased median concentrations of the metals, colour, and turbidity, and are associated with lower amounts of precipitation, higher livestock densities, and gentler reliefs. Towards the lower part of the biplot, catchments are associated with lower pHs and higher topsoil organic carbon contents.

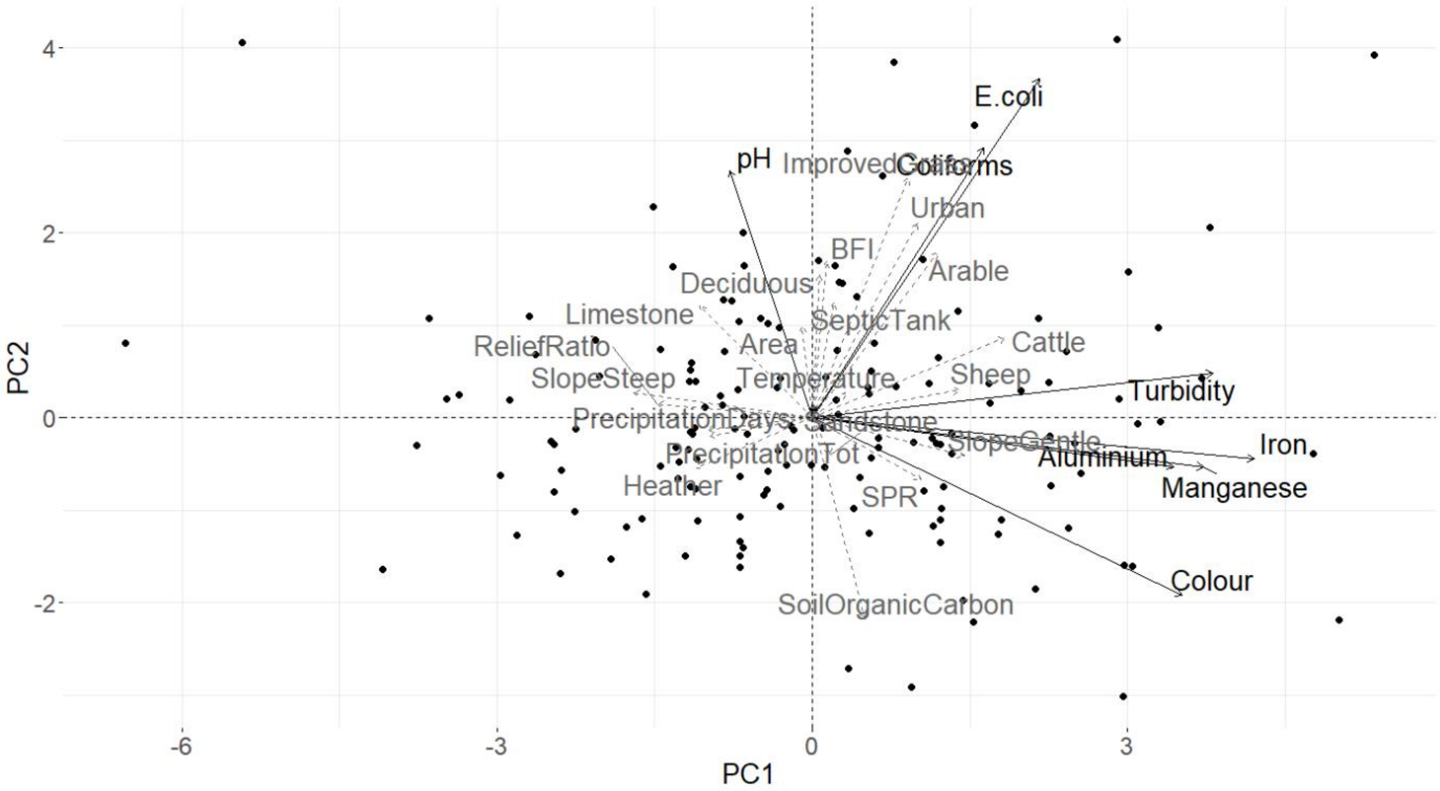


Figure 4.1: Biplot of first two Principal Components (PCs) from the PCA on log-transformed catchment median concentrations (n=154) of water quality indicators, with catchment characteristics superimposed as supplementary variables. Water quality parameters symbolised as black rays, catchment characteristics as grey dashed rays, individual catchments as points. Points/variables in the right of the plot associate positively with PC1, to the left – negatively. Points/variables in the top associate positively with PC2, in the bottom – negatively. The longer the arrow/further from the plot origin the point, the stronger the association. pH = median surface water pH, Coliforms in CFU in 100 ml, *E. coli* in CFU in 100 ml, Turbidity in NTU, Aluminium in µg Al/l, Iron in µ Fe/l, Manganese in µ Mn/l, Colour in mg/l Pt/co. Area = Area in km^2, ReliefRatio = Relief ratio, SlopeGentle = Percentage of gentle slopes, SlopeSteep = Percentage of steep slopes, Limestone = Percentage of limestone, Sandstone = Percentage of sandstone, BFI = Average baseflow index, SPR = Average standard percentage runoff, SoilOrganicCarbon = Average topsoil organic carbon content, Arable = Percentage of arable area, ImprovedGrass = Percentage of improved grassland, Urban = Percentage of urban area, Deciduous = Percentage of deciduous forest, Heather = Percentage of heathland, SepticTank = Number of septic tanks, Sheep = Average number of sheep in the parish, Cattle = Average number of cattle in the parish, Temperature = Mean annual temperature, PrecipitationTot = Mean monthly total rainfall, PrecipitationDays = Mean days with >10 mm of rain per month.

Table 4.1: Loadings table for the first three Principal Components (PCs) from the PCA on log-transformed catchment median concentrations of water quality indicators, and catchment characteristic correlations to PCs. Statistically significant correlations are indicated by ***(p<.001), **(p<.01), or *(p<.05).

	PC1	PC2	PC3
Water quality	Loadings		
Aluminium	0.39***	-0.09	-0.12
Colour	0.39***	-0.33***	-0.09
Iron	0.48***	-0.08	-0.02
Manganese	0.42***	-0.09	0.33***
pH	-0.09*	0.46***	0.66***
Turbidity	0.43***	0.08	0.33***
Coliform	0.18***	0.50***	-0.57***
E. coli	0.24***	0.63***	-0.09
Catchment characteristics	Correlations		
Deciduous woodland	0.01	0.36*	0.07
Improved grassland	0.2*	0.56***	-0.32***
Arable	0.26**	0.39***	-0.04
Urban	0.22**	0.46***	0.00
Heather	-0.24**	-0.12	-0.23*
Cattle	0.4***	0.18*	0.31***
Sheep	0.31***	0.07	0.45***
Septic tanks	0.05	0.27***	-0.11
Relief ratio	-0.32***	0.03	-0.24**
Gentle slope	0.32***	-0.09	0.14
Steep slope	-0.37***	0.06	-0.06
Limestone	-0.23	0.26**	0.12
Sandstone	0.04	-0.09	0.2*
Topsoil organic carbon	0.11	-0.48***	-0.23*
Average BFI	0.06	0.36***	0.15
Average SPR	0.2*	-0.1	0.04
Mean monthly total rainfall	-0.21**	-0.04	-0.29***
Mean number of days per months with >10 mm precipitation	-0.22**	-0.03	-0.24**

4.3.2 RDA

A correlation matrix for the catchment variables showed that correlations are highest between Relief ratio and the slope percentages, between the two precipitation variables, between the land use variables (except heathland), between Percentage of improved grassland and Topsoil organic carbon, and between the precipitation variables and the topography variables (Figure 4.2).

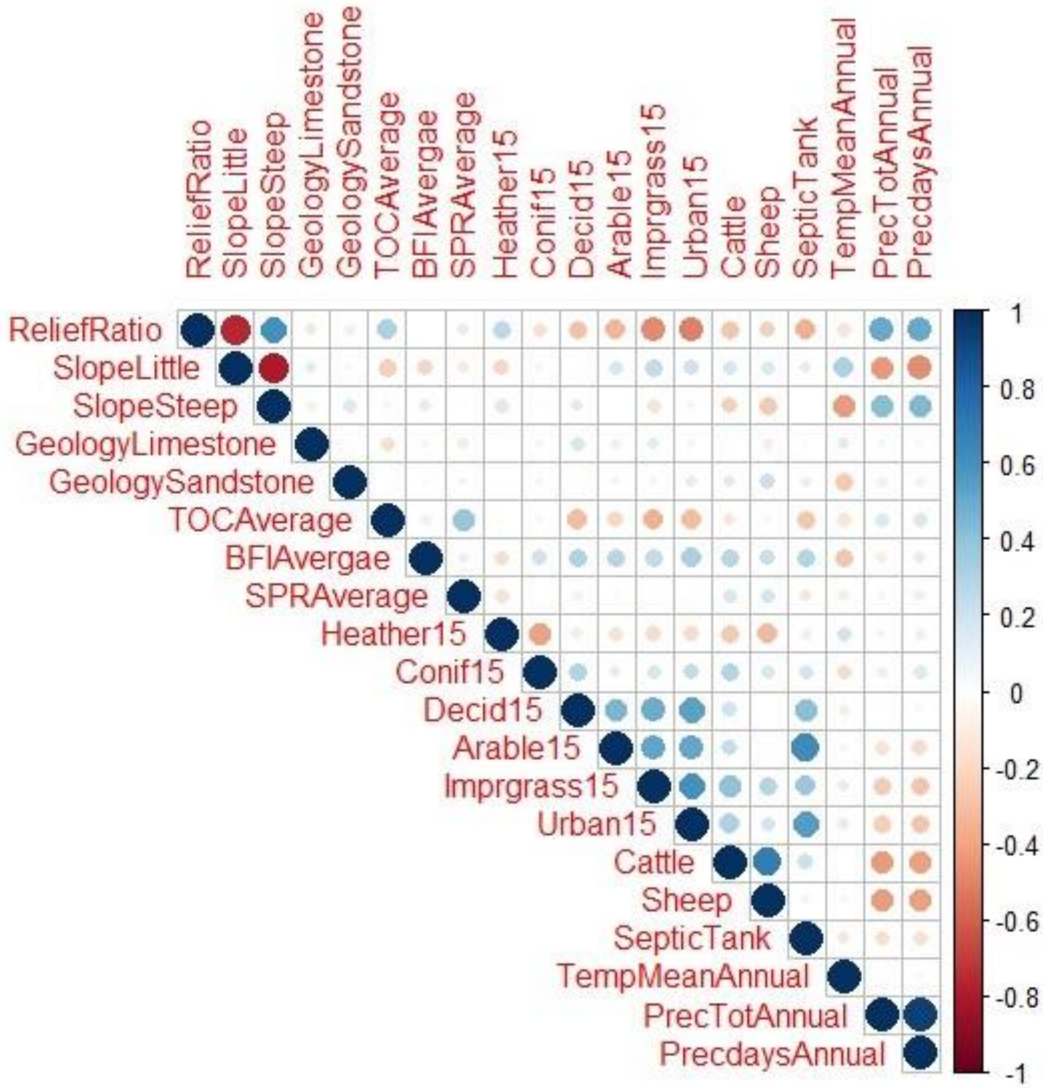


Figure 4.2: Visualisation of the correlation matrix for constraining variables used in the RDA with log-transformed catchment median concentrations (n=154) of water quality indicators as response variables, based on Spearman's rank correlation. ReliefRatio = Relief ratio, SlopeLittle = Percentage of gentle slopes, SlopeSteep = Percentage of steep slopes, GeologyLimestone = Percentage of limestone, GeologySandstone = Percentage of sandstone, TOCAverage = Average topsoil organic carbon content, BFIAvergae = Average baseflow index, SPRAverage = Average standard percentage runoff, Heather15 = Percentage of heathland, Conif15 = Percentage of coniferous forest, Decid15 = Percentage of deciduous forest, Arable15 = Percentage of arable area, Imprgrass15 = Percentage of improved grassland, Urban15 = Percentage of urban area, SepticTank = Number of septic tanks, Sheep = Average number of sheep in the parish, Cattle = Average number of cattle in the parish, TempMean Annual = Mean annual temperature, PrecTotAnnual = Mean monthly total rainfall, PrecdaysAnnual = Mean days with >10 mm of rain per month.

During the RDA, non-significant variables were removed, leaving Relief ratio, Percentage of gentle slopes, Percentage of steep slopes, BFI, Topsoil organic carbon content, Percentage of improved grassland cover, Percentage of arable area, Percentage of urban area, Average numbers of cattle and sheep, and Mean number of days with precipitation above 10 mm per month. Some of the variables that show higher correlations with others were thereby removed, but for the final model, Percentage of steep slopes and Percentage of gentle slopes were additionally removed to avoid multicollinearity, and it was judged that topography would be adequately represented by Relief ratio. As VIFs for the final variables ranged between 1.2 and 2.3 (Table 4.2), collinearity was not considered to be an issue, as a commonly accepted threshold value for VIF is 10 (Alin, 2010).

Table 4.2: VIFs for constraining variables used in the RDA with log-transformed catchment median concentrations (n=154) of water quality indicators as response variables and catchment characteristic variables as constraining variables.

Variable	**VIF**
Relief ratio	1.3
Topsoil organic carbon	1.5
BFI	1.2
Percentage of improved grassland	2.3
Percentage of arable area	1.9
Percentage of urban area	2.1
Average number of sheep	1.6
Average number of cattle	1.8
Mean number of days with precipitation above 10 mm per month	1.5

The first two RDA axes constrained 53% and 32% respectively of total variance of water quality parameters, which corresponded to 17% and 10% of the overall variance.

All water quality variables associate positively with the first RDA axis (RDA1), with only colour having a negligible score (Table 4.3). Of the catchment variables, Percentage of improved grassland, arable and urban areas, mean numbers of sheep and cattle, and BFI also associated positively with RDA1. This indicates that increases in these variables correlate to increased median concentrations for coliform bacteria and *E. coli*, turbidity, and the metal indicators (Figure 4.3). A negative relationship can be observed for Relief ratio, Topsoil organic carbon and mean number of days per month with precipitation above 10 mm.

Colour, turbidity, aluminium, iron, and manganese negatively associated with the second axis of the RDA (RDA2), whereas pH, *E. coli* and coliform bacteria positively associated. This indicates that higher median concentrations especially in colour, iron and manganese usually go together with more acidic pH. Topsoil organic carbon was also negatively associated with RDA2, and Relief ratio positively, indicating that higher colour concentrations occur on organic soils and more gentle reliefs. Agricultural and urban land uses also associated positively, though less strongly, with RDA2, however numbers of sheep and cattle associated negatively. Arable and urban areas as well as improved grassland could therefore be explanatory factors for bacteria, while livestock numbers could contribute especially to turbidity.

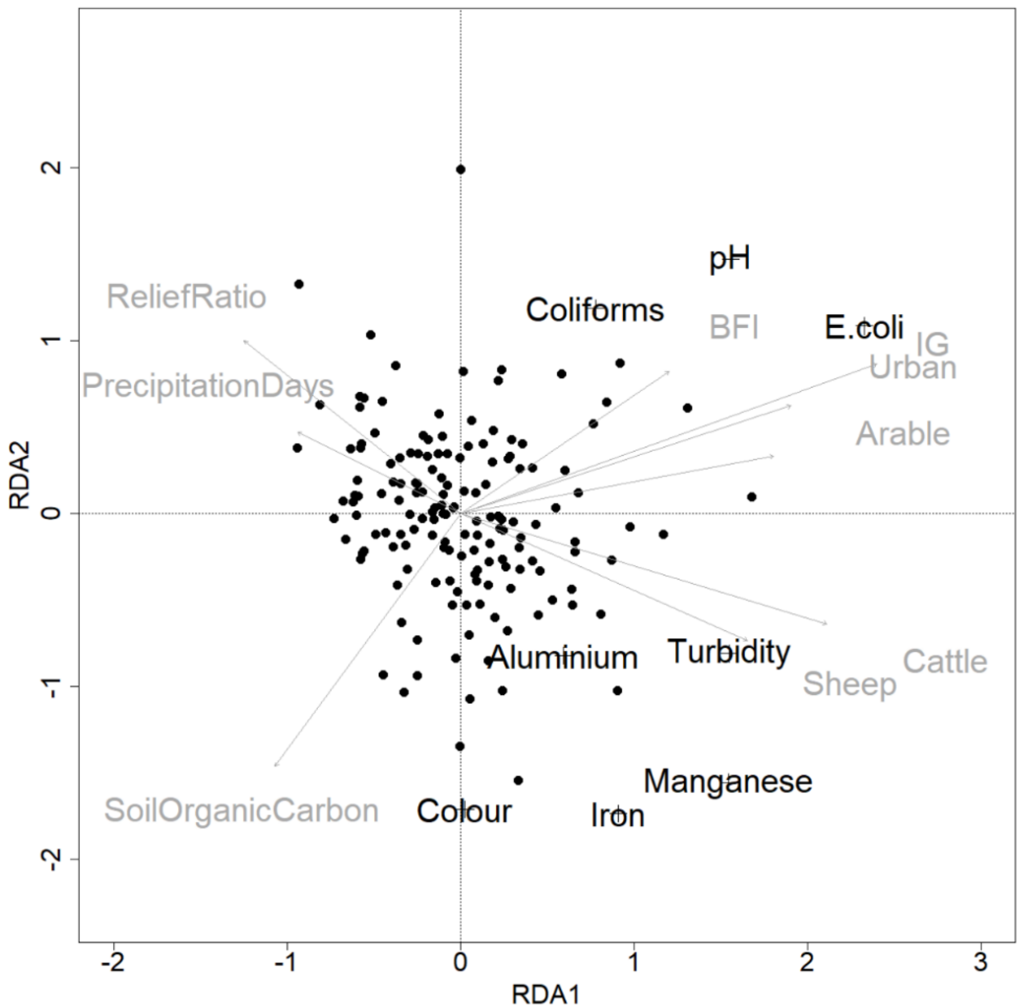


Figure 4.3: Triplot of the RDA with log-transformed catchment median concentrations (n=154) of water quality indicators as response variables and catchment characteristic variables as constraining variables. Water quality parameters symbolised in black writing, catchment characteristics as constraints symbolised as grey rays, and individual catchments symbolised as points. Points/variables in the right of the plot correlate positively with RDA1, to the left – negatively. Points/variables in the top correlate positively with RDA2, in the bottom – negatively. The longer the arrow/further from the plot origin the point, the stronger the correlation. Water quality pH = median surface water pH, Coliforms in CFU in 100 ml, *E. coli* in CFU in 100 ml, Turbidity in NTU, Aluminium in µg Al/l, Iron in µ Fe/l, Manganese in µ Mn/l, Colour in mg/l Pt/co. ReliefRatio = Relief ratio, BFI = Average baseflow index, SoilOrganicCarbon = Average topsoil organic carbon content, Arable = Percentage of arable area, IG = Percentage of improved grassland, Urban = Percentage of urban area, Sheep = Average number of sheep in the parish, Cattle = Average number of cattle in the parish, PrecipitationDays = Mean days with >10 mm of rain per month.

Table 4.3: Score table for the first three axes from the RDA with log-transformed catchment median concentrations (n=154) of water quality indicators as response variables and catchment characteristic variables as constraining variables.

	RDA1	**RDA2**	**RDA3**
Water quality	Scores		
Aluminium	0.24	-0.26	0.14
Colour	0.01	-0.54	0.1
Iron	0.37	-0.55	0.18
Manganese	0.64	-0.5	-0.17
pH	0.64	0.47	-0.4
Turbidity	0.64	-0.26	-0.15
Coliform	0.32	0.38	0.45
E. coli	0.96	0.35	0.22
Catchment characteristics	Biplot scores		
Improved grassland	0.79	0.37	-0.13
Arable	0.6	0.14	0.55
Urban	0.63	0.27	0.49
Cattle	0.7	-0.27	-0.09
Sheep	0.55	-0.31	-0.55
Relief ratio	-0.41	0.42	0.15
Topsoil organic carbon	-0.35	-0.62	0.33
BFI	0.4	0.35	0.02

4.3.3 Cluster analysis

The clustering was run on the standardized median concentrations of the eight water quality indicators for a number k of clusters ranging from 3 to 12. The best silhouette width was achieved for 8 clusters, however this clustering consisted of two clusters with only one member, while the remaining clusters strongly resembled the clusters when the pre-defined number of clusters was 5, 6, or 7 (Table 4.4). The best silhouette width where all clusters had more than one member was achieved when k=5 (Figure 4.4), and it was judged to be the most useful clustering for interpretation of water quality in relation to catchment characteristics.

The best clustering structure still rendered only an overall very weak clustering structure and the clusters were very unequal in size. Overlap can be observed especially between cluster 1 and 2, and cluster 4 and 5 have high heterogeneity (Figure 4.5).

Table 4.4: Silhouette width and size of clusters for different k when clustering catchments (n=154) on standardised median concentrations for the eight water quality indicators using PAM.

k	Silhouette width	Size of clusters
3	0.16	73, 36, 45
4	0.18	65, 32, 40, 17
5	0.21	63, 30, 40, 16, 5
6	0.22	63, 30, 40, 16, 4, 1
7	0.23	63, 30, 40, 15, 4, 1, 1
8	0.24	63, 24, 39, 10, 12, 4, 1, 1
9	0.23	54, 24, 20, 36, 11, 4, 3, 1, 1
10	0.21	52, 14, 16, 35, 11, 17, 4, 3, 1, 1
11	0.18	28, 14, 16, 35, 11, 24, 17, 4, 3, 1, 1
12	0.19	28, 14, 14, 35, 10, 24, 15, 5, 4, 3, 1, 1

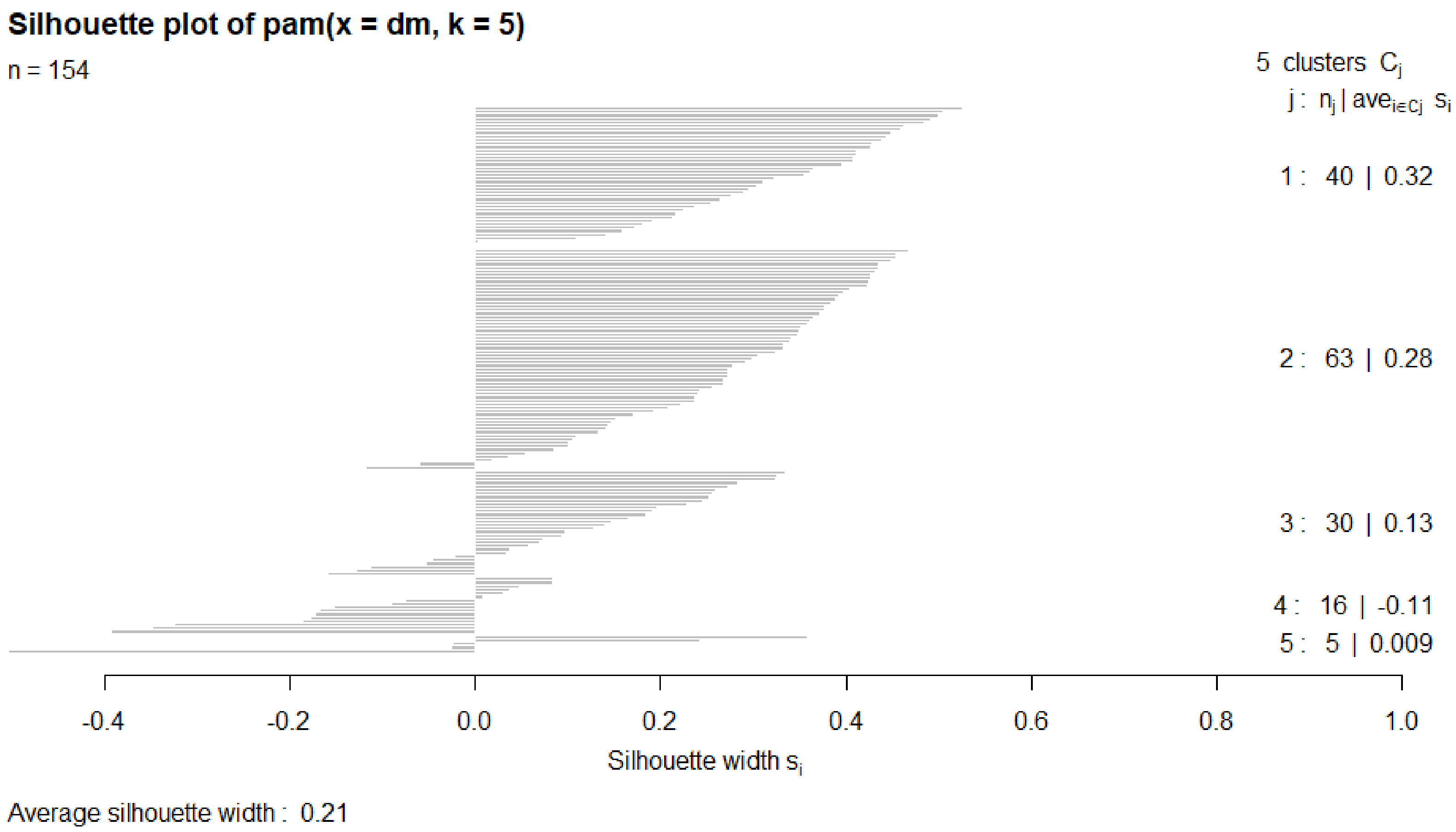


Figure 4.4: Silhouette plot of clusters with k=5 when clustering catchments (n=154) using PAM on standardised median concentrations of eight water quality indicators. Grey bars represent individual catchments, the length and direction of the bar indicate how well the catchment fits into the cluster. The further the bar extends to the right (positive), the further away the cluster is from neighbouring clusters. Catchments with bars extending into negative numbers are potentially allocated to the wrong cluster. Numbers at the right side of the plot inform about the number of catchments in the cluster and the average silhouette width for the cluster, indicating that especially clusters 4 and 5 are very heterogenous with many badly fitting catchments.

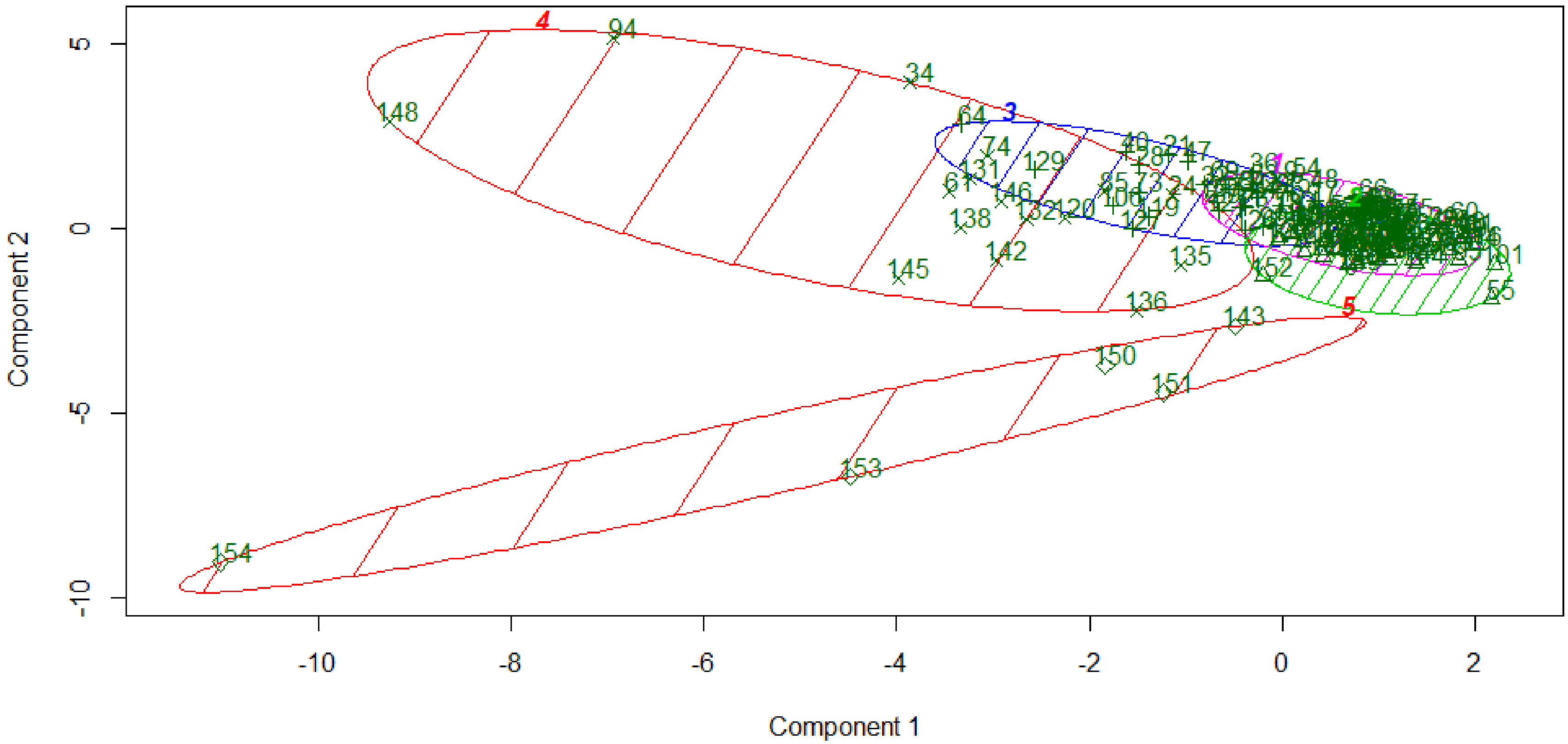


Figure 4.5: 2D representation of the cluster solution for k=5 when clustering catchments (n=154) using PAM on standardised median concentrations of eight water quality indicators. Catchments are symbolised in green, plotted on the first two Principal Components of the standardised water quality median concentration data. Catchments in different clusters are symbolised with different symbols and ellipses drawn around catchments in the same cluster, representing the cluster structure. This visualisation reiterates the poor homogeneity of clusters 4 and 5 and shows significant overlap between cluster 1 and 2 for these 2 dimensions.

However, median concentrations of the water quality parameters were overall statistically significantly different between clusters ($p<0.05$; Figure 4.6).

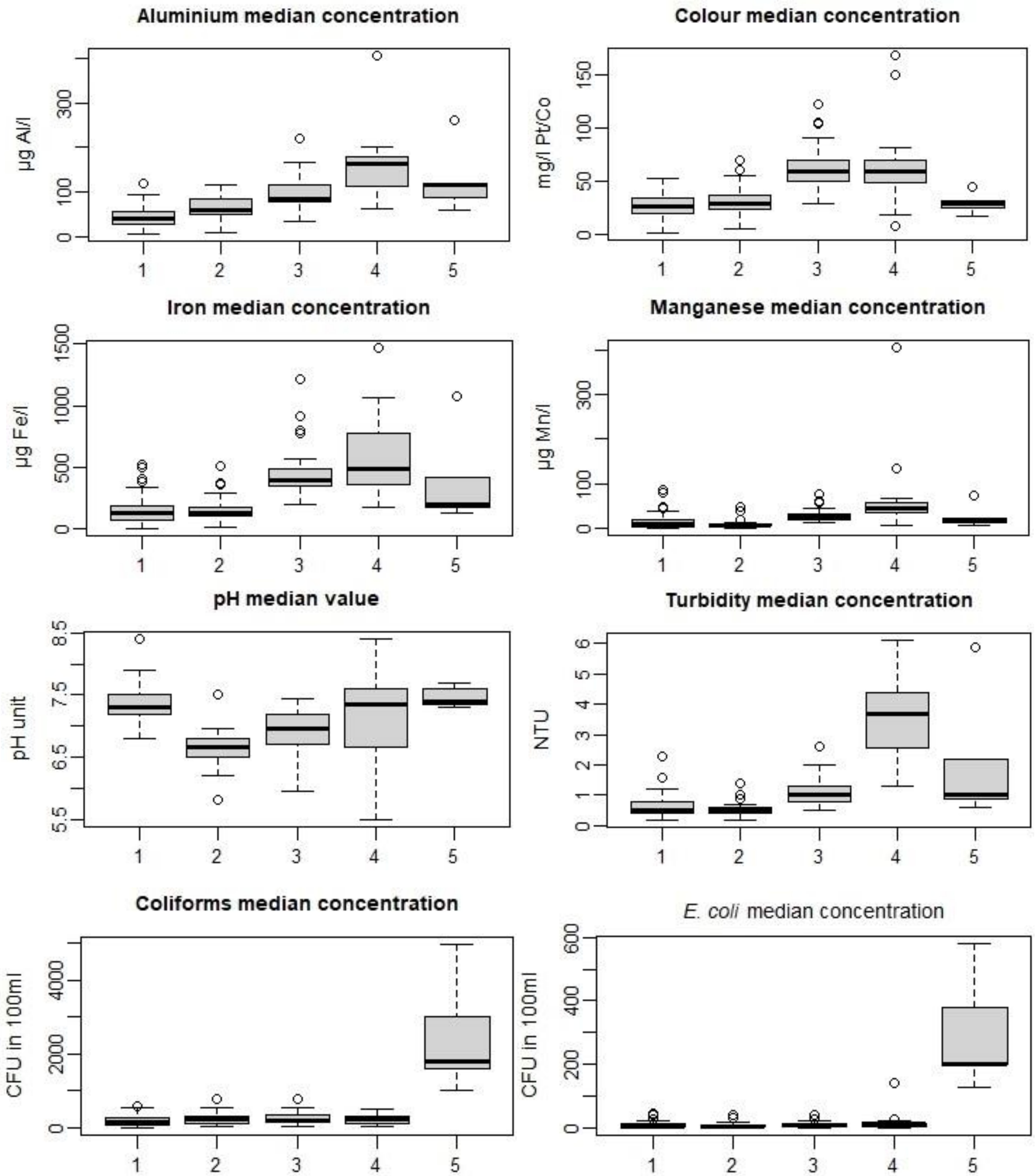
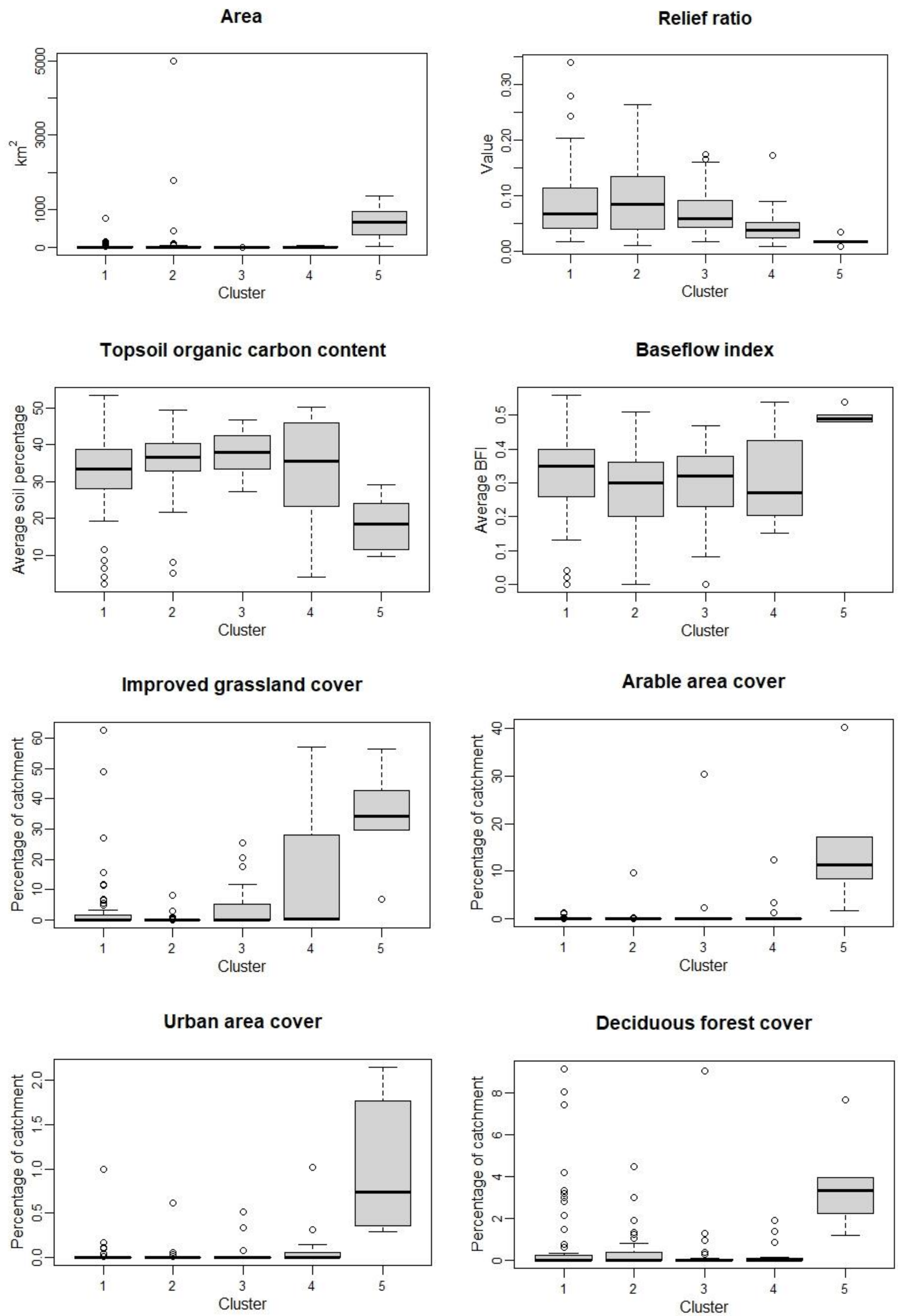


Figure 4.6: Boxplots showing the distribution of median catchment concentrations per water quality indicator per cluster as identified when k=5 when clustering catchments (n=154) using PAM on standardised median concentrations of eight water quality indicators.

There are also significant ($p<0.05$) differences between catchment characteristics per cluster, in Area, Relief ratio, BFI and SPR, Topsoil organic carbon content, Percentage of arable area, urban area, improved grassland, deciduous forest and heathland cover, numbers of sheep and cattle, number of septic tanks, and precipitation (total and days >10 mm; Figure 4.7). Differences are mainly apparent between cluster 5 and the other clusters. The five catchments in cluster 5 have high median concentrations of coliforms and *E. coli*, are larger, have gentler reliefs with smaller organic carbon pools in the soils and a higher BFI, are more intensively used and receive lower amounts of rainfall. Catchments with steeper reliefs are found in clusters 1 and 2. The catchments in these clusters generally show lower median concentrations for all indicators. Cluster 2 contains mostly the most natural and least used catchments and tends to have lower pH medians compared to Cluster 1. Cluster 3 includes catchments with higher colour median concentrations, and only contains catchments with higher amounts of organic carbon in the topsoil, although cluster 1, 2, and 4 also have many catchments with bigger carbon pools. Cluster 4 includes catchments with very high median concentrations for all indicators except bacteria. These seem to be drier catchments with less semi-natural land cover and higher numbers of livestock.


Area
km²
Cluster
Relief ratio
Value
Cluster
Topsoil organic carbon content
Average soil percentage
Cluster
Baseflow index
Average BFI
Cluster
Improved grassland cover
Percentage of catchment
Cluster
Arable area cover
Percentage of catchment
Cluster
Urban area cover
Percentage of catchment
Cluster
Deciduous forest cover
Percentage of catchment
Cluster

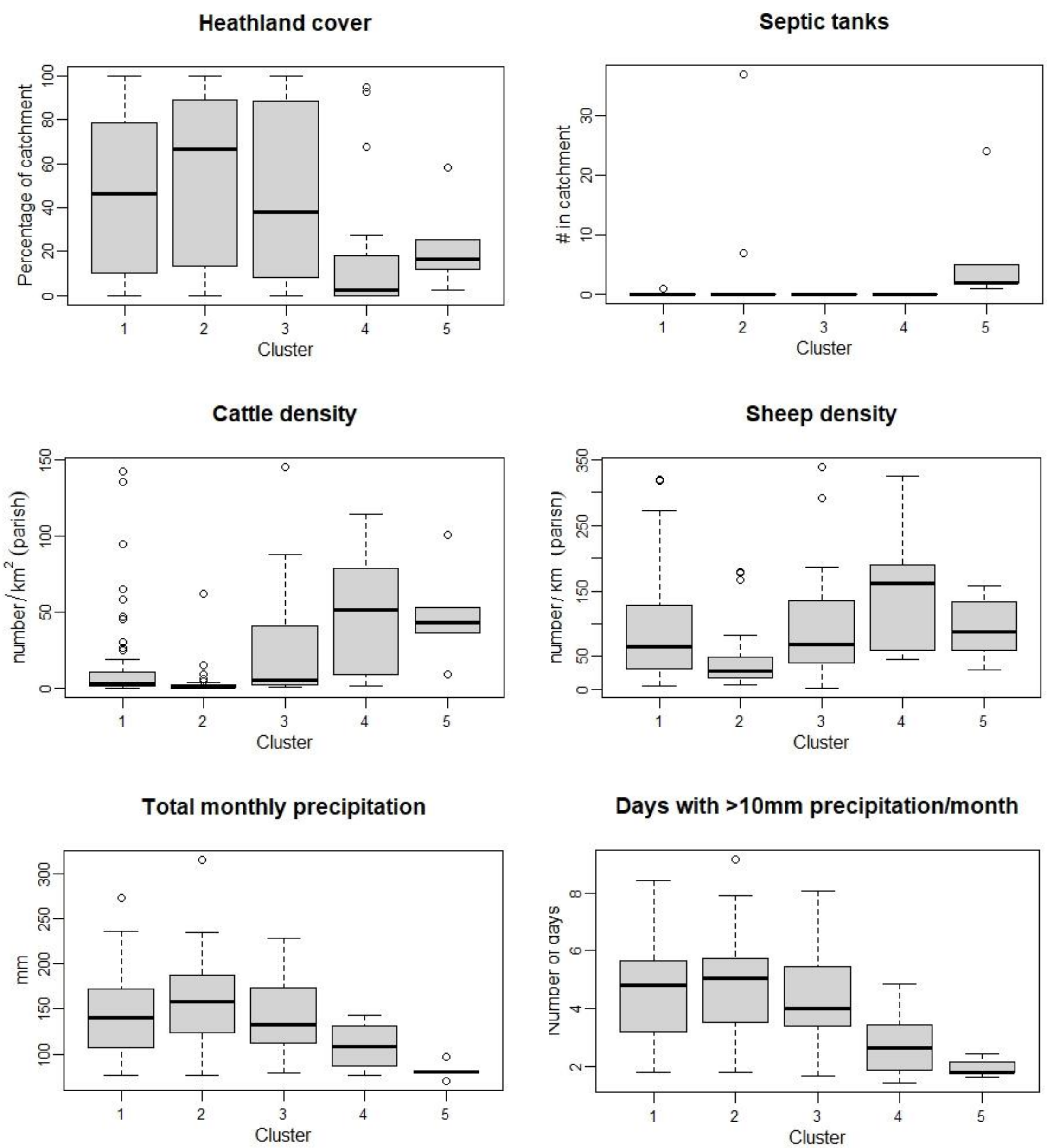


Figure 4.7: Boxplots showing the distribution of catchment characteristics per cluster as identified when k=5 when clustering catchments (n=154) using PAM on standardised median concentrations of eight water quality indicators.

Considering the spatial distribution illustrated in Figure 4.8 reveals that catchments in cluster 1 seem to occur all over Scotland, whereas catchments in cluster 2 are found predominantly in the Northwest, in cluster 3 mainly in the West, in cluster 4 in the East, the South and on the Orkney islands, and the five catchments in cluster 5 are all in the Northeast. Apart from this overall pattern, it is observable that catchments in close proximity belong to different clusters, suggesting that local factors play a role in determining water quality.

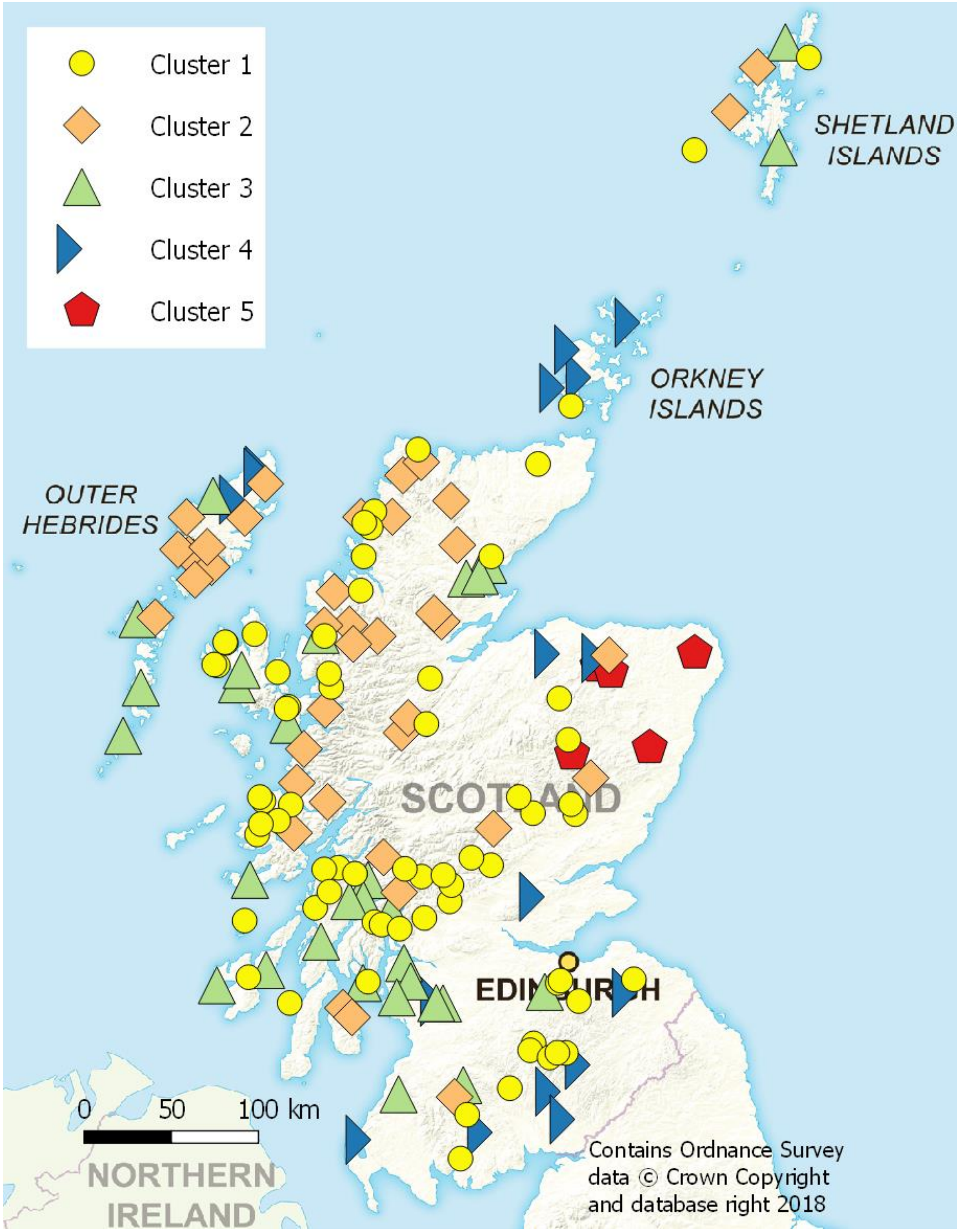


Figure 4.8: Spatial distribution of catchment clusters as identified when k=5 when clustering catchments (n=154) using PAM on standardised median concentrations of eight water quality indicators.

4.3.4 Regression tree modelling

Eight MTPCTs were produced which varied in their complexity. Three MTPCT were very simple with only one separator (for colour, manganese, and coliforms). For all trees, root mean square error (RSME) and R^2 values for training and testing indicated that the predictive capability of the models is very low and that all of them suffered from overfitting to the training set (Table 4.5). However, the models can still be examined for their usefulness in understanding relationships between water quality and catchments characteristics.

The first separator in the MTPCT for aluminium is cattle density, with median and peak concentrations increasing where more cattle is present. The next separator is coniferous forest cover, again with higher median concentrations where more than half of the catchment is covered with coniferous forest, but comparatively lower peak concentrations. A few catchments are then separated with lower mean annual temperatures that show a medium predicted value for median concentration but high peak concentration. The last separator is mean total rainfall per month, with the catchments seeing higher rainfall having the lowest median and peak concentrations. Almost half of the catchments fall into this group.

The MTPCT for colour has only one separator, topsoil organic carbon content, and catchments with bigger organic carbon pools having a higher predicted median and peak concentration. 37 catchments were allocated into this group.

Topsoil organic carbon is also the first separator for iron, with catchments showing higher concentrations where topsoil organic carbon content is higher. This MTPCT has another separator, cattle density, predicting that catchments with a higher mean number of cattle have higher median and peak concentrations. Almost half of the catchments again fall into the group with the lowest median and peak concentration predictions.

For manganese, only two catchments containing the highest sheep densities were separated predicting very high median and peak concentrations.

The MTPCT for turbidity showed highest predicted peak concentrations for catchments with high improved grassland cover (>27.1%), although the highest predicted median concentrations were made for catchments with higher values for SPR. Slightly higher concentrations were also predicted for catchments with high percentage in sandstone. The MTPCT further predicts slightly higher values for catchments where livestock numbers are higher.

The MTPCT for pH predicted values to be highest in catchments with more than 25% improved grassland. The MTPCT splits into several strands, first separating into catchments with cattle and those with almost no cattle. Those catchments with cattle presence then show low pH median and peak values where topsoil organic carbon content is higher. For those catchments with smaller organic carbon pools, median and peak pH values are predicted to be slightly basic except for two catchments with very high predicted peak pH, where mean total rainfall per month is low. Of the catchments without cattle, BFI appears several times as a separator, with three catchments being predicted with very low pH values where the BFI value is between 0.14 and 0.16.

Both MTPCTs for coliform and *E. coli* separated two catchments with very high concentrations first, using urban area cover. The catchments were then not split any further for coliform, but the *E. coli* MTPCT then also identified six catchments with septic tanks as having higher concentrations. Of the remaining catchments, predicted median and peak concentrations were highest for the three driest catchments.

Table 4.5: MTPCTs per water quality parameter. ‘yes’ and ‘no’ refer to the condition above, leading to a new condition or the predicted median and 95th percentile values in [], followed by the number of catchments falling within this prediction. SD = Standard deviation, to compare against RSME = Root mean square error to evaluate model overfitting, and R^2 to evaluate goodness of model fit, for the training and test sets (10-fold cross validation of complete n=154).

<table>
<tr><th rowspan="2">MTPCT – pruned</th><th rowspan="2">SD of complete data set</th><th colspan="2">RSME</th><th colspan="2">R^2</th></tr>
<tr><th>Training</th><th>Testing</th><th>Training</th><th>Testing</th></tr>
<tr><td>Aluminium:
Cattle > 46.9
+--yes: [131.85, 738.2]: 24
+--no: Coniferous > 50.1
        +--yes: [124.5, 334.13]: 6
        +--no: MeanTemperature > 5.01
                +--yes: MeanMonthlyRainfall > 82.1
                | +--yes: [59.46, 166.24]: 113
                | +--no: [98.06, 398.03]: 8
                +--no: [109.83, 732.8]: 3</td><td>[51.81, 408.18]</td><td>[42.9, 344.28]</td><td>[53.68, 430.73]</td><td>[0.31, 0.28]</td><td>[0.06, 0.05]</td></tr>
<tr><td>Colour :
OrganicCarbonContent > 40.24
+--yes: [55.41, 122.47]: 37
+--no: [33.75, 65.84]: 117</td><td>[24.93, 55.04]</td><td>[23.06, 49.23]</td><td>[25.68, 55.99]</td><td>[0.14, 0.19]</td><td>[0.02, 0.04]</td></tr>
<tr><td>Iron:
OrganicCarbonContent > 45.33
+--yes: [640.25, 1886.08]: 12
+--no: Cattle > 46.9
        +--yes: [389.2, 1785.77]: 20
        +--no: [210.81, 713.61]: 122</td><td>[243.87, 1185.41]</td><td>[209.27, 1092.17]</td><td>[251.52, 1228.4]</td><td>[0.26, 0.15]</td><td>[0.03, 0.01]</td></tr>
<tr><td>Manganese:
Sheep > 323.1
+--yes: [218.75, 3049.75]: 2
+--no: [19.53, 183.03]: 152</td><td>[36.99, 688.91]</td><td>[29.16, 605.12]</td><td>[42.39, 765.07]</td><td>[0.37, 0.22]</td><td>[0.005, 0.002]</td></tr>
<tr><td>Turbidity:
ImprovedGrass> 27.1
+--yes: [3.19, 32.44]: 11
+--no: GeologySandstone > 95.54
        +--yes: [1.71, 8.74]: 18
        +--no: Heather > 0.0
                +--yes: Cattle > 9.1
                | +--yes: [1.05, 3.46]: 27
                | +--no: ReliefRatio > 0.062
                | +--yes: [0.5, 1.75]: 55
                | +--no: Sheep > 35.4
                | +--yes: BFI > 0.35
                | | +--yes: [1.25, 5.04]: 5
                | | +--no: [0.75, 2.41]: 14
                | +--no: [0.48, 1.49]: 18
                +--no: SPR > 49.86
                        +--yes: [4.35, 28.34]: 2
                        +--no: [0.73, 2.9]: 4</td><td>[1.09, 13.12]</td><td>[0.72, 10.14]</td><td>[1.08, 17.57]</td><td>[0.57, 0.4]</td><td>[0.18, 0.06]</td></tr>
</table>

<table>
<tr><td>pH:
ImprovedGrass > 25.44
+--yes: [7.76, 8.17]: 12
+--no: Cattle > 1.6
    +--yes: OrganicCarbonContent > 46.77
    | +--yes: [6.4, 6.8]: 4
    | +--no: MeanMonthlyRainfall > 79.59
    | +--yes: [7.2, 7.53]: 84
    | +--no: [7.35, 9.19]: 2
    +--no: BFI > 0.14
        +--yes: BFI > 0.16
        | +--yes: [6.73, 7.14]: 39
        | +--no: [5.87, 6.5]: 3
        +--no: BFI > 0.02
            +--yes: [6.5, 9.94]: 2
            +--no: [7.05, 7.35]: 8</td><td>[0.49, 0.58]</td><td>[0.35, 0.35]</td><td>[0.49, 0.7]</td><td>[0.5, 0.64]</td><td>[0.12, 0.01]</td></tr>
<tr><td>Coliform:
Urban > 1.02
+--yes: [3275, 15250]: 2
+--no: [239.81, 1985.01]: 152</td><td>[497.98, 2507.48]</td><td>[358.17, 1997.77]</td><td>[594.19, 2768.92]</td><td>[0.48, 0.36]</td><td>[0.03, 0.02]</td></tr>
<tr><td>E. coli:
Urban > 1.02
+--yes: [390, 2815]: 2
+--no: SepticTank > 0.0
    +--yes: [126.42, 1030.67]: 6
    +--no: MeanMonthlyRainfall > 77.08
        +--yes: MeanMonthlyRainfall > 234.12
        | +--yes: [8.63, 643.25]: 4
        | +--no: [7.85, 114.72]: 139
        +--no: [16.5, 1014]: 3</td><td>[60.93, 528.8]</td><td>[36.58, 371.24]</td><td>[74.93, 629.1]</td><td>[0.64, 0.5]</td><td>[0.04, 0.03]</td></tr>
</table>

4.4 Discussion

4.4.1 Water quality variability and catchment relationships

The first two PCs of the PCA together explained 66% of variability in the data and gave the first indication that there are two separate main processes leading to differences in water quality profiles. The first relates to more acidic catchments with organic soil types that produce DOC (colour), aluminium, iron, and manganese. It also gives an indication that land use and/or the integrity of the ecosystem might play a role in regulating the concentration of these parameter, as the number of sheep and cattle also associated with these indicators, as do the agricultural land uses, though to a lesser extent. The second process relates to catchments that are more intensively used and/or have a higher population density and show contamination with faecal pathogens. This pattern was reiterated in the RDA, clearly associating urban and arable areas as well as improved grassland with increased median concentration in coliform and *E. coli* and showing that these catchments would tend to have a higher contribution of baseflow, which could be due to a higher proportion of well drained soils. The RDA also clearly showed the association of colour with topsoil organic carbon and lower pH values. It is interesting that relief ratio and precipitation negatively correlate with colour and the metal concentrations, as it could be assumed that increased concentrations in these indicators would be found in the Scottish highland catchments that are typically characterised as being wetter and having steeper reliefs (Langan & Soulsby, 2002). However, this could be an indication that the highest concentrations are found in catchments with gentler relief enabling a thicker layer of organic-rich soil to develop (Evans & Warburton, 2010a; Parry et al, 2015). The negative correlation to precipitation could be due to dilution from higher amounts of rainfall.

While the third PC explained less variability (14%), there are some interesting aspects to be explored. Aluminium, iron, colour, and *E. coli* only very weakly associated with this PC, while manganese, turbidity and pH showed a positive correlation and coliform a negative correlation. This PC thus summarised a process that increases manganese concentrations but doesn't affect the other metals or colour, as well as decoupling coliform and *E. coli* concentrations. It also leads to higher pH values and increased turbidity. Due to the decoupled higher manganese concentrations, groundwater could play a role here as manganese concentrations are typically higher in groundwater (Homoncik et al., 2010). This would fit well with higher values for pH and a weak but significant positive correlation to sandstone, as groundwater in this type of bedrock typically has a slightly higher pH (Homoncik et al., 2010). There is also a negative correlation to the two precipitation variables, again indicating these catchments may have a higher contribution of groundwater to the water body. The negative correlation to coliforms, and its decoupling from *E. coli* is harder to interpret. There is a negative correlation to improved grassland, but a positive correlation to numbers of sheep and cattle. Despite this latter positive correlation, these catchments might not be intensively used for agriculture, generally showing low concentrations for *E. coli* and coliform bacteria.

The cluster analysis, while showing that it is not possible to find an acceptable typology using median concentrations of these eight water quality indicators, still yields results that can be linked to the findings from the PCA and RDA. Geographically, catchments in Scotland can very broadly be distinguished as highland catchments predominantly in the West of Scotland with steeper slopes, high precipitation, and more natural land cover; and catchments in the lowland areas of the East and South of Scotland that are drier and less steep and have more agricultural use (Gillen, 2013). Especially cluster 3 with higher median concentrations in colour and metals and more acidic pH median values could be seen as representing highland catchments, although reliefs are gentler compared to clusters 1 and 2. This could indicate more occurrence of blanket peat and thicker depth of the peat soil which would accumulate better on gentler reliefs (Evans & Warburton, 2010a). These catchments would well represent the variability explained through PC1 and the identified relationships to catchment characteristics. Cluster 5, with high median concentrations of *E. coli* and coliform, and the higher percentages of arable, improved grassland and urban areas could be representing the extreme of the agricultural lowland catchments, mirroring the water quality variability explained with PC2. Water quality and catchment characteristics in cluster 4 are a lot more heterogenous, which is unsurprising given the low cluster homogeneity indicated in the silhouette plot (Figure 4.4). Cluster 1 and 2 incorporate the catchments with predominantly lower median values for the water quality indicators, the main difference observable here is that cluster 1 includes catchments with higher median pH values and cluster 2 those with more acidic pH.

The MTPCTs, while not usable as predictive models, still reiterate the observed relationships to catchment characteristics. Colour and iron both showed that the carbon content of the soils is a good separator to identify catchments with higher concentrations. For iron, manganese and aluminium, livestock numbers also play a role in increasing concentrations. For aluminium, coniferous forest cover was also associated with higher concentrations, which is in line with findings that aluminium concentrations are elevated in forested catchments due to a slower acidification reversal (Battarbee et al., 2014). It could also point to release of contaminants through mechanical disturbance from forestry activity (Van Dijk & Keenan, 2007). It is interesting that coniferous forest cover was not picked up in any other MTPCT or in the other methods. This suggests forestry expresses varying influence depending on its combination with other risk factors, and on water quality conscious management (Reynolds, 2007; Warrington et al., 2017). Elevated levels of turbidity were attributed in the MTPCTs to several factors, including more intensive land use (improved grassland), sandstone, and a high SPR. This shows that turbidity is influenced by different processes, either natural or anthropogenic. Mechanical disturbance and removal of vegetation cover from agriculture can lead to higher erosion through surface and sub-surface mobilisation of particulates and delivery into the channel network (Thompson et al., 2013). The variety of factors working separately and in combination explains why the associations of turbidity in the PCA and RDA were less clear, with turbidity associating positively to all of the first three PCs. Cluster 4 includes many catchments with higher median concentrations of turbidity, and looking at its spatial distribution, implies association with soils developed on underlying sedimentary lithologies, although this could not be established through the catchment data. The MTPCT for pH showed higher pH values associated with improved grassland cover, which is unsurprising as this type of land use usually has a higher soil pH than unimproved areas (Grayston et al., 2001). It also showed lower pH values associated with higher topsoil organic carbon, which again can be related to more acidic upland catchments high in organic soils. Finally, the MTPCTs for coliform and *E. coli* related high concentrations to urban land cover. While urban areas are certainly associated with higher contamination of faecal pathogens (Barbosa et al., 2012), urban area cover is positively correlated with arable area and improved grassland cover, and it can be assumed that all these factors play a role in these catchments, again indicating lowland areas. It is perhaps surprising that improved grassland or livestock densities were not identified as explanatory variables for these parameters, but Neill et al. (2018) found that the extent of arable or pasture land was not a good predictor for *E. coli* concentration especially in smaller catchments, as point sources would have a greater influence than diffuse sources. Although septic tanks as point sources were included in the *E. coli* MTPCT, there are only few catchments with septic tanks. Furthermore, the occurrence of single contamination events, e.g., through access of livestock to a stream, depends on management (Newell-Price et al., 2011) which would not be reflected in the data.

4.4.2 Identification of a catchment typology and risk factors

Each of the different empirical analysis approaches provided some insight into patterns in variability of the water quality indicators and their relationship to catchment characteristics. PCA identified overall trends and controls in the water quality data, giving the first indication of a broad distinction with different water quality profiles. RDA supported the identification of explanatory catchment characteristics. The cluster analysis added a spatial perspective, allowing to identify national scale influences as well as showing that local conditions can override larger scale influences on water quality. The MTPCTs reiterated most important explanatory variables for individual water quality indicators, allowing a better interpretation of overarching processes and local influences. All these approaches showed limitation to interpretability, but the combination of these different techniques supported a more comprehensive exploration of the data and a more homogenous understanding of different patterns and relationships, which aided interpretation of causation and risk factors.

A clear typology with catchments that show distinct water quality profiles with matching catchment characteristics could not be achieved. Due to the broad-scale analysis, the loss of nuance in water quality data from summarising into median and 95th percentile concentrations, difficulties in finding data to adequately represent catchments conditions, and the general complexity of catchments and its relationship to water quality, it is perhaps not surprising that many relationships in water quality and catchment characteristics remain unclear.

Summarising the water quality data dismisses any intra-annual variability that could help to distinguish catchments and influences on water quality. Differences in response to extreme events for example would help to understand catchment vulnerabilities and hence risk factors. The low sampling frequency means that especially in catchments with a lower number of samples, extremes are likely to be missed. On the other hand, a bias might exist especially in catchments with a higher sampling frequency as sampling can be reactive to high concentrations. This might exacerbate differences in baseline concentrations.

Due to the high spatial and temporal diversity in headwater streams, relationships between catchment characteristics and water quality becomes weaker in small catchments (Abbott et al., 2017). Catchment data that were available through national surveys often had a resolution that ill matched the small catchments size, so higher resolution data would probably improve the analysis. Water body type may also mask catchment effects, for example through differences in buffering capacity for high concentration following extreme events. This may mean that similar catchment profiles are associated with different water quality outcomes. Catchment data also didn't reflect the location of specific characteristics within the catchment, or the combination of characteristics, which may lead to differences in water quality depending for example on how well connected the source is to the receiving water body.

To improve the analysis and find clearer relationships between catchment characteristics and water quality outcomes, the quality of the catchment characteristics data could be improved with higher resolution data, and water quality data could be looked at in more detail e.g., in terms of intra-annual variability. Responses to extreme events could also be analysed for a better understanding of catchment vulnerabilities. This would enable a better estimation of how catchments may respond to changes in pressures, and also enable a typology.

Despite these limitations, a first high-level risk screening was achieved that identifies predominant pressures in broadly different types of catchments, with associated water quality issues. The drinking water catchments analysed cover the variety of environmental conditions found across Scotland, and water quality mirror this variability. However, a broad distinction can be made between "upland" catchments, dominated by more natural land cover and highly organic and acidic soils that yield higher concentrations especially in colour, but also in aluminium, iron, manganese, and turbidity; and "lowland" catchments with higher pH values, more intensive agricultural use, and higher concentrations of coliform, *E. coli*, and turbidity concentrations. Therefore, more acidic, peaty upland catchments prone to water-logged conditions can generally be identified as candidate catchments for higher colour risk, with indications that anthropogenic influence leading to peat degradation and erosion can increase this risk. Catchments with more intensive agricultural use and with a higher population density are candidate catchments for higher faecal pathogen risk. These types of catchments, and the identified water quality issues, can now be looked at in more detail to identify drivers and pressures, understand differences in catchment responses and their relationships to catchment conditions, and thus enable to identify different catchment vulnerabilities that allow to estimate risk from future changes in climate and land use.

4.5 Further approach

Despite great variability in water quality between catchments and difficulties to clearly separate them into different “types” of catchments, a distinction into two very broad categories was made, and for each a water quality indicator can be seen as diagnostic: “upland” catchments with elevated concentrations of colour, and “lowland” catchments with elevated concentrations in *E. coli*. Focusing on these two water quality indicators separately, to arrive at better insights into processes leading to contamination, and of controls over these within the catchments, could provide a deeper understanding of vulnerable catchments and associated risks from changes in climate and land use, and be seen as representing a broader suite of water quality issues.

Colour concentrations are already elevated in many catchments, requiring treatment to meet the standard (Table 3.1). Upward trends in DOC have been observed throughout the Northern hemisphere over the past decades and it is expected that DOC release from Scottish peatlands will rise further in the future (Evans et al., 2005; Ritson et al., 2014, Sawicka et al., 2017). Climate change may exacerbate the problem through more frequent heavy rainfall events (ASC, 2016; Burt & Howden, 2013) and increases in erosion and runoff rates (Li et al., 2016). While the analysis generally linked larger carbon pools to higher colour concentrations, specific risk factors are more difficult to disentangle. Exploring intra-annual variability in concentrations and responses to climatic factors is needed to further understand impacts of climate and anthropogenic factors in increasing colour concentrations in surface waters, and to specifically identify catchments most at risk of seeing increases in colour concentrations that may exceed current treatment capacity.

E. coli concentrations have been linked to anthropogenic land uses which are predominant in the East of Scotland. However, some areas of Scotland, especially the zones between uplands and lowlands, have been highlighted as becoming more suitable for agricultural activity as current constrains will be reduced with rising temperatures (Brown et al., 2010). This may lead to agricultural intensification within these areas, increasing the risk for bacterial contamination, and a suite of other water quality indicators that are usually associated with agriculture, such as nitrate or pesticides (Kay et al., 2009). Identifying catchments with these potentials for land use change, and exploring what this would mean for water quality, is the next step in identifying high risk catchments for water quality degradation from land use change.

Identifying high-risk catchments with specific water quality issues from identified pressures will allow a prioritisation for further investigation as well as mitigation and adaptation actions, to aid risk assessment and planning at programme level. It can also be a starting point to indicate best methods for building resilience to change, depending on catchment conditions, and thus constitutes a step towards pro-active management rather than re-active intervention.

Following the identification of current water quality profiles and issues and of two diagnostic water quality indicators, this research now investigates in more detail catchment relationships to colour (5) and *E. coli* (6) concentrations in order to screen for catchments most at risk from climate and land use changes.

5. Risk screening: Investigating links between catchment characteristics and colour concentrations in raw water

Assessment 1.2 aims to find out how climate change is likely to act on current water quality issues and what this indicates for other water quality concerns. The following analysis explores these questions with regard to colour, which is a common concern in many Scottish Water catchments (see 4). First, mechanisms of colour production are explained to set the background and rationale for the analysis (5.1), then methods are described for each step of the analysis (5.2), the results are then presented and discussed (5.3), and finally limitations are reviewed and reflections given on ways to improve the analysis, and further steps (5.4).

Water colour in Scotland is predominately due to the presence of dissolved organic carbon (DOC), which is supported by the results reported in 4.3. The concentration of organic carbon in many Scottish rivers has doubled over the last decade of the 20th and the first decade of the 21st centuries (Moxley, 2014), and continues to increase (de Wit et al., 2021). Increasing trends in dissolved organic carbon (DOC) have also been observed across the Northern Hemisphere (Monteith et al., 2007; Sawicka et al., 2017). The causes for this observed trend are debated with increasing temperatures (Cole et al., 2002; Freeman et al., 2001), recovery from acidification (Evans, Chapman et al., 2006; Monteith et al., 2007), changes in hydrology (Tranvik & Jansson, 2002), and land management and peatland drainage (Worrall, Armstrong & Adamson, 2007; Worrall, Armstrong & Holden, 2007; Yallop & Clutterbuck, 2009) among the suggested causes.

DOC in drinking water is a matter of aesthetics (discolouration), it influences the treatment process, and its presence leads to the formation of disinfection by-products that cause health concerns (Sillanpää et al., 2018; Villanueva et al., 2015). The observed trends therefore raise concerns over rising costs for the removal of DOC from drinking water as well as the capability of existing water treatment assets to effectively cope with increasing concentrations. Limited information on changing DOC concentrations and its linkages to climate and land use changes adds complexity for water utilities when planning investment and adaptation strategies (Ritson et al., 2014).

Many of the source water catchments have median concentrations well above the prescribed value at the consumers' tap, making colour concentrations an important part of the treatment and rising concentrations a concern. Removing organic carbon from raw water is costly and being able to reduce the amount of organic carbon as well as to plan for possibly rising concentrations is therefore an important part of strategic planning for water safety in terms of water quality. Indications of where to expect (further) increases in concentrations would provide a basis to understand where and what kind of actions may be needed in short or long term and focus effort on catchments and regions most at risk.

5.1 Mechanisms of colour production

In upland areas of Scotland, DOC is derived mainly from terrestrial organic matter, and there is a correlation between soil carbon pools and stream DOC concentrations (Dawson et al., 2008). The rate of biological decomposition of organic matter leading to the production of DOC in soil is driven especially by soil temperature and moisture: higher temperatures increase microbial activity, leading to a higher production of TOC (Freeman et al., 2001). Decomposition is also influenced by moisture, where waterlogged conditions inhibit DOC production (Clark et al. 2009). These patterns lead to higher concentrations of soil DOC in summer.

DOC solubility in water is necessary to enable transport from the soil to surface water and is influenced by the acidity of the soil and the ionic strength of the soil solution, with decreasing acidity and increasing ionic strength reducing solubility (Evans et al., 2012). DOC transport from the soil is linked to precipitation events and changing flow paths through the soil horizons (Dawson et al., 2002; Wen et al., 2020). Mineral soils can retain and store carbon, the extent of which depends on organic matter content, mineralogy, clay content and pH (Chapman & Palmer, 2016). In organo-mineral soils, storm events may induce a shift from flow through the mineral horizon to flow through and over the organic horizon, inducing an increase in DOC concentration in surface water (Stutter, Dunn & Lumsdon, 2012). In contrast, catchments with a thick organic horizon may see a dilution or no change in concentration from a storm event (Clark et al., 2007).

Once it has reached surface water, biological uptake, photo-degradation, and sedimentation may reduce the amount of DOC. Studies show that the uptake of organic matter is strongly dependent on nutrient availability and time (Chapman & Palmer, 2016). In lakes, absorption and re-suspension could influence seasonal variation (Hamilton-Taylor et al., 1996). Similarly, photo-mediation may be an important process in lakes for peat-derived organic matter, with a study by Moody et al. (2013) showing an average loss of DOC of 73%. Köhler et al. (2013) found that a large lake in Sweden showed little brownification despite other lakes in the region showing increasing colour trends, which they attribute to low residence times of iron and DOC.

Seasonal patterns of DOC in surface waters are hence explained through increased decomposition of organic matter, increased solubility, and increased release of DOC with high discharges, leading to higher concentrations of surface water DOC in autumn months (Cooper et al., 2006; Tipping et al., 2007). Thus, temperature and discharge, together with other factors such as soil moisture, acidity, or antecedent flow, influence DOC variability over the year (Clark et al., 2005; Futter & de Wit, 2008; Köhler et al., 2009). Differences in seasonal patterns of DOC between different surface water are influenced by soil properties, weather, hydrological flow paths, and processes within the water body, with catchments being dominated by different drivers. Winterdahl et al. (2014) found that catchments in Sweden differed in their sensitivity to either discharge or temperature. They studied 136 rivers or streams, using monthly total organic carbon (TOC) data and correlating it to air temperature and discharge. Analysing these together with other variability indices (normalized average annual range, coefficient of variance, seasonality index, and R^2 of an ANCOVA model explaining variance in DOC by discharge, temperature, month and the interaction of discharge and month), they proposed a first order classification of catchments into “flow-driven” (associating this with different flow pathways at different flow regimes), “temperature-driven” (associating this with peat dominated catchments), “snowmelt-dominated” (where DOC dynamics are dominated by spring floods), and “nonseasonal” (associating this with large water bodies and long DOC residence time within the water). Winterdahl et al. (2016) subsequently examined 12 Swedish headwater catchments using the Riparian flow-concentration Integration Model, to understand different sensitivities to temperature and discharge, and to project median DOC concentrations under downscaled climate projections. They found that all catchments saw higher median DOC concentrations in warmer and wetter climates, and that catchments sensitive to temperature experienced the largest increases. They also hypothesise that catchments may change in their sensitivity due to climate induced changes in soil properties. An understanding of main controls on intra-annual changes could hence support an understanding of how changing temperatures and precipitation patterns will influence different types of catchments, in turn reflecting on long-term DOC trends.

Differences in overall DOC concentrations between catchments are influenced by differences in climate, soil properties, topography, and land cover and vegetation. Musolff et al. (2018) used partial least squares regression to estimate median DOC concentrations and concentration variability (expressed as ratio of interquartile range and median) from catchment characteristics and found that median DOC is mainly explained by soil wetness and slopes, with catchments with high wetness and low slopes having highest DOC concentrations. Cool et al. (2014) used iterative generalised least squares to estimate the influence of catchment characteristics on DOC concentrations in Quebec, finding that slope negatively influenced concentrations whereas total precipitation from the antecedent 10 days, mean temperature from the antecedent 60 days, and coniferous and mixed forest cover positively influenced DOC concentrations. Monteith et al. (2015) used multiple regression with backward variable elimination based on Akaike Information Criterion to find which catchment characteristics explain mean DOC concentrations in UK upland catchments. They found that organic soil types (peat and peaty gleys) were the strongest predictor of DOC, and that effective precipitation negatively influenced DOC concentrations, attributing this to a diluting effect. DOC also exponentially increased with declining altitude, which could be a combination of several factors, for example increased primary production, or increased temperature. They conclude that increasing temperatures could have implications for long-term DOC release in these catchments, but that soil type is likely to influence sensitivity to future warming.

5.2 Methods

The risk screening aims to 1) investigate differences in colour concentration patterns in a range of raw water sources throughout Scotland, in order to 2) understand how catchment characteristics influence overall colour concentrations and intra-annual colour variability, 3) understand how climate parameters can exert different influences over colour concentration patterns depending on these characteristics, so that 4) catchments can be classified according to their sensitivity to changes in climate, and perhaps in land use. This will ultimately allow to 5) identify high risk catchments for increasing colour concentrations.

The analysis is set up to test and explore the following assumptions:

- Catchments will show a concentration pattern for colour where concentrations in surface waters start to rise in summer and peak in autumn, with concentrations going down in winter, due to mechanisms of DOC production and transfer in soils.
- Catchments will show different relationships between colour concentrations in surface waters and flow as well as temperature, depending on soils and flowpaths through the soil horizons. Peat dominated catchments are expected to show a relationship with temperature, whereas catchments with organo-mineral soils are expected to show a reaction to flow events.
- Depending on temperatures, and on properties of the soil in terms of drainage and potentially artificial drainage, catchments will show differences in wetting-up periods and hence different time lags between rainfall events and DOC concentrations in surface waters.
- Land use and management have different impacts on DOC and colour concentrations for catchments with different sensitivities.
- DOC will be lost especially in reservoirs and lakes due to photo-mediation and biological uptake and seasonality may be influenced by absorption and re-suspension. This will make it more difficult to discern catchment-related influences on DOC and colour concentrations.

5.2.1 Colour as a proxy for DOC

It is assumed that in Scotland, colour is a good proxy for DOC, although colour can be also induced by iron or manganese (Kritzberg & Eckström, 2012), or be influenced by type of compounds (referred to as the quality of DOC), with some compounds being more coloured than others. DOC data are available for a very limited number of Scottish Water catchments, whereas TOC concentrations are monitored more frequently. Data was obtained from Scottish Water for 127 catchments for a period of 4 years (2013-2016), with sample frequency varying per catchment from every four weeks to weekly. While TOC includes suspended sediment, it should be an improvement over colour to represent DOC (Winterdahl et al., 2014). Linear regression was performed on TOC and colour concentrations per catchment to test usability of colour as a proxy for DOC and potential limitations (see B.4 for R code). As a comparison, multiple linear regression models were developed that included iron, manganese, and turbidity as further independent variables, as these parameters are also leading to water colouration. To judge the strength of the relationships, the R^2 value of the model was used. A value of 0.4 was chosen as a cut-off, below which the relationship was supposed to be weak enough to assume other factors playing a significant role to determine colour.

5.2.2 Shape-based clustering

To understand patterns of seasonal concentrations and to explore differences, colour time series for 154 catchments from 2011-2016 were used for shape-based clustering (B.5). Shape-based clustering is a technique for unsupervised grouping of shapes, in this case the shape of the curve of colour concentrations over the year. The aim was to find similarities in dynamics of colour concentrations, such as timings and speed of increases and decreases.

As time series were unevenly spaced, daily values were interpolated, using linear interpolation. Linear interpolation is easy to achieve and has been shown to work well in creating daily time series from water quality sample data with less than 28-day gaps, especially for some water quality parameters, including DOC (Gnauck, 2004). It was therefore chosen as an appropriate interpolation method to aid in the creation of yearly profiles for colour. The daily data points were then logged to avoid an overwhelming influence of extreme outliers. Yearly profiles were extracted by calculating means for every day of the year. The yearly profiles were separated by type of source (reservoir, river, loch) for subsequent analysis and clustered using the k-Shape algorithm. K-Shape is a partitional clustering method based on shape-based distance (SBD), using a centroid shape (time series). The series are z-normalised. The algorithm chooses a random shape as a starting point. K-Shape clustering has been shown to work efficiently and accurately on time series (Paparrizos & Gravano, 2016). Clustering was repeated for a changing number k of clusters and the results plotted to allow visual inspection of the cluster.

5.2.3 Climate sensitivity analysis

To understand differences in catchment sensitivity to soil temperature, flow, and wetting up processes, the relationship between TOC concentrations and temperature as well as flow was investigated. As proxies, a) total rainfall of the three preceding days to the sampling date was used for flow conditions after a short-term rainfall event, b) total rainfall of the 60 days prior to the sampling date was used for catchment wetness, and c) mean temperature of the 60-day period prior to the sample date was used for soil temperature. Rainfall and temperature data were obtained from the UK Met Office. The datasets were split by season (December-February for winter, March-May for spring, June-August for summer, and September-November for autumn). Spearman's rank correlation tests were run individually for each catchment for each of these variables for each season (B.6), resulting in a set of 12 Spearman's rho and corresponding p-values per catchment. A relationship was judged as moderate if Spearman's rho > 0.4 and as strong if Spearman's rho > 0.6 if p=<0.05.

The catchments were then sorted into 5 possible categories: "Temperature", "Rainfall". "Rainfall+Temperature", "Wetup" and "None", depending on the variables for which strong or moderate relationships could be observed, with priority given for seasons of peak TOC concentrations. To visualise the relationships, and further examine for patterns between the relationship strengths, and between the relationships and the identified grouping, PCA on Spearman's rho values was performed.

Differences in catchment characteristics between the groups were visually inspected using violin plots and statistically analysed with Kruskal-Wallis and Wilcoxon tests. Redundancy analysis on Spearman's rho was used to test explanatory power of catchment characteristics for the different correlation strengths. Independent variables included in the analysis were catchment area in km^2, elevation relief ratio, relief ratio, percentage of Southwest facing aspect, average topsoil organic carbon content, average baseflow index, percentage of semi-natural land cover, percentage of arable land cover, percentage of improved grassland cover, percentage of coniferous forest cover, number of deer, sheep and cattle density, amount of SER, accumulated annual temperature above 5.5°C, and median pH values. Automatic stepwise backward modelling based on Akaike Information Criterion (AIC) was used to find significant variables. AIC is a widely used estimator of model error that tends to favour inclusion of variables with subtle effects and is applied especially in a predictive context (Cavanaugh & Neath, 2018). For the final model, some variables were added back in to aid interpretation.

5.2.4 Catchment characteristics relationship to median TOC concentrations

Multiple linear regression was used with catchment median TOC concentrations as the dependent variable, and catchment characteristics as independent variables, to determine how catchment characteristics influence median TOC concentrations. This was performed for all catchments together as well as separately for each group identified in the climate sensitivity analysis (B.7). Catchment characteristics initially included as independent variables were the same as in the redundancy analysis (catchment area in km^2, elevation relief ratio, relief ratio, percentage of Southwest facing aspect, average topsoil organic carbon content, average baseflow index, percentage of semi-natural land cover, percentage of arable land cover, percentage of improved grassland cover, percentage of coniferous forest cover, average number of deer, sheep and cattle, amount of SER, accumulated annual temperature above 5.5°C, and median pH). Bayesian Information Criterion (BIC) was used in a stepwise backward regression to find the significant variables. The dependent variable was log transformed to bring residuals nearer to a normal distribution. Adjusted R^2 was used to evaluate goodness of fit. Ten times repeated 10-fold cross validation and the resulting normalised Root Mean Squared Error (RMSE/(Max-Min)) was used to check for overfitting. If models didn't perform well, some independent variables with suspected collinearity were removed from the initial model to see if final model performance improved.

Multiple linear regression with the same catchment characteristics as independent variables was also used to determine how catchment characteristics influence median TOC concentrations for all the catchments together, but also using interaction terms to further identify potential impacts of climate and land use on different underlying sensitivities. Stepwise backward regression based on BIC was used (B.8) to find significant independent variables. Interaction terms were introduced batchwise, with one batch each for interactions between SER and all other variables, between AAT and all other variables, between category and all other variables, and some selected interactions between the remaining variables. Adjusted R^2 and 10 times repeated 10-fold cross-validation was used to evaluate goodness of fit.

5.3 Results and discussion

5.3.1 Colour-TOC relationships

Of the 127 catchments for which a colour-TOC model was produced, 34 resulted in a model with an R^2 value below the cut-off (Table C.1). If re-running the models for all catchments, and including iron and manganese, only 19 catchments were below the R^2 cut-off (Table C.2). If also including turbidity, this number was 18 out of 127 catchments (Table C.3). For some catchments, including iron, manganese and turbidity resulted in reduced model fit. Hence for the majority of the catchments, TOC alone produces a good estimate of colour value, confirming that DOC is a main contributor for colour, while for some catchments, iron, manganese, and turbidity also substantially contribute to colour (Figure 5.1). Still, some catchments clearly do not have a linear relationship between TOC and colour, or TOC, iron, manganese, turbidity, and colour. This could be due to a change in the quality of DOC, meaning that the types of compounds produced within the catchment change over the year (Vidon et al., 2014). This results in different properties with regard to colour and would hence not allow to establish a linear relationship.

Secondly, intercepts of the models above the R^2 cut-off value varied, with most of the models with high R^2 values having negative intercepts. This is an indicator that colour measurements are missing organic carbon that is not visible, i.e., not detectable through the standard measurement. Organic material produced by algae for example tends to be less coloured (Henderson et al., 2008), so negative intercepts could also be an indicator in which catchments relationships between catchment characteristics and TOC concentrations could be confounded by in-lake processes.

In terms of operational practice, where colour is measured on a continuous basis to inform treatment, this may not be suitable for catchments where a high amount of DOC is produced that is not detectable through colour measurements.

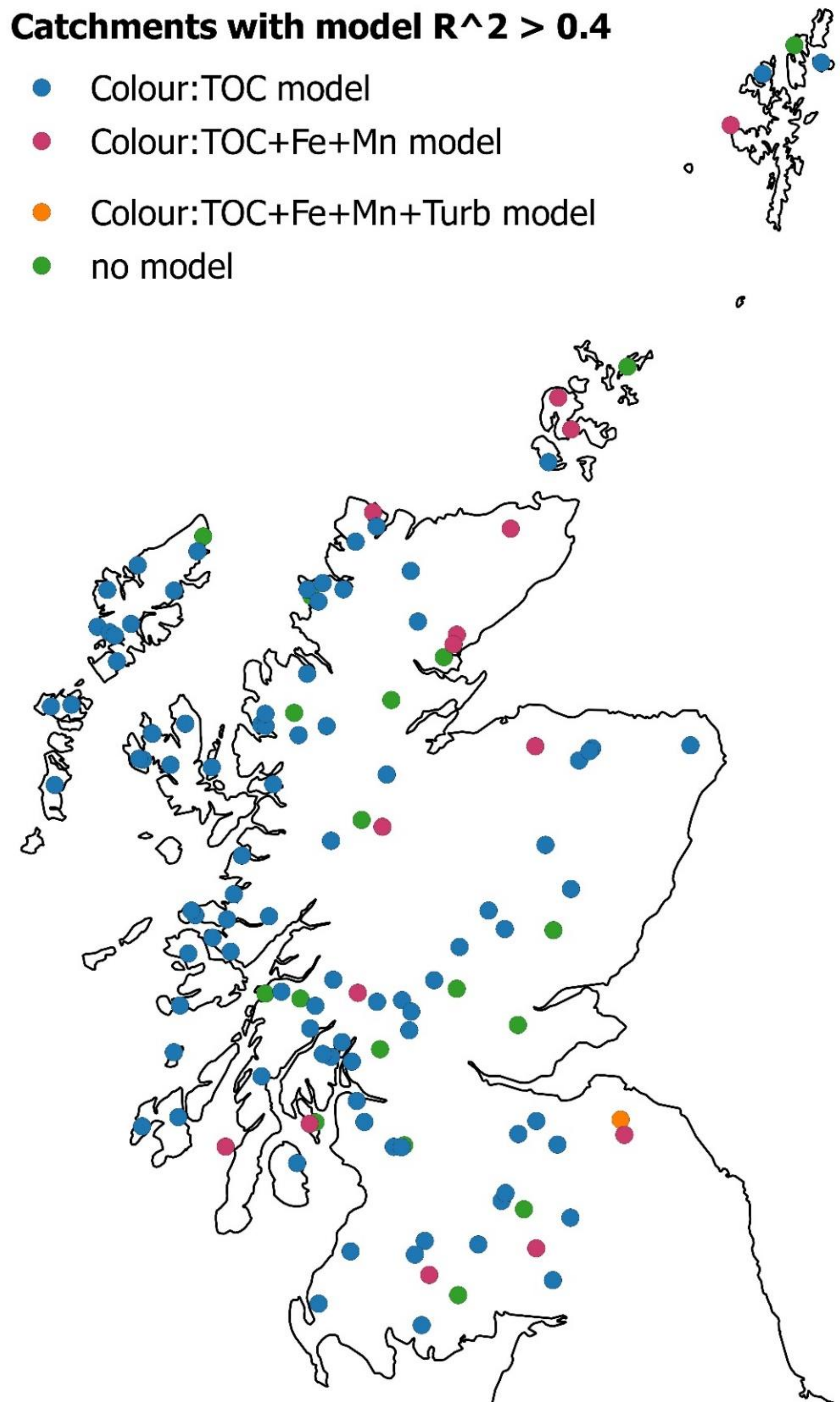


Figure 5.1: Location of catchments with satisfactory ($R^2 > 0.4$) models for colour, by inclusion of independent variables (TOC, iron, manganese, turbidity). Blue are catchments where only including TOC as an independent variable to predict colour as dependent variable in a linear regression produced a model with an R^2 value above 0.4, violet are catchments where additional inclusion of iron and manganese brought the model to an R^2 value of above 0.4, while for the orange catchment, turbidity had to be additionally included. The green catchments did not yield a satisfactory model even when all four independent variables were included.

5.3.2 Shape-based clustering

Visual inspection of the patterns emerging from the shape-based clustering using different pre-defined numbers of clusters indicated for reservoirs and lochs, that some catchments roughly showed the expected curve of rising concentrations in summer and peaks in autumn (looking like a “sine curve” on the yearly time line), and that some catchments showed a time series looking more like a “V-shape” over the year, with a dip in concentration in the summer and a peak in late autumn/winter. River catchments showed more irregular shapes, but most catchments roughly followed the hypothesised seasonal curve.

It is possible that linear interpolation draws out dips or spikes in the timeline, making them appear over an extended time, as extreme values will increase or decrease all values from the sample point before and after the extreme. Especially for time series with low sampling frequency (monthly or longer) and hence large gaps between samples, this can mean “hills” or “valleys” appear when there should be short spikes or dips. The use of linear interpolation and extreme values is observable in some timeseries, especially with sudden dips (see Figure 5.2 & Figure 5.3), which would appear where sampling frequencies are shorter (weekly or two-weekly). For most catchments in this analysis, samples were taken at this interval, so these effects span a shorter time period than the observed predominant shapes, which are observable over several months. It is therefore assumed that the effects are predominantly true and not an artefact of interpolation.

After this initial inspection, the hypothesis was formed that lochs and reservoirs could be separated into groups depending on the shape of their “seasonal profile” as either a “sine curve” or a “V-shape”. To do this, the clustering was repeated separately for all lochs together and all reservoirs together, for 5 times with a number of clusters k=2. For each clustering, one cluster was defined as the “sine” and the other as the “V-shape” group. The reservoirs and loch catchments were then sorted into two groups with either the expected “sine curve” like shape, or the diverging “V-shape” curve, depending on which group they fell in more often. Three lake catchments could not be assigned (Table C.4).

The yearly timelines of the catchments within these groups now showed roughly the same shapes of either a “sine” curve or a “V” (Figure 5.2). Rivers all roughly showed a “sine” shape (Figure 5.3).

The catchments were finally separated into 4 groups: rivers (56 catchments), V-shape (29), sine-shape (66), and inconclusive (3), with the latter three categories comprising lakes and reservoirs. Spatially, a pattern was not distinguishable although two groups of “V-shape” catchments can be seen that are spatially proximate, one in the North of Scotland around Kincardine, and another one South of Edinburgh (Figure 5.4)

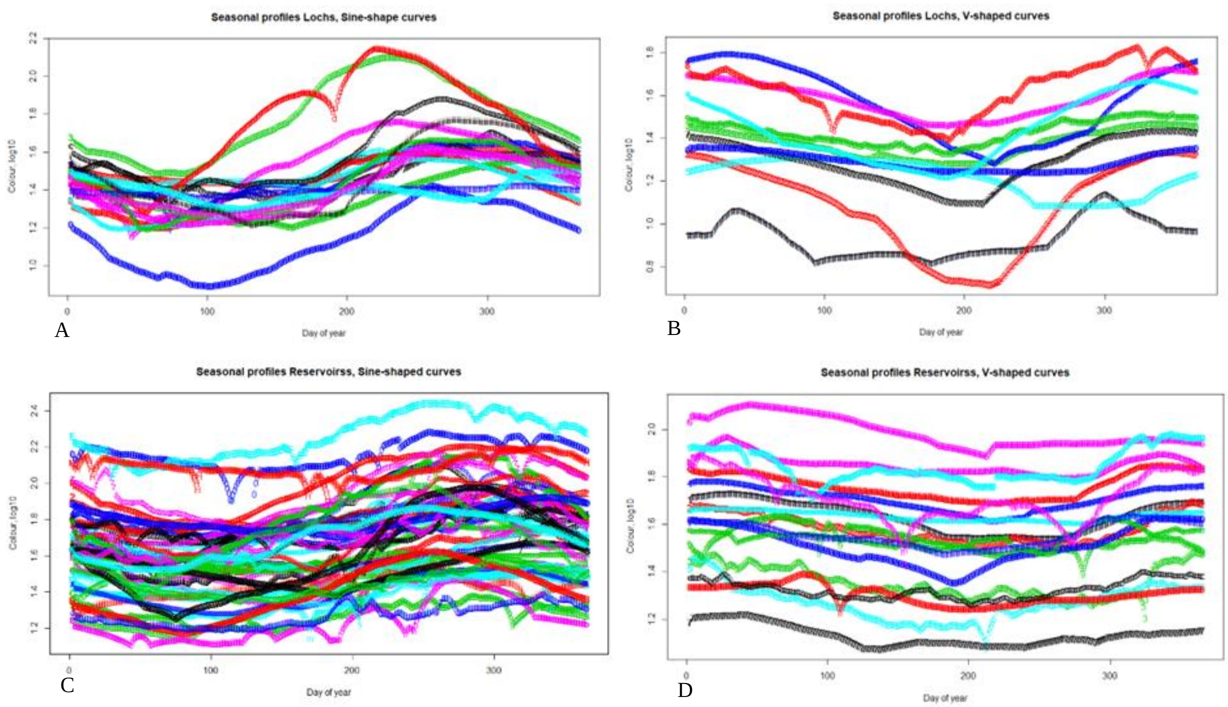


Figure 5.2: Seasonal profile for each of the groups identified in shape-based clustering: A. "Sine" shape lochs (19); B. "V-shape" lochs (11); C. "Sine" shape reservoirs (47); D. "V-shape" reservoirs (18). Seasonal profiles extracted from colour time series 2011-2016 with daily values interpolated, logged, and means taken for every day of the year. “Sine” shape profiles show the expected curve throughout the year with a dip in spring and peak in autumn, “V-shape” profiles show a delayed curve with a dip in summer and peak in winter.

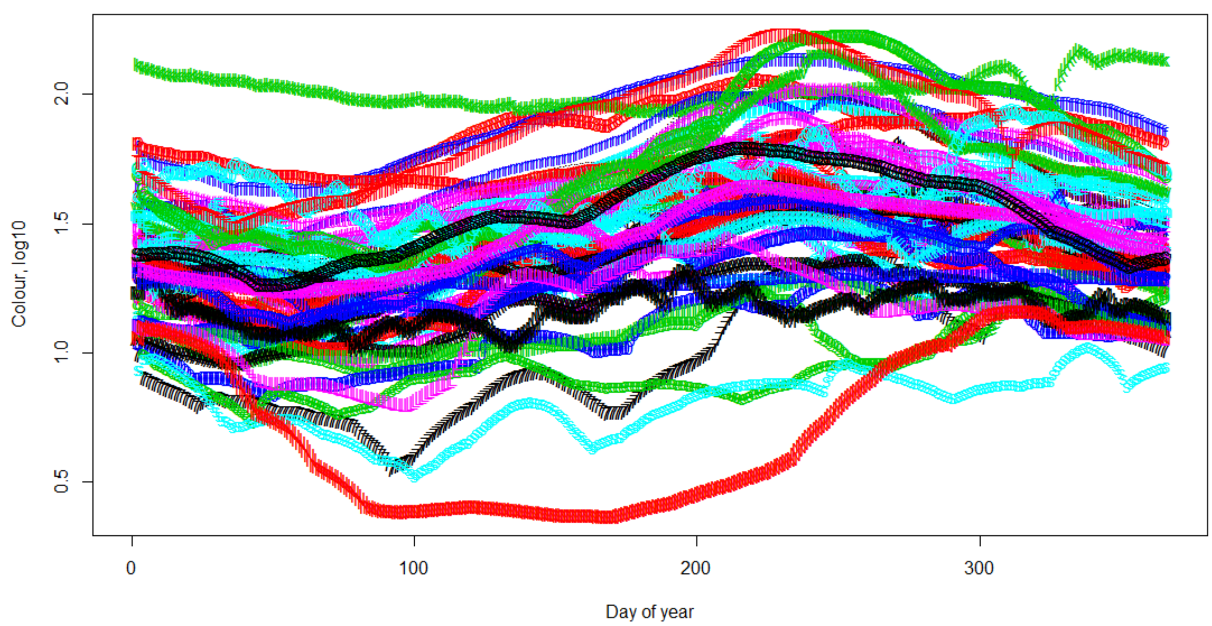


Figure 5.3: Seasonal profile for rivers (n=56). Seasonal profiles extracted from colour time series 2011-2016 with daily values interpolated, logged, and means taken for every day of the year.

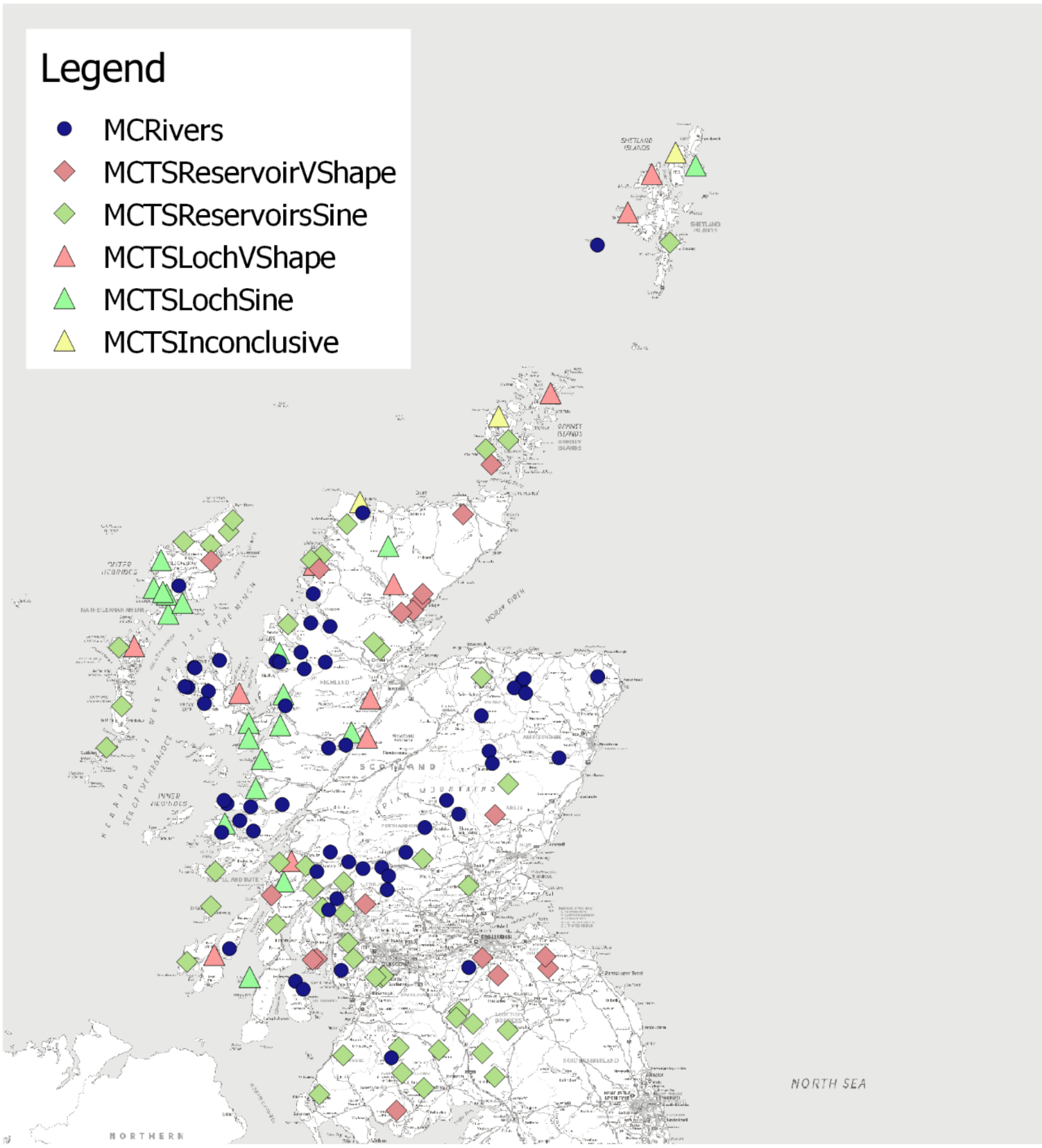


Figure 5.4: Spatial distribution of catchments, by category allocated through shape-based clustering. “MCRivers” are river catchments that were not clustered further, “MCTSReservoirVSshape” are reservoir catchments allocated into the cluster characterised by a seasonal profile that dips in summer and peaks in winter (“V-shape”), “MCTSReservoirsSine” are reservoir catchments allocated into the cluster characterised by a seasonal profile that dips in spring and peaks in autumn (“sine”), “MCTSLochVShape “ are loch catchments with a “V-shape” seasonal profile, “MCTSLochSine” are loch catchments with a “sine” seasonal profile, and “MCTSInconclusive” are catchments that could not be clearly allocated.

The catchments showing a “V-shape” curve in their seasonal profile have a later dip in concentration in the summer and a later rise and peaks of concentrations as expected, in late autumn and winter. This can be observed if looking at the mean of concentrations between the groups over the seasons (Table 5.1 & Table 5.2).

Table 5.1: Catchment median concentrations for colour in mg/l Pt/Co, for the whole year and split by season, shown for each of the categories identified through shape-based clustering (“River” = river catchment (n=56), “Sine” = catchment with seasonal profile that dips in spring and peaks in autumn (n=66), “V-shape” = catchment with seasonal profile that dips in summer and peaks in winter (n=29)).

	All	Spring	Summer	Autumn	Winter
All	38.95	33.07	46.37	53.18	38.77
River	32.02	26.63	49.51	46.38	27.29
Sine	46.67	37.16	51.38	66.58	45.64
V-shape	34.17	35.21	29.9	35.84	43.63

Table 5.2: Significance of the difference in colour concentration means between categories identified through shape-based clustering, by season, according to Wilcoxon tests (* p<0.05, ** p<0.005, *** p<0.001; “River” = river catchment (n=56), “Sine” = catchment with seasonal profile that dips in spring and peaks in autumn (n=66), “V-shape” = catchment with seasonal profile that dips in summer and peaks in winter(n=29)).

	Sine	V-shape
River	Median* Spring* Summer Autumn** Winter***	Median Spring Summer* Autumn Winter***
Sine		Median Spring Summer** Autumn*** Winter

Such a pattern in colour concentration could be due to either a later start in production of DOC, to a lag between the production and the arrival at the outlet, to processes that remove DOC before it reaches the water treatment works, or to other factors inducing colour over the winter season.

In case of a later start of production of DOC within the catchment, the curve could be expected to be flatter rather than later, or it should show a spatial pattern where these catchments are situated within generally colder areas. There are also no significant differences (tested with a Kruskal-Wallis test) in mean temperature in spring or summer between these groups which could explain this phenomenon (Figure 5.5A&B). If catchments have a shallower organic horizon and better drainage, they will have to wet up before a flow through the organic horizon, where DOC is produced and stored, takes place, which could lead to a later rise in concentrations (Stutter, Dunn & Lumsdon, 2012). This might also explain the overall lower concentrations in these catchments as DOC would have more chance to reach the mineral horizon, where it gets retained (Ussiri & Johnson, 2004). Lower concentrations could also be due to generally slightly lower organic carbon pools in these catchments (Figure 5.5C), although the difference is statistically only significant (p=0.008) between the river and the "V-shape" catchments. These catchments should also show better drained soils, or alternatively land uses which can be assumed to favour artificial drainage, but this is not the case. However, another aspect is mean monthly precipitation: There are significant differences between the groups in the amount of mean monthly rainfall, with the "V-shape" catchments receiving less rainfall than the river and "sine" shape groups (Figure 5.5D), so it could take longer for the catchment to wet up and flush out DOC produced over the summer months.

Occasional high concentrations that are flushed out over the summer months might be buffered in the water body, so the concentration at the outflow only rises later when much larger and steadier amounts are washed out of the catchment. If this was the only reason, the effect should be most pronounced in large water bodies. There is no indication that this effect takes place if catchment size taken as a proxy for water body size (Figure 5.6).

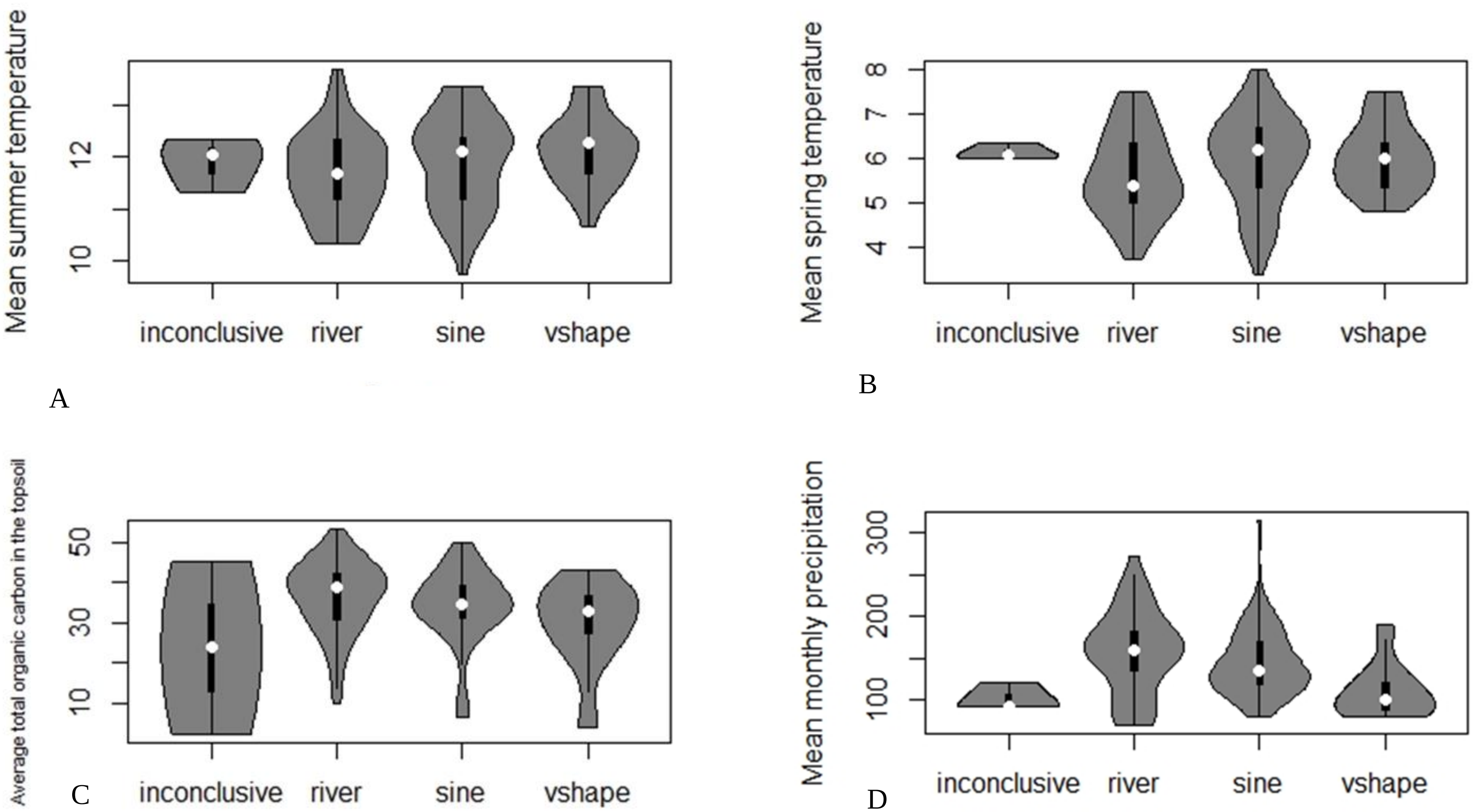


Figure 5.5: Violin plots showing distribution of A. mean spring temperatures and B. mean summer temperatures (long-term average 1981-2010), C. average total organic carbon in the topsoil, and D. mean monthly precipitation in mm (long-term average 1981-2010), within each category identified through shape-based clustering (“inconclusive” = catchment that could not be allocated (n=3), “river” = river catchment (n=56), “sine” = catchment with seasonal profile that dips in spring and peaks in autumn (n=66), “vshape” = catchment with seasonal profile that dips in summer and peaks in winter (n=29)).

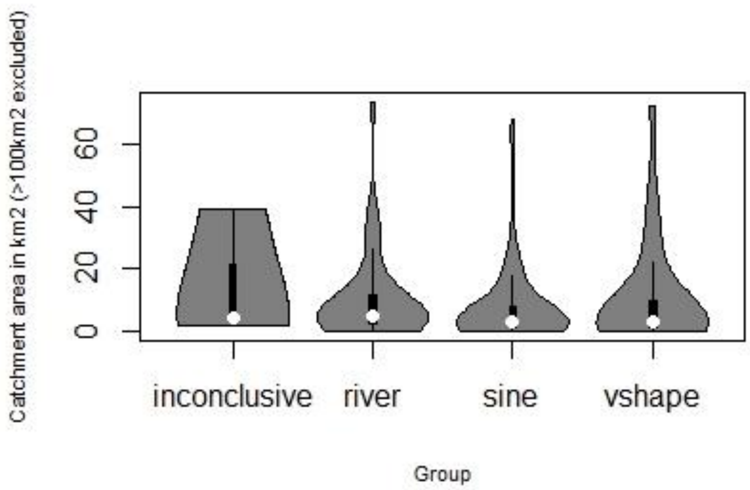


Figure 5.6: Violin plots showing distribution of catchment area in km2 within each category identified through shape-based clustering (“inconclusive” = catchment that could not be allocated (n=3), “river” = river catchment (n=49), “sine” = catchment with seasonal profile that dips in spring and peaks in autumn (n=62), “vshape” = catchment with seasonal profile that dips in summer and peaks in winter (n=28)). Catchments with areas >100 km^2 have not been represented to allow better clarity of the plot.

In terms of processes that reduce the amount of DOC, organic uptake as well as photodegradation would do this and explain a dip in summer, but less so a rise in late autumn/winter, although as the concentrations between the “sine” shaped lochs and reservoirs and the “V-shaped” lochs and reservoirs are similar in winter and spring, these concentrations might represent a “background” concentration. To understand the possible influence of this effect, parameter such as the number of hours of direct sunlight, the water temperature, or a measurement of nutrient availability and algae productivity could inform further analysis.

Lastly, iron, manganese and sediment would also induce colour. Iron and manganese that occur in the catchment usually show correlation to DOC as they are mobilised by similar processes (see 3.2.2), therefore it seems more likely that they would amplify the curve rather than delay it. If colour was predominantly influenced by suspended sediment, erosion during autumn and winter could lead to the observed concentration pattern.

The reason for these differences in concentration “profiles” could well be a combination of all these factors. To further test the possibility of differences in wetting up, comparing a year with a wet summer (2012) to a year that was very dry (2018) could help to identify if this determines the pattern. If a later rise in colour was due to DOC being flushed out later because the catchment would take longer to wet up, it could be expected to see more of a “sine” curve in these catchments in a wet year, compared to a dry year. Rerunning the analysis using only data from 2012, and only from 2018, showed that most catchments did not change behaviour, while some showed behaviour as expected under this hypothesis (switching from normally “sine-shaped” to “V-shaped” in 2018, or from normally “V-shaped” to “sine-shaped” in 2012), but a few also showed the opposite effect (switching from a normally “V-shaped” to a “sine”-shape in 2018, or normally a “sine”-shape to a “V-shape” in 2012). There seemed to be a geographical pattern to this where the catchments showing the opposite behaviour tended to be in the Northwest (Figure 5.7). This region had anomalously dry weather in 2012, but normally tends to be wetter than the rest of Scotland.

In terms of median concentrations, there was no significant difference between these two years. It was also no clear trend observable that switching from a “sine” shape to a “V-shape” or vice versa would cause median concentration to rise or fall. In terms of variation in concentrations, rivers are most variable if comparing the 95th percentile to the 5th percentile of catchment concentrations (group means are 70.89 for rivers, 60.84 for “sine” shaped catchments and 38.71 for “V-shaped” catchments), and “V-shaped” lochs and reservoirs least variable (with the variability in mean significantly different to the river and “sine” shaped catchments).

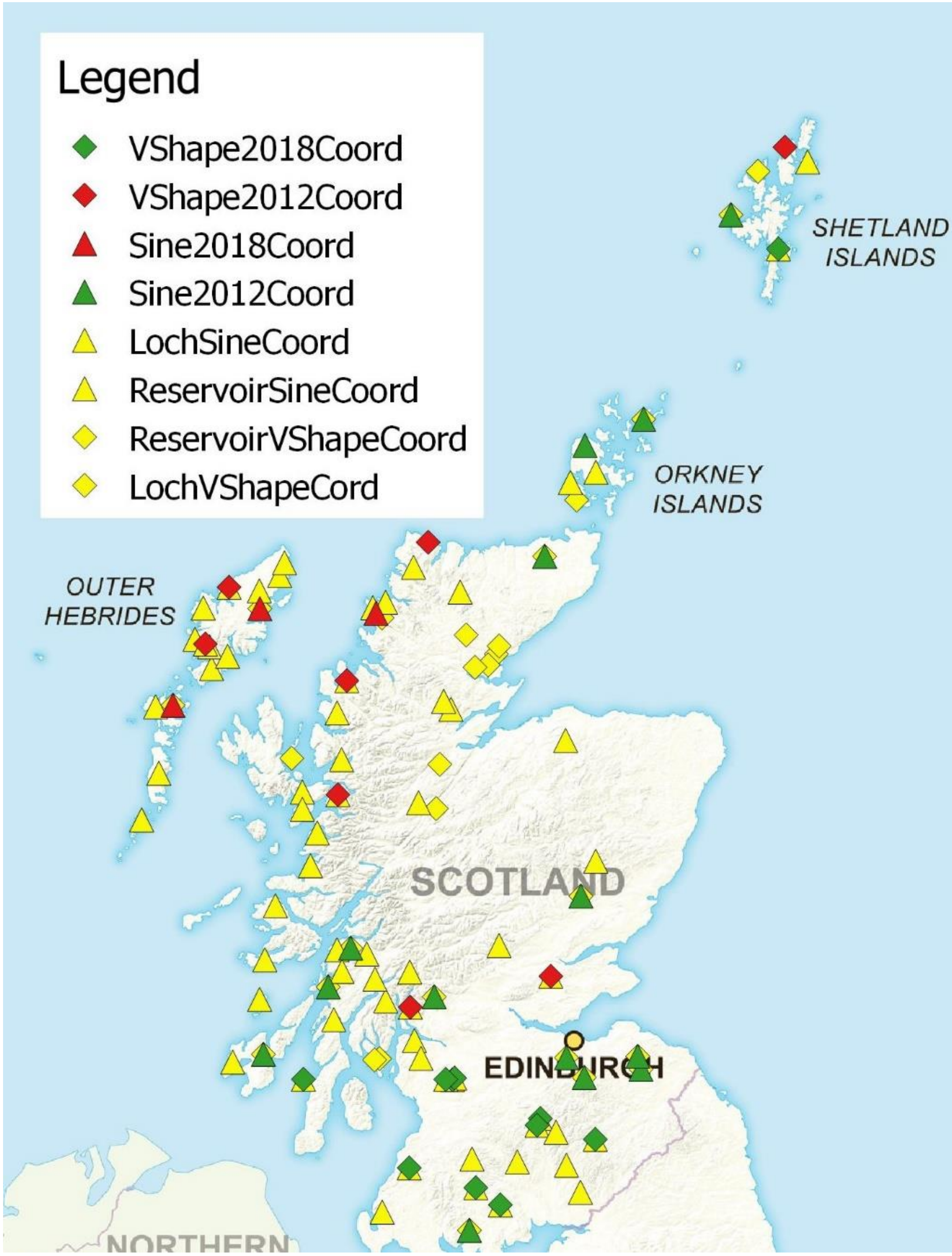


Figure 5.7: Spatial representation of comparison of catchment seasonal profile in a wet (2012) or dry (2018) year, based on allocation to a cluster using shape-based clustering. Green diamond = catchment switching from normally "sine-shaped" to "V-shaped" in 2018 (behaving as expected under the hypothesis that the "V-shape is influenced by longer/delayed wetting up of the catchment); red diamond = catchment switching from a normally "sine"-shape to a "V-shape" in 2012 (behaving opposed to expected); red triangle = catchment switching from a normally "V-shape" to a "sine"-shape in 2018 (behaving opposed to expected); green triangle = catchment switching from normally "V-shaped" to "sine-shaped" in 2012 (behaving as expected).

From the analysis, it remains unclear which effects are contributing to the observed differences in concentration over the year, and if and what this would mean in the context of climate change. If photodegradation and organic uptake play a role in reducing the amount of DOC, increased radiation could lead to reductions in concentrations (Chapman & Palmer, 2016). This could lead to further reductions in DOC in these catchments where the effect is already observable. On the other hand, algae can contribute to DOC production (Wyatt et al., 2012), so increases in algae productivity due to climate change might outbalance this effect. This may be dependent for example on nutrient availability, so a more detailed, multi-indicator analysis might be needed to get a clearer picture for these catchments.

Further analysis would also be needed to determine how much wetting up processes contribute to the observed pattern, and if differences in overall concentrations and concentration peaks between these groups are linked to this pattern or to other catchment characteristics (such as the organic carbon pool in the catchment, the hydrology of the catchment, etc.). This would help to determine if changes in concentration patterns due to e.g., decreases in precipitation in the summer could influence overall concentrations or peak concentrations, and which catchments would be most sensitive.

5.3.3 Climate sensitivity analysis

Catchments were sorted into a "sensitivity" category by looking at the correlations to the chosen climate variables (3 days antecedent (short) rainfall, 60 days antecedent (long) rainfall, and 60 days mean temperature). Significant spearman's rho in the season when peak concentrations occurred were looked at first, with moderate or strong correlation to short rainfall leading to allocation to the category "Rainfall", correlation to short rainfall and temperature to "Rainfall+ Temperature", correlation to long rainfall to "Wetup", and correlation to temperature to "Temperature". If there were no significant moderate or strong correlations, then the season before the peak would be considered, then the season after. If the peak season only correlated to short rainfall, but the temperature was also correlated in the season before or after, then the category would be "Rainfall+Temperature". If there were no correlations in these seasons, or the correlations were inconclusive, the category would be "None". If both short rainfall and long rainfall correlated, the category would be

“Rainfall”, if both long rainfall and temperature correlated, the category would be “Wetup”. In borderline cases, the overall correlations would be looked at to see if there was an indication which category would be a better fit. Some negative correlation also occurred, mainly for long rainfall in winter or autumn, occasionally also for short rainfall, or for temperature in autumn or spring. There was no case of consistent negative correlations across a catchment, and the negative correlations were therefore not regarded for category allocation. Correlations with a rho value below 0.4 were also not regarded, even if statistically significant.

Setting the significance value at <0.05 meant that likelihood of familywise errors was very high, because a great number of individual Spearman’s rank correlation tests were done. With 12 tests each on 127 catchments and a p-value for each test set at <0.05, the likelihood of making a type I error (rejecting the null hypothesis when it should be accepted) was almost 100%. Controlling the error with this number of tests would however mean that an extremely low p-value would have to be chosen, greatly increasing the likelihood of making a type II error (accepting the null hypothesis when it should be rejected). Even a p-value that controls the likelihood of a type I error for one catchment (i.e.,12 tests) at 0.01 would greatly reduce the number of statistically significant relationships, meaning that most catchments would fall into the “None” category. It was judged that for this exploratory analysis, the high risk of making a type I error would be better than a high risk of making a type II error, i.e., it would be better to allocate some catchments to the wrong category, rather than be able to only distinguish very few catchments for each remaining category. Hence the familywise error wasn’t controlled. This seems to be more congruent with a tiered risk assessment approach and especially a first risk screening, where it would be better to overestimate rather than underestimate risk.

The results of the Spearman’s rank correlations, and the final allocation into one of five categories (“Temperature” – 18 catchments, “Rainfall” - 18, “Rainfall+Temperature” - 22, “Wetup” - 49, “None” - 20) are shown in Table C.5.

The PCA on Spearman’s rho shows that the correlations between TOC concentrations and the climate variables form quite strong clusters (Figure 5.8). The first 2 principal components explain 47% of the variation in the data (26% and 21% respectively). There is a grouping for the temperature correlations (lower left quadrant), opposed to correlations with longer periods of rainfall in summer and autumn (upper right quadrant), and an unrelated grouping for correlations to the shorter periods of rainfall and the longer periods of rainfall in winter and spring (upper left quadrant). The catchments accordingly concentrate in the lower part of the plot for the “Temperature” group, the left part for the “Rainfall” and “Rainfall+Temperature” group, and the right part for the “Wet up” group. The “None” catchments also predominantly occur in the right part of the biplot. The biplot also shows that catchments distribute quite evenly around the origin of the biplot and there is overlap between the groups, which is to be expected as the PCA took all rho values into account regardless of their significance. On the whole, the PCA seems to confirm the manual grouping.

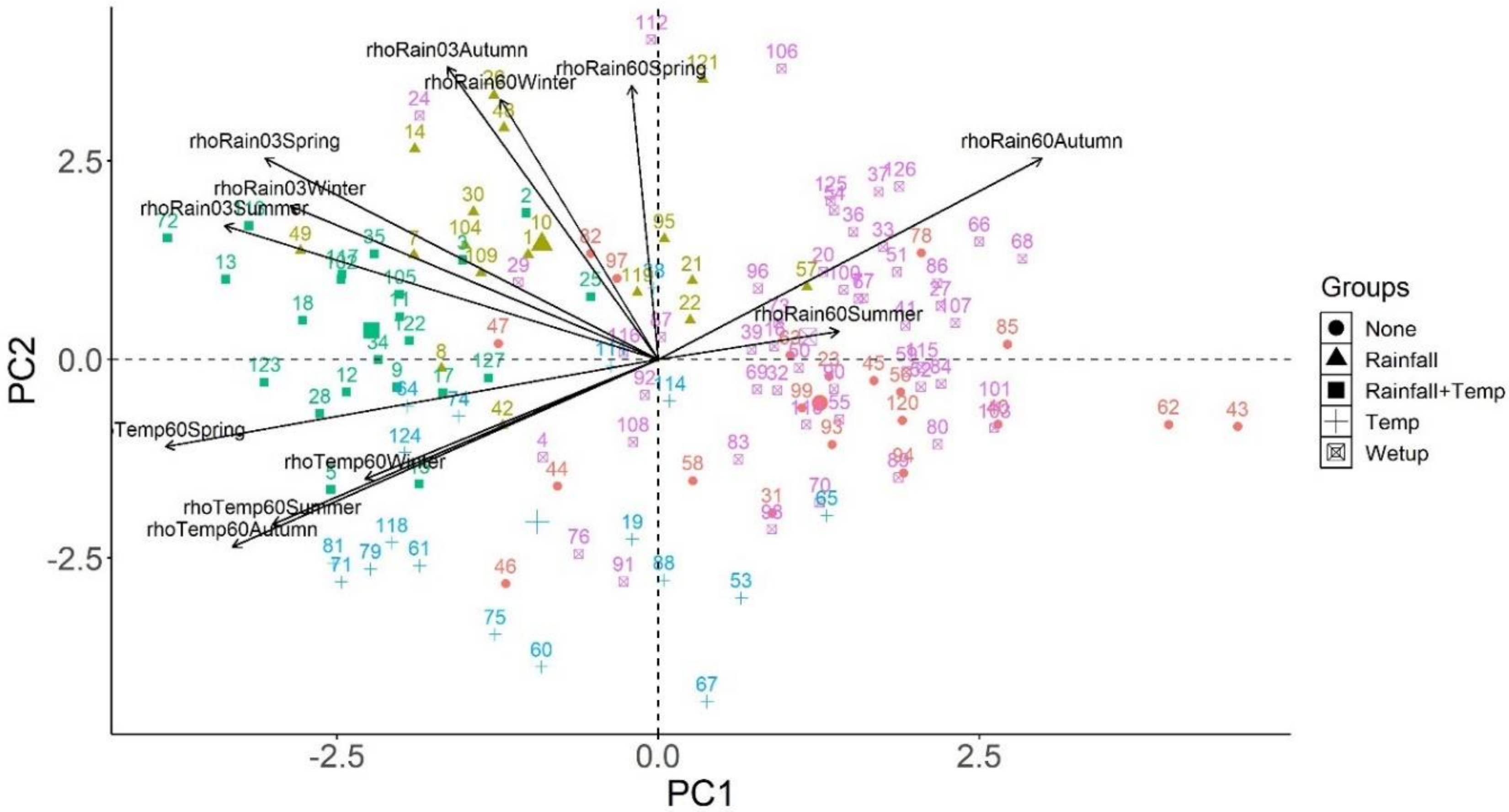


Figure 5.8: Biplot of the first two PCs (explaining 47% of the variation in the data) of a PCA based on the rho values of Spearman's rank correlation tests run per catchment on sampled TOC concentrations from 2013 – 2016 and corresponding 3-day antecedent rainfall, 60-day antecedent rainfall, and 60-day antecedent mean temperature, split by season. Catchments are symbolised according to the manually identified category ("None" = no correlation to the climate variables observable; "Rainfall" = positive correlation to 3-day antecedent rainfall; "Rainfall+Temp" = positive correlation to 3-day antecedent rainfall and 60-day antecedent mean air temperature; "Temp" = positive correlation to 60-day antecedent mean air temperature, "Wetup" = correlation to 60-day antecedent rainfall).

I. Group characterisation

The most apparent difference between the categories is that the “Rainfall” and “Rainfall+Temperature” categories are almost all river catchments. This could be due to a buffering effect for reservoirs and lakes, where any high concentrations that come out of the catchment following a storm event are mixed with lower concentrations before they reach the intake. The fact that two of the three river catchments within the “Temperature” category are large river catchments would support this.

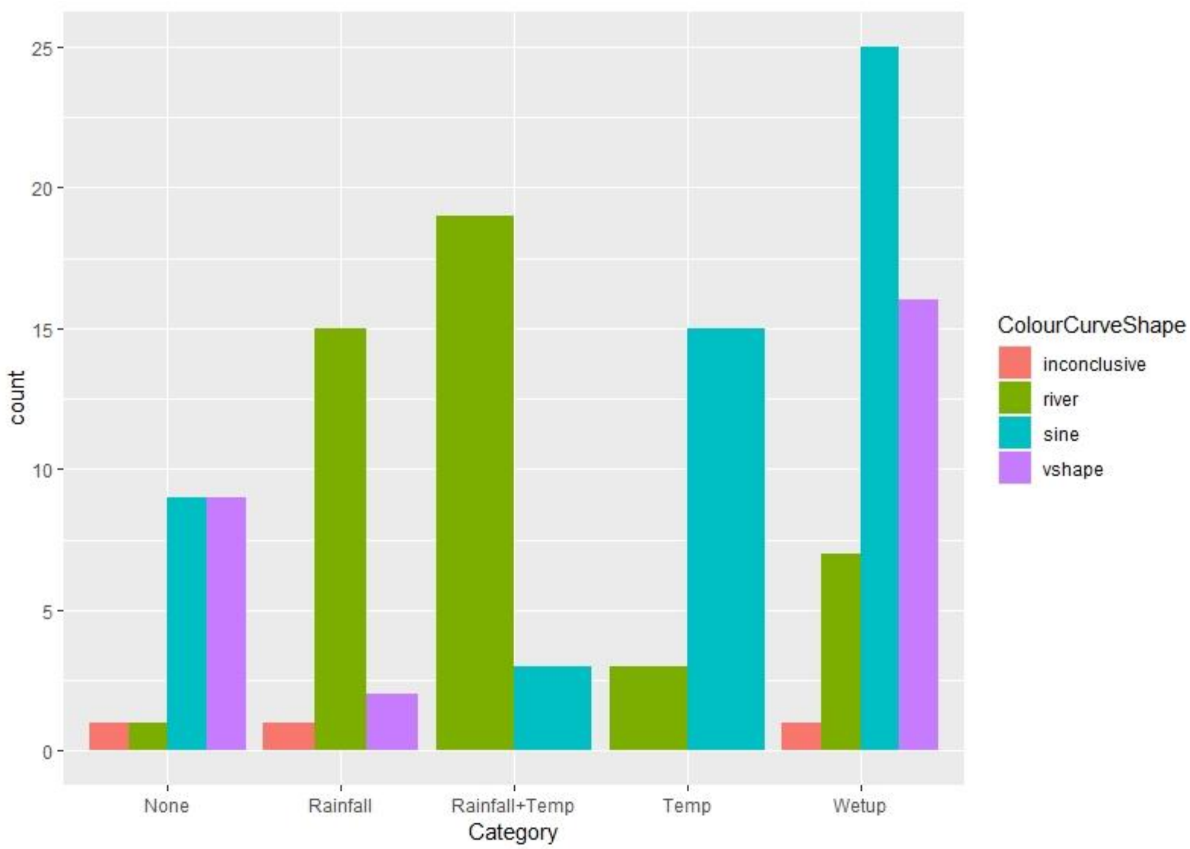


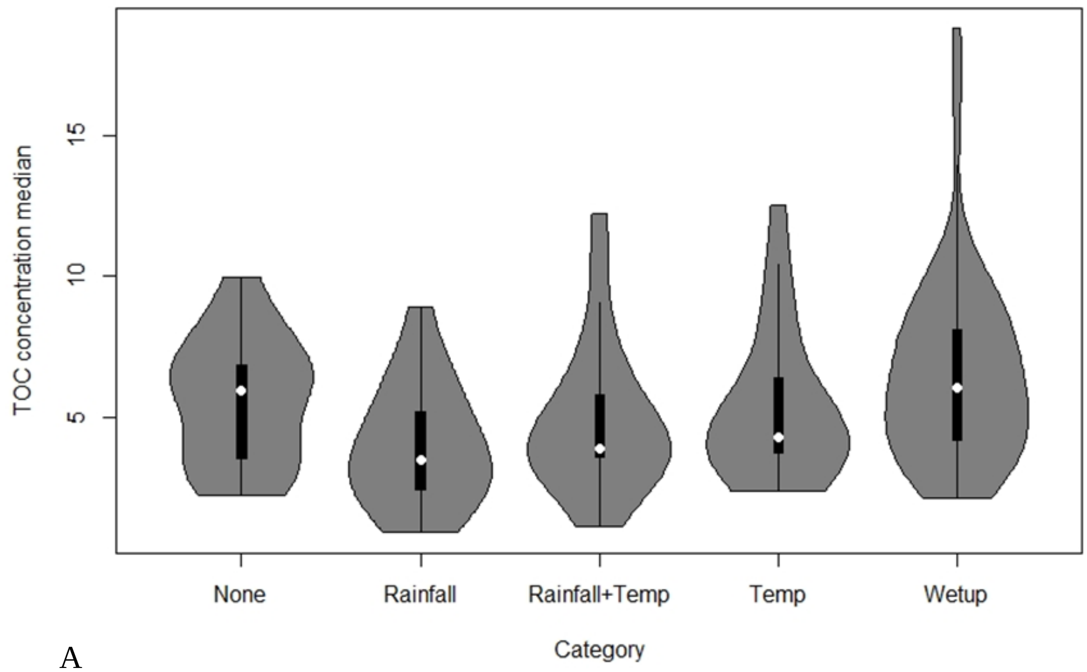
Figure 5.9: Number of catchments in each category (allocated through the results of the Spearman's rank correlation tests run per catchment on sampled TOC concentrations from 2013 – 2016 and corresponding 3-day antecedent rainfall, 60-day antecedent rainfall, and 60-day antecedent mean temperature, split by season; “None” = no correlation to the climate variables observable; “Rainfall” = positive correlation to 3-day antecedent rainfall; “Rainfall+Temp” = positive correlation to 3-day antecedent rainfall and 60-day antecedent mean air temperature; “Temp” = positive correlation to 60-day antecedent mean air temperature, “Wetup” = correlation to 60-day antecedent rainfall), by cluster allocated in the shape-based clustering (“inconclusive” = catchment that could not be allocated, “river” = river catchment, “sine” = catchment with seasonal profile that dips in spring and peaks in autumn, “vshape” = catchment with seasonal profile that dips in summer and peaks in winter).

Examining the colour curve shape categories, the “Temperature” category consists of mostly of “sine-shaped” lochs and reservoirs. Most of the “v-shape” lochs and reservoir belong in the “Wetup” or “None” category (Figure 5.9).

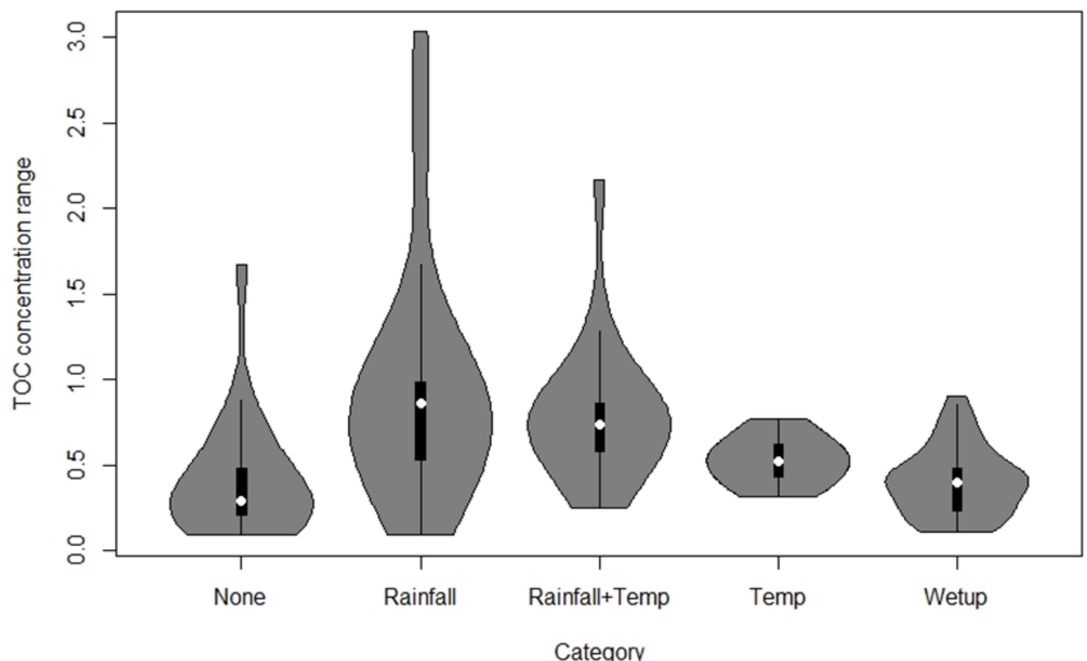
While there is no overall pattern of finding very high or very low median TOC concentrations in specific groups, the highest TOC median concentrations can be found in the “Wetup” group, while the lowest median concentrations can be found in the “Rainfall” group (Figure 5.10A). The difference in mean is statistically significant between these two groups ($p=0.015$, Wilcoxon test). When looking at the variability of concentrations, the “Rainfall” group shows highest variability in terms of difference between maximum and minimum concentrations, normalised against the mean ((TOC maximum concentration-TOC minimum concentration)/TOC mean concentration) (Figure 5.10B). Statistically, the difference in mean is only significant between the “Rainfall” and the “Temperature” groups and the “Temperature” and the “Wetup” groups.

The fact that the variability seems to be higher in the “Rainfall” and maybe also the “Rainfall+Temperature” groups would fit together with the hypothesis that there is a buffer effect in lochs and reservoirs, whereas sudden spikes in concentrations following high rainfall events cannot be buffered in rivers.

There are further distinctions when looking at the catchment profiles between these categories. Average topsoil organic carbon content is lower in most catchments in the “None” and “Rainfall” group, with statistically significant differences ($p<0.05$, tested with a Wilcoxon test) between the ”None” and the “Rainfall+Temperature”, and the “Wetup” groups, as well as between the “Rainfall” and “Rainfall+Temperature” groups (Figure 5.11A).



A



B

Figure 5.10: Violin plots showing the distribution of A. TOC median concentrations and B. TOC concentration variability for the catchments within each category allocated through the results of the Spearman's rank correlation tests run per catchment on sampled TOC concentrations from 2013 – 2016 and corresponding 3-day antecedent rainfall, 60-day antecedent rainfall, and 60-day antecedent mean temperature, split by season; ("None" = no correlation to the climate variables observable; "Rainfall" = positive correlation to 3-day antecedent rainfall; "Rainfall+Temp" = positive correlation to 3-day antecedent rainfall and 60-day antecedent mean air temperature; "Temp" = positive correlation to 60-day antecedent mean air temperature, "Wetup" = correlation to 60-day antecedent rainfall).

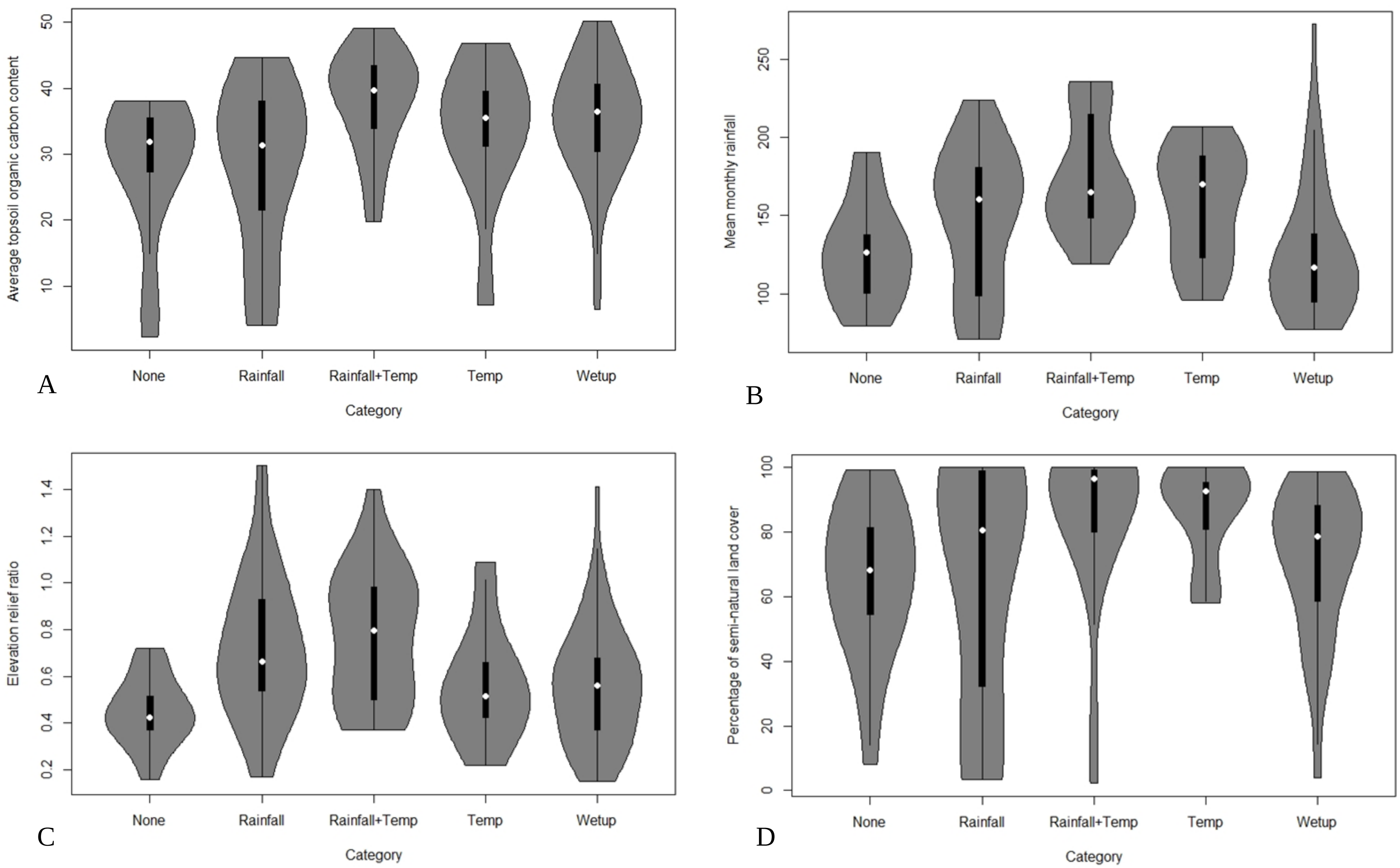


Figure 5.11: Violin plots showing the distribution of A. average amount of organic carbon in the topsoil, B. amount of mean monthly rainfall in mm (long-term average 1981-2010), C. elevation relief ratios, and D. percentage of semi-natural land cover for the catchments within each climate sensitivity category allocated through the results of the Spearman's rank correlation tests run per catchment on sampled TOC concentrations from 2013 – 2016 and corresponding 3-day antecedent rainfall, 60-day antecedent rainfall, and 60-day antecedent mean temperature, split by season; (“None” = no correlation to the climate variables observable; “Rainfall” = positive correlation to 3-day antecedent rainfall; “Rainfall+Temp” = positive correlation to 3-day antecedent rainfall and 60-day antecedent mean air temperature; “Temp” = positive correlation to 60-day antecedent mean air temperature, “Wetup” = correlation to 60-day antecedent rainfall).

Rainfall amounts also differ between the groups, with lower amounts of rainfall in the “None” and the “Wetup” groups, although the latter shows wide variability. Differences in mean are statistically significant between the “None” and the “Rainfall+Temperature” and “Temperature” groups, and between the “Wetup” and the “Rainfall+Temperature” and “Temperature” groups (Figure 5.11B). There is no significant difference in mean annual temperature between the groups.

In terms of topography, the main observable difference seems to be in the “Rainfall” and “Rainfall+Temperature” groups, with more catchments with higher elevation relief ratios (Figure 5.11C). These observable differences are statistically significant between these two groups and the “None” group, as well as between the “Wetup” and the “Rainfall+Temperature” group. Elevation relief ratio is a method to calculate hypsometric integral, which in turn can be used to estimate erosion status, with higher values being more prone to erosion (Singh et al., 2008). Higher values for elevation relief ratio could hence mean a quicker transfer of precipitation to the water body, which would increase the sensitivity to small rainfall events as rainfall would tend to produce surface and shallow subsurface runoff, increasing the transfer of DOC produced in the upper aerobic part of the organic soil horizon. In contrast, shallower reliefs could either mean a longer residence time and a higher chance of infiltration into the mineral horizon, where DOC may be retained, or conditions for a build-up of thicker organic soil layers producing more DOC.

It is also observable that semi-natural land cover is higher in the “Temperature” and “Rainfall+Temperature” categories (Figure 5.11D), with statistically significant differences in mean between these two groups and the “None” and “Wetup” groups.

Just going by these few catchment characteristics, the catchments in the “Rainfall+Temperature” and the “Temperature” categories seem to be the wetter ones with higher organic carbon pools and more natural landcover. While the “Rainfall+Temperature” mainly contains river catchments, the “Temperature” group mainly contains lochs or reservoirs, otherwise they seem to resemble each other in terms of catchment characteristics. It could be hypothesised that these are catchments with well-connected carbon pools that produce DOC throughout the warmer months, which gets regularly washed out into the surface water source due to generally wetter soil conditions.

This would explain the correlation to temperature. The sensitivity to short term rainfall could then stem from a lower buffering capacity in the river sources.

Further interesting insights might be derived from the comparison between the “Rainfall” and the “Rainfall+Temperature” groups, which together contain the majority of river catchments. In contrast to what Winterdahl et al. (2014) found, rivers seem to almost always respond to storm events. However, the relationship to temperature grows weaker in catchments with lower organic carbon content. Interestingly, in terms of TOC concentrations, the groups do not differ very much either in median concentrations or concentration variability, which might have been expected if there are differences between the size of the organic carbon pools in the catchment.

The “Wetup” group contains the largest number of catchments, and some of these catchments also showed a correlation to temperature, making the distinction between these groups more difficult. It is interesting that the difference in their TOC concentration variability is statistically significant, with the “Wetup” catchments being slightly more variable. The most notable difference between these two groups is that the “Wetup” catchments receive less rainfall, which could support the hypothesis that the produced DOC builds up in the soil over the summer and only gets washed out once the catchments get wetter towards the end of the year. This could fit together with the fact that while the majority of catchments in this category belong to the “sine” shaped group, a considerable number was characterised as “V-shape” in the shape based clustering, which was also hypothesised to have to do with wet-up processes (see 5.3.2). The “None” catchments, while also receiving less rainfall, also have lower organic carbon pools than the “Wetup” or “Temperature” groups. The carbon pool is strongly associated with TOC concentration, but the median TOC concentrations within the group are not lower than in comparison to the “Temperature” or the “Wetup” group. As discussed, higher concentrations can also be due to in-lake processes such as DOC generation by phytoplankton and macrophytes. This might play a role here, although a correlation to temperature might then be expected, as a 60 day mean air temperature would also be a proxy for water temperature, which in turn influences organic production of DOC.

Spatial distribution shows no clear pattern (Figure 5.12). “Temperature” and “Rainfall+Temperature” catchments are mainly found along the West coast. This goes well with the pattern emerging from the above observations that these catchments would tend to be peaty, wet catchments with semi-natural land cover. The difference between these two groups is mostly the different type of water body (river vs loch/reservoir). We also observe “Rainfall” and “Rainfall+Temperature” catchments in close proximity, so we can assume that differences in climate are not an important factor here – local conditions would probably play a bigger role, such as topography, aspect, soils, and land use and management. Data describing the catchment may not be sufficiently detailed to observe these differences. This also goes for comparisons between the “Temperature” and “Wetup” catchments.

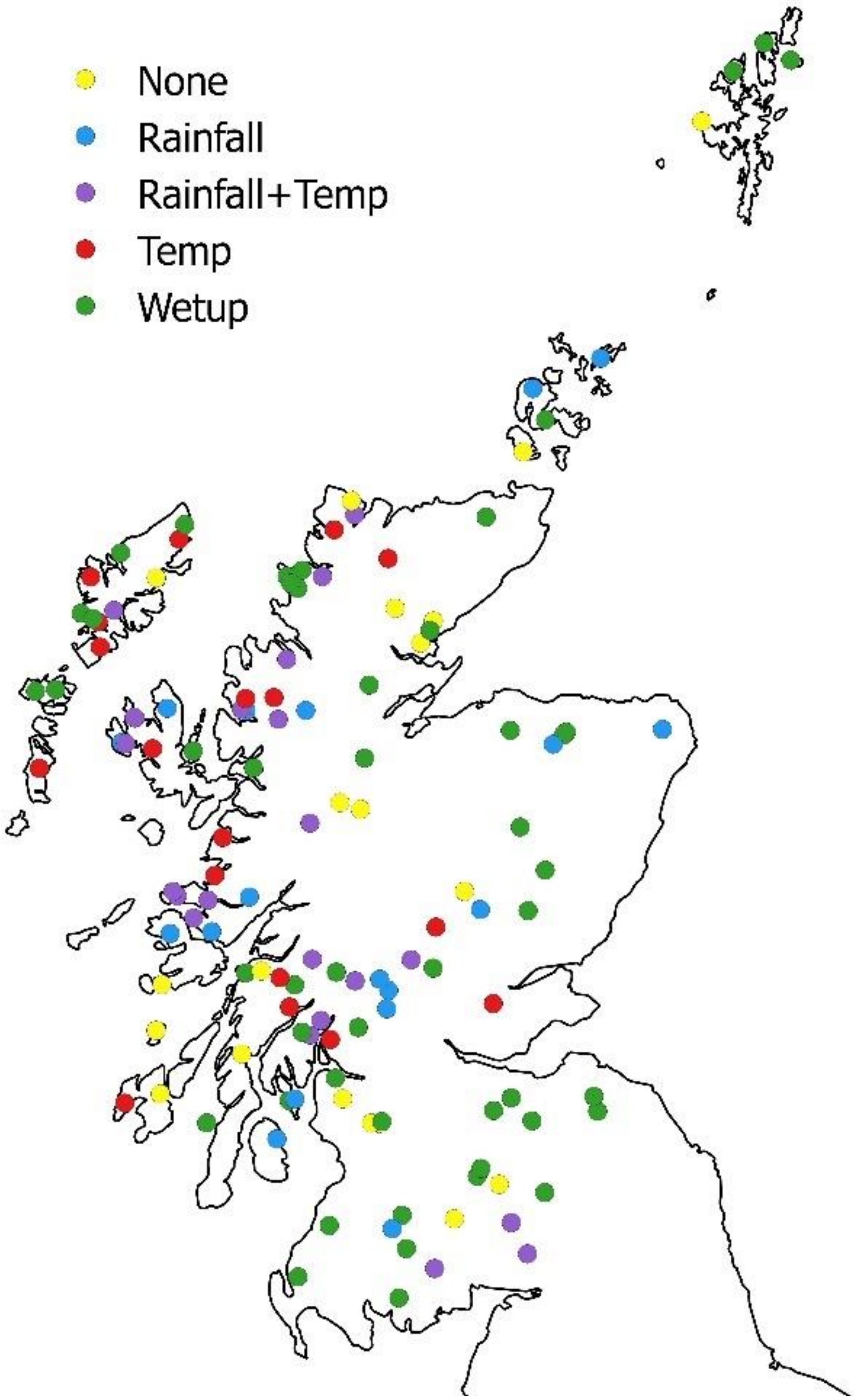


Figure 5.12: Spatial distribution of catchment categories allocated through the results of the Spearman's rank correlation tests run per catchment on sampled TOC concentrations from 2013 – 2016 and corresponding 3-day antecedent rainfall, 60-day antecedent rainfall, and 60-day antecedent mean temperature, split by season; ("None" = no correlation to the climate variables observable; "Rainfall" = positive correlation to 3-day antecedent rainfall; "Rainfall+Temp" = positive correlation to 3-day antecedent rainfall and 60-day antecedent mean air temperature; "Temp" = positive correlation to 60-day antecedent mean air temperature, "Wetup" = correlation to 60-day antecedent rainfall).

II. Relationships to catchment characteristics

The variables that were identified for the optimum model for the redundancy analysis were elevation relief ratio, relief ratio, percentage of semi-natural land cover, percentage of coniferous forest cover, percentage of arable area, percentage of improved grassland cover, amount of SER, AAT above 5.5°C, and median pH. The final redundancy analysis additionally used area and average topsoil organic carbon content as they were judged to be critical to help interpret the findings. There was some correlation between the independent variables originally included in the model (Figure 5.13), some of which were retained in the final model. However, as this modelling served as a first exploration of explanatory nature of the catchment characteristics, this was judged not to be an issue.

An ANOVA showed that area, percentage of coniferous forest cover, percentage of arable area, and average topsoil organic carbon were not significant variables in the final model. The first two axes of the redundancy analysis (Figure 5.14) explained 84% of the constrained variation, but only 24% of the overall variation. The short rainfall relationships showed a correlation with topographic indexes, reiterating the findings from above and suggesting a relationship between steeper reliefs to dominance of storm event driven DOC export. They also show a negative correlation to AAT and area, suggesting cooler temperatures and smaller catchments (Table 5.3).

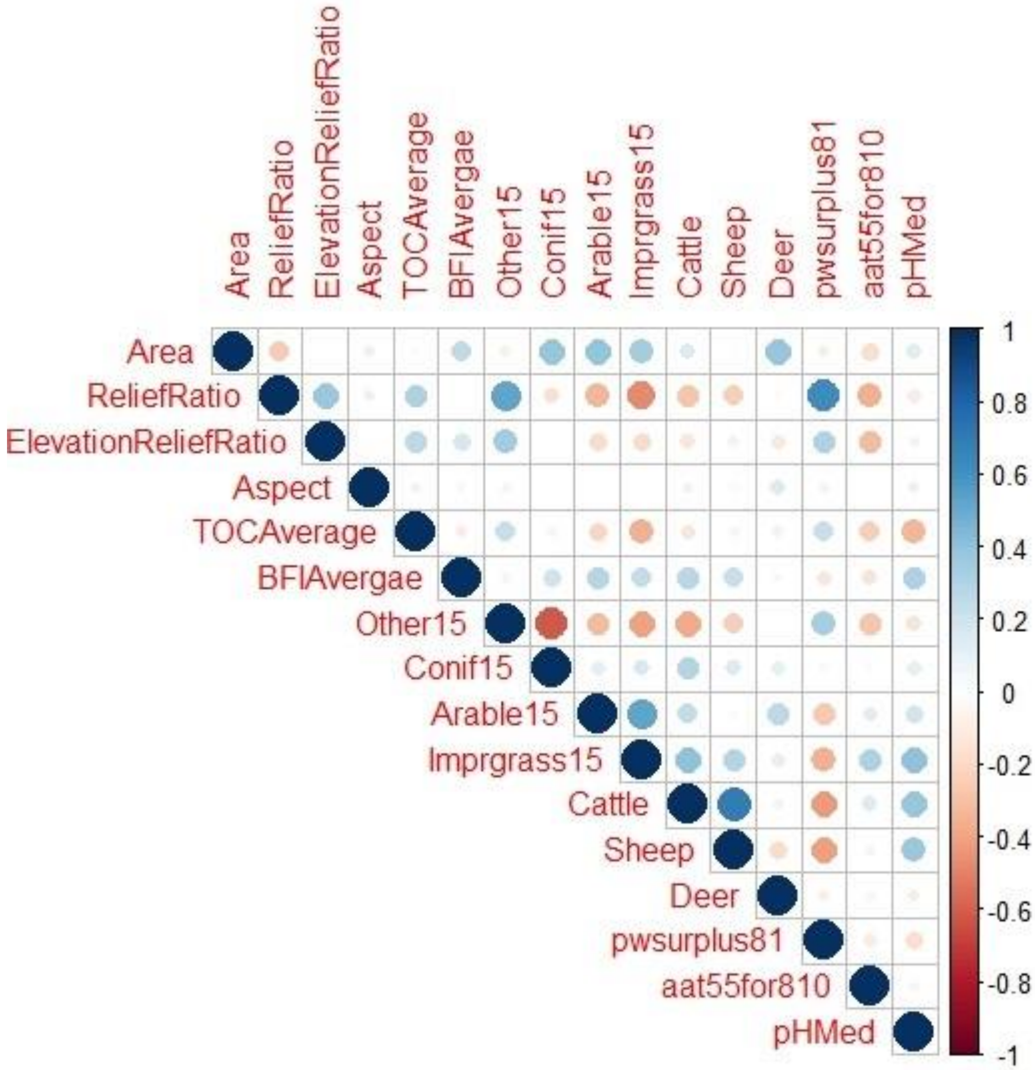


Figure 5.13: Correlation matrix based on Spearman's rank correlation for constraining variables used in the redundancy analysis to determine explanatory power of catchment characteristics on catchment rho values from the Spearman's rank correlation tests run per catchment on sampled TOC concentrations from 2013 – 2016 and corresponding 3-day antecedent rainfall, 60-day antecedent rainfall, and 60-day antecedent mean temperature, split by season. Area = Area in km^2, ReliefRatio = Relief ratio, ElevationReliefRatio = Elevation relief ratio, Aspect = Percentage of aspects facing South and Southwest, TOCAverage = Average topsoil organic carbon content, BFIAvergae = Average BFI in the catchment, Other 15 = Percentage of semi-natural land cover, Conif15 = Percentage of coniferous forest, "Arable15 = Percentage of arable area, Imprgrass15 = Percentage of improved grassland, Cattle = Average density of cattle in the parish, Sheep = Average density of sheep in the parish, Deer = Number of deer in the catchment, pwsurplus81 = SER for the period 1981-2000, aat55for810 = AAT for the period 1981-2000, pHMed = median pH value of the water quality sample data for the catchment.

Table 5.3: RDA biplot scores from the redundancy analysis using catchment characteristics as constraining variables and catchment rho values from the Spearman's rank correlation tests run per catchment on sampled TOC concentrations from 2013 – 2016 and corresponding 3-day antecedent rainfall, 60-day antecedent rainfall, and 60-day antecedent mean temperature, split by season, as dependent variables (n=127).

	RDA1	**RDA2**
Spearman's rho climate sensitivity analysis	Scores	
Short rain – spring	-0.35	0.37
Short rain – summer	-0.42	0.21
Short rain – autumn	-0.13	0.55
Short rain- winter	-0.16	0.4
Long rain – spring	0.21	0.45
Long rain – summer	0.19	0.16
Long rain – autumn	0.54	0.21
Long rain – winter	-0.05	0.4
Temp – spring	-0.81	0.12
Temp – summer	-0.38	-0.19
Temp – autumn	-0.7	-0.25
Temp - winter	-0.35	0.05
Catchment characteristics	Biplot scores	
Area	0.06	-0.09
Elevation relief ratio***	-0.51	0.54
Relief ratio***	-0.71	0.29
Semi-natural*	-0.41	-0.16
Coniferous forest	0.01	0.18
Arable	0.19	0.3
Improved grassland***	0.25	0.26
Topsoil organic carbon content	-0.32	0.02
SER***	-0.81	-0.19
AAT**	0.11	-0.35
pH median***	0.28	0.55

Smaller catchments could certainly be expected to react more quickly to rainfall events, but the effect is only small. The relationships to temperature were correlated to SER amounts, percentage of semi-natural land cover and average topsoil organic carbon. This is again in agreement with findings above suggesting these catchments occur in wetter and less intensively used areas with larger organic carbon pools. Catchments related to longer term rainfall amounts showed correlations to percentage of arable area and improved grassland, as well as median pH values. This suggests more intensive agricultural usage, less acidic soils, and drier conditions. It cannot be deducted if these characteristics would contribute to the observed effect of delayed TOC peaks, or if the catchments with processes and characteristics that lead to this pattern are naturally more capable of sustaining agricultural land uses.

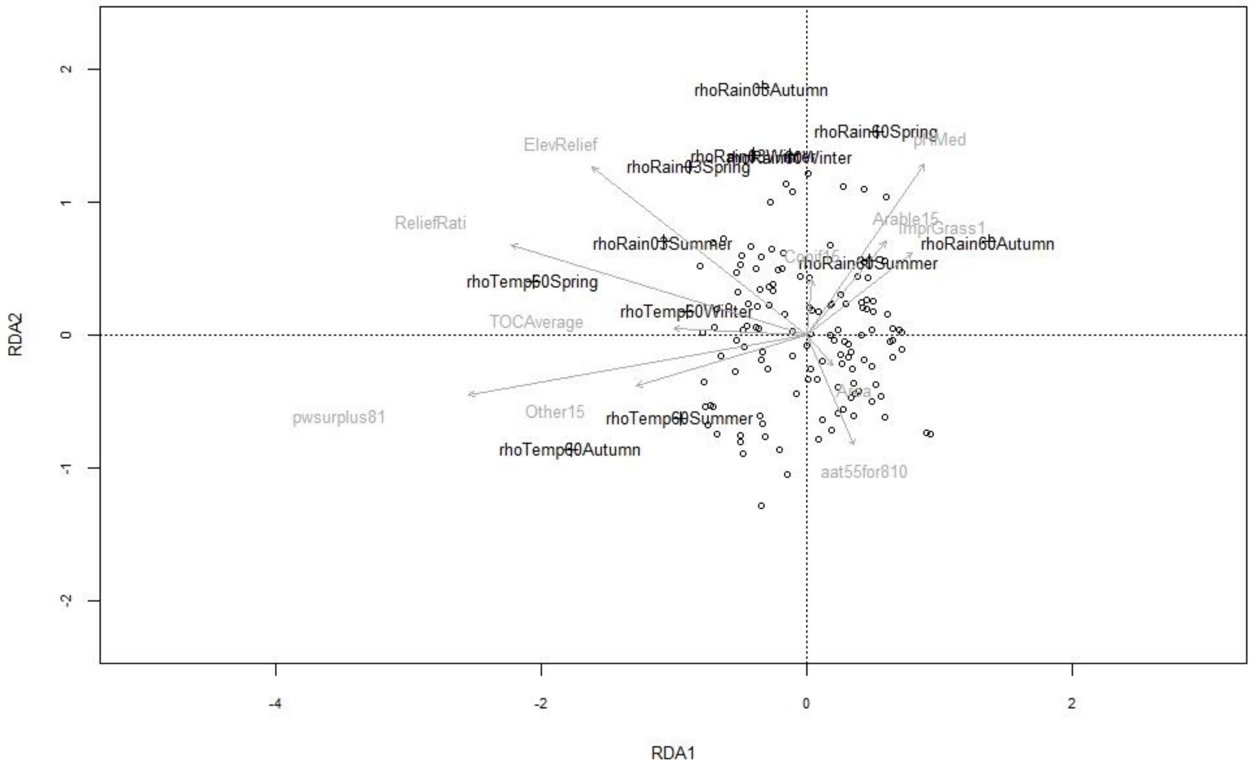


Figure 5.14: Triplot from the redundancy analysis using catchment characteristics as constraining variables and catchment rho values from the Spearman's rank correlation tests run per catchment on sampled TOC concentrations from 2013 – 2016 and corresponding 3-day antecedent rainfall, 60-day antecedent rainfall, and 60-day antecedent mean temperature, split by season, as dependent variables (n=127). Water quality parameters symbolised in black writing, catchment characteristics as constraints symbolised as grey rays, and individual catchments symbolised as points. Points/variables in the right of the plot correlate positively with RDA1, to the left – negatively. Points/variables in the top correlate positively with RDA2, in the bottom – negatively. The longer the arrow/further from the plot origin the point, the stronger the correlation. ReliefRatio = Relief ratio, ElevRelief = Elevation relief ratio, pHMed = median pH value of the catchment water quality sample data, Conif15 = Percentage of coniferous forest cover, Arable15 = Percentage of arable area, ImprGrass15 = Percentage of improved grassland, Other15 = Percentage of semi-natural land cover, TOCAverage = Average topsoil organic carbon content, pwsurplus81 = SER for the period 1981-2000, aat55for810 = AAT for the period 1981-2000.

In summary, the findings suggest that the “Temperature” and “Rainfall+Temperature” catchments tend to be peaty, wet catchments with semi-natural land cover. The correlation of relief ratio and elevation relief to the shorter period of rainfalls indicates that sensitivity to short-term rainfall events increases with steeper reliefs; it could also be a correlation to peat depth, as we would expect less peat to accumulate on steeper slopes, offering a smaller pool of organic carbon (Parry et al., 2015). It also remains unclear if more intensive land use contributes to catchments being in the “Wet up” category, and if drainage leads to the observed effects, or if the catchments under such land use would naturally be better drained catchments, and so need a longer time period to wet up and flush out DOC. The role of in-lake processes also remains unclear, especially as several factors and drivers could be overlapping.

5.3.4 Catchment characteristics relationships to median TOC concentrations

DOC/TOC concentrations vary throughout the year in most catchments, with variability within the year often greater than between years (Musolff et al., 2018). Median concentrations therefore not necessarily reflect how problematic a catchment is in terms of water treatment, as it is peak concentrations that will usually cause problems. However, median concentrations can be used to represent a “baseline” for a catchment and show which catchments routinely produce higher concentrations. While the main driver for this is the soil carbon pool in the catchment (see 4), other catchment conditions will influence how much DOC reaches the water body. Climate change could directly alter processes in the catchment leading to changes in concentrations, or indirectly influence factors that in turn change mechanisms in either the source, the pathways, or the surface water body and lead to changes in TOC concentrations. Looking at drivers for the “baseline”, median TOC concentrations depending on the identified different catchment sensitivities could help to understand how climate change might influence catchments differently and to project changes in risk.

I. Regression per group

If assuming that in the different groups of catchments identified through the correlations between TOC concentrations and climate variables, different mechanisms take place that lead to the production and transfer of DOC, these groups should be treated separately when trying to find relationships of catchment characteristics to median TOC concentrations. A model was also developed for all the catchments together for comparison.

All groups, with the exception of the “None” category, produced acceptable models in terms of variation explained and cross validation indicators (Table 5.4). For some groups (“Temperature” and “Rainfall”), the modelling process of stepwise backward variable elimination by BIC produced models that were not significant. In this case, variables were removed one at a time based on potential collinearity, until an acceptable model was produced. The original “None” model is significant and explains a high proportion of variation, but normalised RSME suggests it is very overfitted. Removing independent variables did not lead to a more acceptable model.

In the overall model, average topsoil organic carbon, percentage of improved grassland cover, the average number of sheep, and AAT positively correlate with median TOC while area, SER, and median pH negatively correlate. Of these, it at first seems surprising that improved grassland would increase TOC concentrations, as it can be assumed that these areas are less rich in organic soils, and grazing decreases soil organic carbon (Eze et al., 2018). On the other hand, it is possible that more intensive use creates disturbances leading to increases in suspended sediment and TOC (Glendell & Brazier, 2014). Furthermore, it has been shown that woodland and farmland can substantially contribute to DOC (Ritson et al., 2019), and that DOC draining from farmland tends to be less coloured and older (~300-700 years), in comparison DOC draining from peat dominated soils (~10-15 years), with the release of this older DOC possibly due to land management rather than climate factors (Evans, Freeman et al., 2006). In terms of climate, the model suggests that warmer catchments with longer growing seasons produce more TOC, and that higher summer rainfall could have a diluting effect, leading to lower TOC concentrations.

Table 5.4: Multiple linear regression models per climate sensitivity category with TOC median concentration as dependent variable. Adjusted R2, p-value and normalised RSME from 10 times 10-fold cross validation are provided as measures for model performance.

Category	Number of catchments	Regression model	Adjusted R^2	P value	Cross validation: RSME/(Max-Min)
All	127	log(TOC median) = 2.4*** - 0.0003xArea**-2.37xRelief ratio** + 0.02xTopsoil organic carbon*** + 0.01xImproved grassland cover** + 0.001x#Sheep* - 0.0005xSER* + 0.0006xAAT** - 0.27xMedian pH	0.47	$8xe^{-15}$	0.13
Temperature *(Median pH removed)*	18	log(TOC median) = 0.61 + 1.53xElevation relief ratio** + 0.03xAspect** - 2.16xBFI**	0.53	0.003	0.19
Rainfall *(Arable area and improved grassland cover removed)*	18	log(TOC median) = 2.22*** – 4.66xRelief ratio* + 0.01xSemi-natural land cover** –0.003x#Sheep – 0.003xSER*	0.65	0.001	0.16
Rainfall+ Temperature	22	log(TOC median) = 0.6 + 0.006xAspect + 0.04xTopsoil organic carbon*** - 0.01xSemi-natural land cover*** + 0.01x#Cattle***	0.72	0.000008	0.1
Wetup	49	log(TOC median) = 2.9** – 0.001xArea** - 4.24xRelief ratio** + 0.02xTopsoil organic carbon content** + 0.03xImproved grassland cover*** + 0.001x#Sheep* + 0.001xAAT** - 0.37xMedian pH**	0.57	0.0000003	0.18
None	20	log(TOC median) = 3** - 0.0007xArea* + 3.19xElevationReliefRatio**+4.83xRelief ratio + 0.01xTopsoil organic carbon content + 1.18xBFI - 0.06xSemi-natural land cover** - 0.05xConiferous forest cover** +13.9xArable area - 0.1xImproved grassland cover** + 0.003x#Deer -0.005x#Sheep* + 0.02x#Cattle** + 0.002xAAT**	0.89	0.001	0.78

The “Temperature” group consists of mainly loch and reservoir catchments, with larger carbon pools and wetter conditions. Elevation relief ratio and percentage of south and southwest facing aspect are positively correlated. A higher elevation relief ratio indicates that the catchment has a higher proportion of area on higher elevations, and south and southwest aspects are more likely to capture precipitation. Baseflow index is negatively correlated. As higher DOC concentrations are associated with storm flow conditions (Buffam et al., 2001; Hood et al., 2006; Tiwari et al., 2014), a higher proportion of base flow could contribute to dilution of DOC concentrations. While it can be assumed that aspect is a precipitation related variable, the more direct climate variables, SER and AAT, are not included in the model. This is a little surprising as we could assume that especially within this group, a longer growing period would also lead to longer bacterial activity and hence more DOC production. However, this may reflect an inhibiting effect of wetness, with soil moisture limiting the amount of DOC produced (Clark et al., 2009), or a regular wash out into the water body. Similarly, if the conditions are generally quite wet, smaller differences in precipitation may not show in concentrations as DOC does not build up in the soil.

The “Rainfall+Temperature” group differs from the “Temperature” group in its sensitivity to short term rainfall and the predominance of river sources. In this group median TOC also correlates positively with percentage of south and southwest facing aspect, as well as average topsoil organic carbon content, and average number of cattle. The percentage of semi-natural land cover is negatively correlated. The latter two variables suggest that agricultural use of these areas could contribute to TOC. There is no evidence that grazing contributes to the release of DOC (Chapman & Palmer, 2016), but as we are looking at TOC, it is possible that cattle grazing contributes to particulates in the water through poaching. It is also possible that the catchments with higher percentages of agricultural land uses simply have other properties that increase TOC concentrations, or that it is indicative of a geographic pattern that has a range of effects.

The ”Rainfall” group consists mainly of river catchments with lower organic carbon pools. Here, relief ratio, average number of sheep and SER are negatively correlated to TOC median concentration, whereas percentage of semi-natural land cover is positively correlated. Steeper reliefs could indicate shallower layers of organic soils, and semi-natural land cover a higher organic carbon pool in the catchment. While it seems initially counterintuitive that a higher number of sheep would reduce TOC median concentrations, it might reflect that higher numbers of sheep would be kept on land with higher quality of plants for grazing, corresponding to richer and less organic soils (García et al., 2012). This model includes a climate variable, SER, which is negatively correlated, however it is unclear if this is a causal relationship, with higher rainfall amounts e.g., causing dilution, or if this is also related to other conditions.

For the “Wetup” group, area, relief ratio and median pH are negatively correlated, and average topsoil organic carbon content, percentage of improved grassland cover, average number of sheep, and AAT are positively correlated. The model is similar to the overall model. This group is distinguished from the “Temperature” group mainly by drier conditions. The positive correlation of improved grassland and sheep also suggests an effect of agricultural use. The positive correlation of AAT again could be due to a longer and more intensive growing period, inducing higher microbial activity, and higher primary productivity leading to more biomass, or to an increased likelihood of drought which also, in short term, leads to increased DOC concentrations (Clark et al., 2005). Alternatively, it could also indicate more DOC production within reservoirs and lochs from algae. It is observable that the catchments where we find a negative intercept in the TOC-colour models belong to the “Wetup” or “None” group, which we would expect to find under this hypothesis.

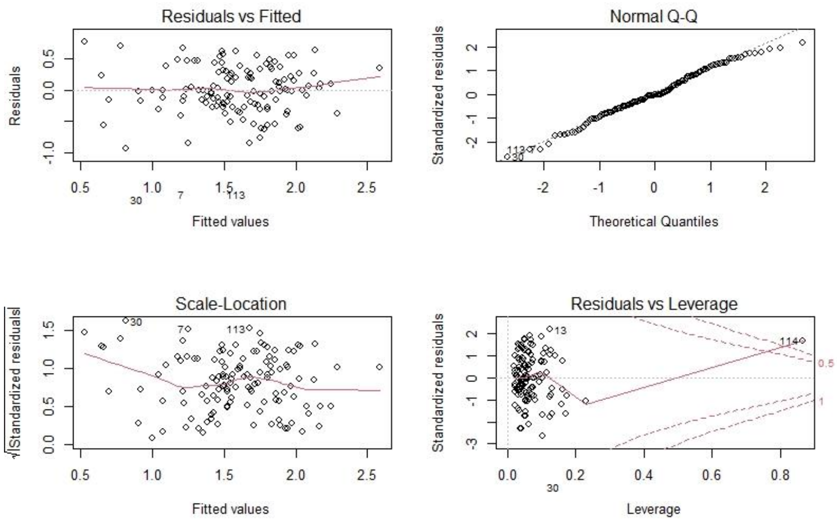
II. Pooled regression with interaction terms

The final model for median TOC with interactions that were identified through batchwise inclusion explained 57% of variability, which is an improvement over the model without interactions (47%). Visual inspection of residuals as well as cross validation also suggests it is a slightly better model (Figure 5.15). The only category interaction that was retained when including only category interactions in the model was between category and coniferous forest cover, however this interaction was removed during the modelling process when combined with the interactions from the other batches. The model that results when including only category interactions is shown in Table 5.5 for comparison.

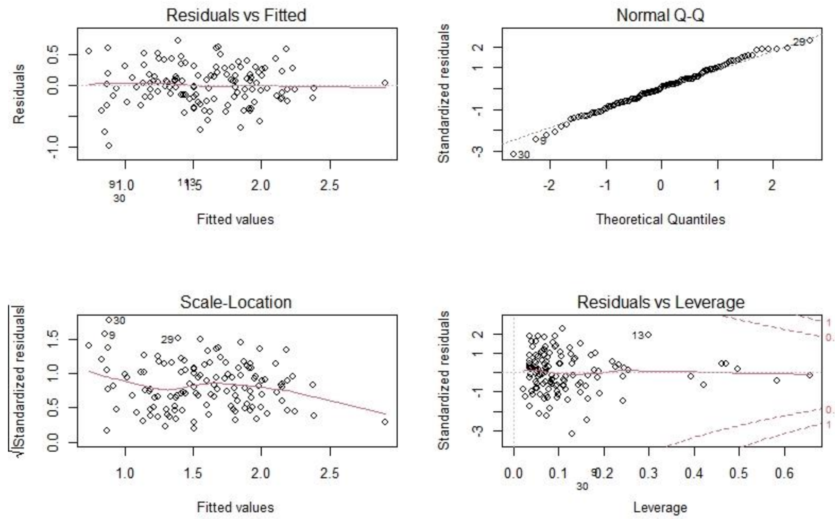
Table 5.5: Multiple linear regression models with log-transformed median TOC concentration as response variable (n=127; *** p<0.001, ** p<0.01, * p<0.05, - was eliminated during the modelling process, N/A not included in the modelling process).

Variable	**Coefficients and significance**			**Comment**
	1. No interactions	2. All interactions	3. Interactions – categories only	
Variability explained	47%	57%	54%	
RSME/(max-min)	0.13	0.11	0.13	
Category Rainfall Rainfall+Temperature Temperature Wetup	N/A	N/A	 -0.17 -0.27 -0.16 0.03	In comparison to the “None” category, catchments on the “Rainfall”, “Rainfall+Temperature” and “Temperature” group have slightly lower TOC median concentrations.
Area	-0.0002**	-	-	Larger catchments correlate to slightly lower TOC medians.
Relief ratio	-2.37**	-1.55*	-	Higher relief ratio correlates with lower TOC medians. This might be due to shallower depths of organic soils horizons in steeper catchments.
%age Southwest aspect	-	-	0.006*	Higher percentage of south and southwest facing aspects is assumed to correspond to wetter conditions.
Average topsoil organic carbon	0.02***	0.02***	0.02***	The higher the carbon content in the soil, the higher the TOC median.
% coniferous forest cover	-	0.006*	0.003	TOC medians increase with coniferous forest cover.
% improved grassland cover	0.001**	-	0.01**	TOC medians increase with improved grassland cover.
# deer		0.04	-0.007**	Higher deer numbers correspond to higher/lower TOC median concentrations.
# sheep	0.001*	-0.001	0.001**	Higher sheep numbers correspond to lower/higher TOC median concentrations.
SER	-0.0005*	0.01***	-0.0006**	Higher SER corresponds to higher/lower TOC median concentrations.
AAT	0.0006*	0.0015***	0.0008***	Higher AAT corresponds to higher TOC median concentrations.
Median pH	-0.27***	0.17	-0.22**	Increasing pH corresponds to lower/higher TOC median concentrations.
SER:AAT		-0.000003**	N/A	SER and AAT negatively interact, so high values in SER and AAT will decrease TOC medians. Where AAT is high,

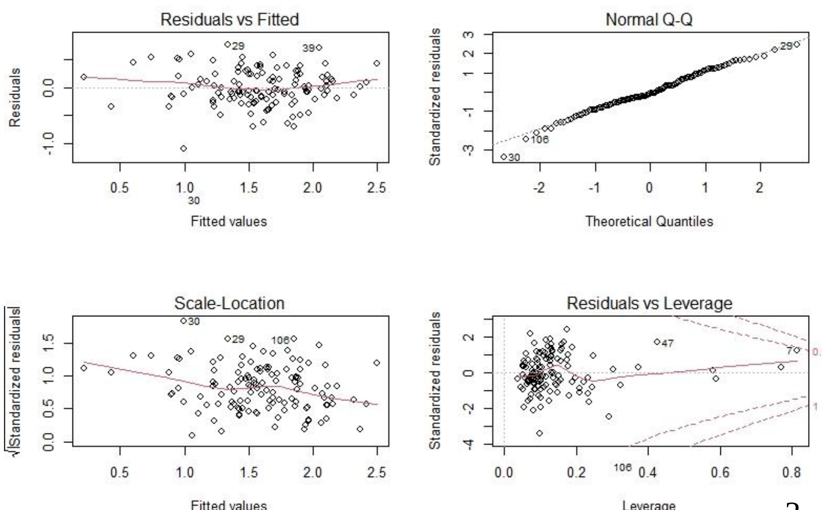
				differences in SER will make a more pronounced difference to TOC medians, and where SER is high, differences in AAT will be more marked in TOC medians.
# sheep:SER		-0.00001***	N/A	SER and number of sheep negatively interact, so high values of SER and sheep number will decrease TOC medians.
SER:Median pH		-0.001***	N/A	SER and median pH negatively interact, so high values of pH and SER will decrease TOC medians, with changes in SER making a higher difference to TOC medians in catchments with higher pH values.
Average topsoil organic carbon:# deer		-0.002*	N/A	Average topsoil organic carbon content and number of deer negatively interact, so increasing deer numbers correspond to lower TOC median concentrations, especially in catchments with higher organic carbon content.
Relief ratio:%age coniferous forest cover		-0.06*	N/A	Relief ratio and percentage of coniferous forest cover negatively interact, so higher coniferous forest cover corresponds to lower TOC median concentrations especially where relief ratio is higher.
Relief ratio:# deer		0.275**	N/A	Relief ratio and number of deer positively interact, so especially where relief ratio is higher, higher deer numbers correspond to increased TOC median concentrations.
Category:coniferous forest cover Rainfall Rainfall+Temperature Temperature Wetup		N/A	 -0.015** 0.006 0.01 -0.005	Compared to the “None” category, coniferous forest cover decreases TOC median concentration in the “Rainfall” and the “Wetup” category.



1



2



3

Figure 5.15: Residual plots for multiple linear regression models with log-transformed median TOC concentration as response variable and catchment characteristics as independent variables (n=127). Numbers correspond with models in Table 5.5. For each model, four plots are presented: in the upper left corner, residuals are plotted against fitted values, to assess if residuals have non-linear pattern that could indicate non-linear relationships, with distinguishable patterns indicating that no-linear relationships are present that the model cannot account for. In the upper right corner, a quantile-quantile plot for residuals allows evaluation of normal distribution of residuals (points should follow the straight line), as error terms must be normally distributed in a regression model. In the lower left corner, a scale-location plot shows if residuals are spread equally along the range of predictors, to check assumption of equal variance (homoscedasticity), with a straight line and randomly spread points indicating homoscedasticity. In the lower right corner, standardized residuals are plotted against leverage to find influential outliers, with points outside the red dashed lines influencing the regression line.

Including interactions terms improves the model in comparison to the one without interactions terms (57% variation explained vs. 47%). While some other variables were included, and some coefficients changed direction, the main interest is in the interaction terms and the direction of their coefficients.

Although SER increases TOC median concentration, the model includes negative interactions between SER and several characteristics. This counterbalances the positive influence and leads to an overall influence of SER reducing TOC medians, especially where these characteristics (higher AAT, more sheep, and higher pH) are present. This could be due to dilution effect and/or reflect drought effects, as drought may lead to a subsequent rise in TOC concentrations for a prolonged period of time (Pärn & Mander, 2012). SER is projected to decrease in most areas of Scotland, with highest reductions projected for the South and East. The model thus highlights these areas as having the highest risk of seeing increases in TOC median concentrations, because the reducing influence of SER diminishes.

The most interesting interaction term in this model in terms of climate change impacts is between SER and AAT with a negative coefficient. Although the interaction has a negative coefficient, meaning that higher values of AAT reduce TOC medians with this interaction, AAT as a single term has a positive coefficient, which means it acts positively on the median TOC concentration overall. Therefore, the model indicates that in cooler areas, SER reduces concentrations, whereas in warmer areas, higher temperatures lead to increases in concentration, with catchments being sensitive to precipitation and drier areas having higher median concentrations. This again suggests drought effects. An increase in AAT is predicted for most Scottish Water catchments, and that would lead to increases in median TOC concentration, especially in combination with decreasing rainfall during the summer.

The interaction of sheep and effective rainfall also appears with a small, but highly significant negative effect. It is hard to infer a causal relationship from this. Using this model for forward prediction means that this interaction makes catchments with higher numbers of sheep more vulnerable to reductions in SER.

The negative interaction between pH median value and SER suggests that pH has different influence depending on wetness conditions. More basic conditions have a reducing influence on TOC median especially where SER is high. This interpretation however does not conform very well with the currently widely acknowledged hypothesis that peatlands release more DOC under deacidification (Monteith et al., 2007). Similarly counterintuitive is the negative interaction between average topsoil organic carbon and number of deer, suggesting that a higher number of deer leads to a reduction in TOC median in catchments with higher average of organic carbon in the topsoil.

In contrast, higher deer numbers correspond to higher TOC medians where relief ratio is higher, which would be easier to interpret as disturbance of soil from trampling could lead to easier transport of TOC to the water source. Relief ratio also negatively interacts with coniferous forest cover, so catchments with steeper reliefs would have lower TOC medians if coniferous forest cover is higher. Generally, coniferous forest can increase concentrations through drainage and disturbance of ground from forest operations and clear felling (Van Dijk & Keenan, 2007). As impacts can depend largely on management and local conditions, it is unsurprising that coniferous forest cover has not been more clearly identified as an influence in the models. The interaction suggests that coniferous forest can however have a beneficial effect on steep reliefs, maybe due to slowing down surface flow.

The only interaction that remained in the model if only interactions were included between the climate sensitivity "category" and the other variables is between category and coniferous forest, with a significant negative effect for the "Rainfall" catchments. These catchments are the slightly steeper river catchments reacting to short term rainfall, so this interaction could be interpreted similarly to the relief ratio – coniferous forest interaction in the other model, as forest cover slowing down surface flow and providing a chance for TOC to be held up in the catchment.

5.3.5 Interpreting the results with regard to implications for climate and land use changes

In the pooled modelling, the category interactions did not help to improve the models. The models derived from separating the groups were partly also difficult to interpret. There are several explanations for this, including data issues, model choice, lack of homogeneousness of the groups, multiple variables describing similar effects, or lack of meaning of the groups for predicting TOC medians.

In terms of data, data used for describing the catchments is mainly free and easily available in Scotland but may not always be a good representation for these predominantly very small catchments due to scales or from averaging. Examples include livestock data which are at parish level, meaning that there may in fact be no livestock present in the catchment even if the parish has a high density, or vice versa; while topsoil organic carbon has been averaged over the catchment, but there could be a large carbon pool that is poorly connected to the surface water, or a small, but very well-connected carbon pool. Results from the statistical modelling are still indicative, but shouldn't be overinterpreted. It needs to be kept in mind that even if expected effects are not showing in the model, this does not rule out that they do not exist, but may simply not be represented in the data. Furthermore, only linear models were used for describing the relationships between catchment characteristics and sensitivity categories/TOC medians. This may mask relationships that are not linear, or threshold effects, although none were immediately apparent from plotting of the data.

Even if some relationships remain unclear from the statistical modelling, the catchments in the groups clearly showed different correlations to climate variables. It is therefore likely that they will react differently to changes in climate and land use. Classification generally imposes an artificial grouping that may represent some members of the groups very well and others very badly. We can see a lot of overlap and rather fuzzy "boundaries" between the groups, so hypotheses formulated for each group with regard to possible implications of climate change may be over- or underestimating the effect on individual catchments within the group. With this in mind, the findings so far are summarised and interpreted for each group, and implications for effects of climate change discussed on this basis.

I. Temperature driven catchments

These catchments have been identified from the analysis above as wet, peaty catchments with mainly semi-natural land cover. TOC concentrations correlate well to 60 days mean air temperatures. From the modelling on TOC medians, it can be seen that baseflow in these catchments seem to dilute TOC concentrations, and higher elevation relief ratios, indicating higher proportions of high elevations and higher erosion, as well as more exposure to weather from South and Southwest aspects increase median TOC concentrations. This suggests DOC is mainly exported in surface flow. Variability in these catchments is also rather low, so it seems likely that these catchments are mainly wet throughout the year, with a regular transfer of DOC from the organic soil horizons to the surface water, and probably some buffering capacity from larger water bodies, making them less sensitive to storm events.

A decrease in precipitation in the summer months may affect these catchments if drought events become more frequent. In case of drought, these water sources could see an increase in variability as DOC accumulates over dry periods and is then flushed out in a storm event, rather than being more steadily exported, unless the water body is capable of buffering this. Studies suggest a decrease of DOC concentrations during droughts is followed by a subsequent increase that may last for several years, although evidence is lacking with regard to long-term effects of drought (Chapman & Palmer, 2016). Frequent drought events may also make peatland more vulnerable to degradation and erosion, with both effects being linked to increases in DOC concentrations (Evans & Warburton, 2010b).

Temperature does not come out as very strongly related to TOC median concentrations, although one would assume a connection especially with AAT as a proxy for period of microbial activity and primary production, and hence DOC production. Wetness however is an inhibitory factor for DOC production, as soil conditions remain mainly anaerobic (Clark et al., 2009). A decrease in precipitation would mean that increases in temperature became a factor for increasing DOC production, once these areas fall below a threshold value for wetness.

Finally, changes in climate may influence vegetation in these catchments, leading to shifting concentrations, as there are indications that peatlands dominated by *Calluna* (in drier areas) leach higher amounts of DOC than those dominated by *Sphagnum* (Ritson et al., 2014).

These interpretations suggest that climate change may shift these catchments to become more sensitive to discharge, and they may see an increasing long-term trend.

II. Rainfall and temperature driven catchment

These catchments are similar in their characteristics to the “Temperature” catchments, but have steeper reliefs and consist mainly of rivers rather than lochs and reservoirs. TOC concentrations are related to 60 days mean air temperature, but also to 3-day antecedent rainfall periods. They seem to behave similar to the “Temperature” catchments, but steeper reliefs could lead to a quicker runoff of precipitation over the surface and subsurface, washing out DOC produced in the topsoil. As rivers, the water sources have lower buffering capacity, meaning that these events can more easily show in spikes of concentration. Assuming that drought events would act as described above, leading to DOC spikes from accumulation during dry periods and flushing out during storm events, it is likely that these water bodies have less buffering capability than those in the “Temperature” group and also see an increase in variability and peak concentrations.

The model for median TOC found a potential influence of agricultural activity, so changes in climate that would make these areas more capable of agricultural land use, might also induce increases in TOC concentrations.

III. Rainfall driven catchments

These catchments have smaller carbon pools and steeper reliefs, and TOC concentrations show correlations to 3-day antecedent rainfall periods. They tend to have lower TOC median concentrations but show the biggest variability in concentrations. The group consists mainly of river catchments. The regression model for this group was harder to interpret with some counterintuitive coefficient directions. It is possible that negative correlations to relief ratio, average number of sheep and SER, together with a positive correlation to semi-natural land cover, indicate that this group is more diverse in the size of the carbon pool, with wetter, more natural, less intensively used catchments producing more DOC due to higher carbon content in the soils.

IV. Wetup effect catchments

These catchments show a correlation to a 60-day antecedent rainfall period. They are the most variable group and include many of the catchments with atypical colour concentration curves. This could be due to catchment processes but also to in-lake/reservoir processes, making it very difficult to attribute catchment characteristics to TOC concentrations and hence estimate likely effects of climate change. The basic assumption is that these are catchments with drier soil conditions where a higher amount of rainfall is needed to start the transfer of DOC from the topsoil into the surface waters. Higher AAT have been linked to higher TOC concentrations in this group. This indicates that in these catchments, wetness is not an inhibiting factor for DOC production and the length of the growing season would influence the amount of DOC produced in the soil. With increasing temperatures, more DOC is likely to be produced, however with decreases in summer precipitation, it may be less likely to be washed out into the surface water. Accumulation over the summer months could then lead to more DOC being released to the surface water in autumn and winter when catchments get wetter. This would be worth analysing further especially for those catchments that already peak in winter.

Alternatively, or additionally, in-lake processes could influence concentrations either by degradation of DOC, or production from algae. There may also be a pronounced buffering effect, meaning that there could be high concentrations coming from the catchment as a reaction to short term rainfall events that only build up gradually.

These catchments have also been associated with more intensive agricultural use in the redundancy analysis. The regression model for this group shows a positive correlation of percentage of improved grassland and average number of sheep with median TOC concentrations. This means that this group could contain catchments where a significant proportion of DOC comes from farmland, with DOC exported from baseflow (Evans, Freeman et al., 2006). In these catchments, land management is probably a more important factor for risk of decreasing water quality than climate directly.

It can also be assumed that catchments with degraded peatland would be found in this group, which could be exacerbated through changes in climate. Drought conditions would make the soil more sensitive to mechanical influences (such as from grazing) and lead to higher erosion risks in subsequent storm events.

V. No correlation catchments

Catchments in this category seem to be relatively dry, flat catchments with smaller organic carbon pools, although they show comparatively high TOC median concentrations. The PCA visualises that there is considerable overlap between this groups and the “Wetup” group. While these catchments didn’t show correlation to the climate indicators, it is possible that there still exists a relationship, but it wasn’t observable, especially if the period of antecedent rainfall doesn’t match the conditions of the catchments (a longer or shorter period could be more suitable). The concentrations also seem to be relatively stable for most catchments within this group. The modelling approach didn’t lead to a satisfactory model for TOC median concentrations. This could be due to diversity within this group.

VI. Land use implications

The modelling on TOC medians with interaction terms allow some very general conclusions for management in different types of catchments. While grazing has not been linked to increases in TOC concentrations, modelling has linked the average number of sheep in the catchment to increases in TOC concentrations, especially in drier areas. This could indicate that decreases in precipitation could make catchments more sensitive to grazing, maybe due to mechanical disturbance. The modelling also indicated that trees could reduce TOC concentrations on steeper reliefs, maybe due to slowing down surface and subsurface runoff.

5.3.6 Conclusions for the risk screening

The analysis of the different groups leads to a first loose classification of the catchments in terms of TOC production that can be used for some conjectures about impacts of climate change:

An increase in AAT increases the possibility for more DOC production. While the catchments in the “Temperature” group may not react to this due to high wetness, it is possible that this may change with decreasing SER. It is therefore assumed that increases in annual accumulate temperature poses a risk to this group.

A decrease in SER has been discussed in all the groups to likely lead to a build-up of DOC in the soil that gets flushed out in subsequent storm events. This could translate to spikes in concentration in those water bodies with lower buffering capacity, in the “Rainfall” and “Rainfall+Temperature” groups. While variability has not been analysed directly, it is assumed that these catchments are at risk of seeing increases in DOC concentration variability and peaks.

While it is difficult to generally project impacts from increases in temperature and decreases in precipitation for the “Wetup” catchments, it is assumed that the main factor is precipitation, with decreases likely leading to conditions that make the catchments more sensitive to disturbances, and thus more leaching of DOC. Although there are little conclusions for the “None” catchments, they are treated as the “Wetup” catchments due to observed overlap between the groups.

Using these conjectures, a first risk screening (Figure 5.16) can be carried out by looking at projected changes in AAT, influencing overall (median) TOC concentrations for the “Temperature” catchments (increasing where AAT increases), and in SER, indicating risk for overall TOC concentrations for the “Temperature”, “Rainfall+Temperature”, “Wetup” and “None” groups (increasing where SEP decreases), and finally of risk for TOC concentration variability for the “Rainfall” and “Rainfall+Temperature” groups (increasing where SER decreases).

These risk maps especially highlight the catchments in the South and Southeast of Scotland as having higher risk of seeing increases in TOC concentrations from decreases in SER. Of these catchments, the “Rainfall” and “Rainfall+Temperature” catchments are also highlighted for risk of increases in variability of concentrations. Additionally, some catchments on the West coast are identified as being of higher risk from increases in AAT. These results are interesting as increases have mainly been seen so far in the Western catchments (personal communication, Scottish Water). The risk maps highlight that these catchments may be less at risk from climate change, but those that have so far not been identified as problematic over rising DOC trends might become so in future.

Alternatively, the best multiple linear regression model for all catchments together can be used to derive estimates for TOC median values, using AAT and SER projections for the period 2041-2060 (Table C.6), and a risk map produced based on changes between current and projected TOC median (Figure 5.17).

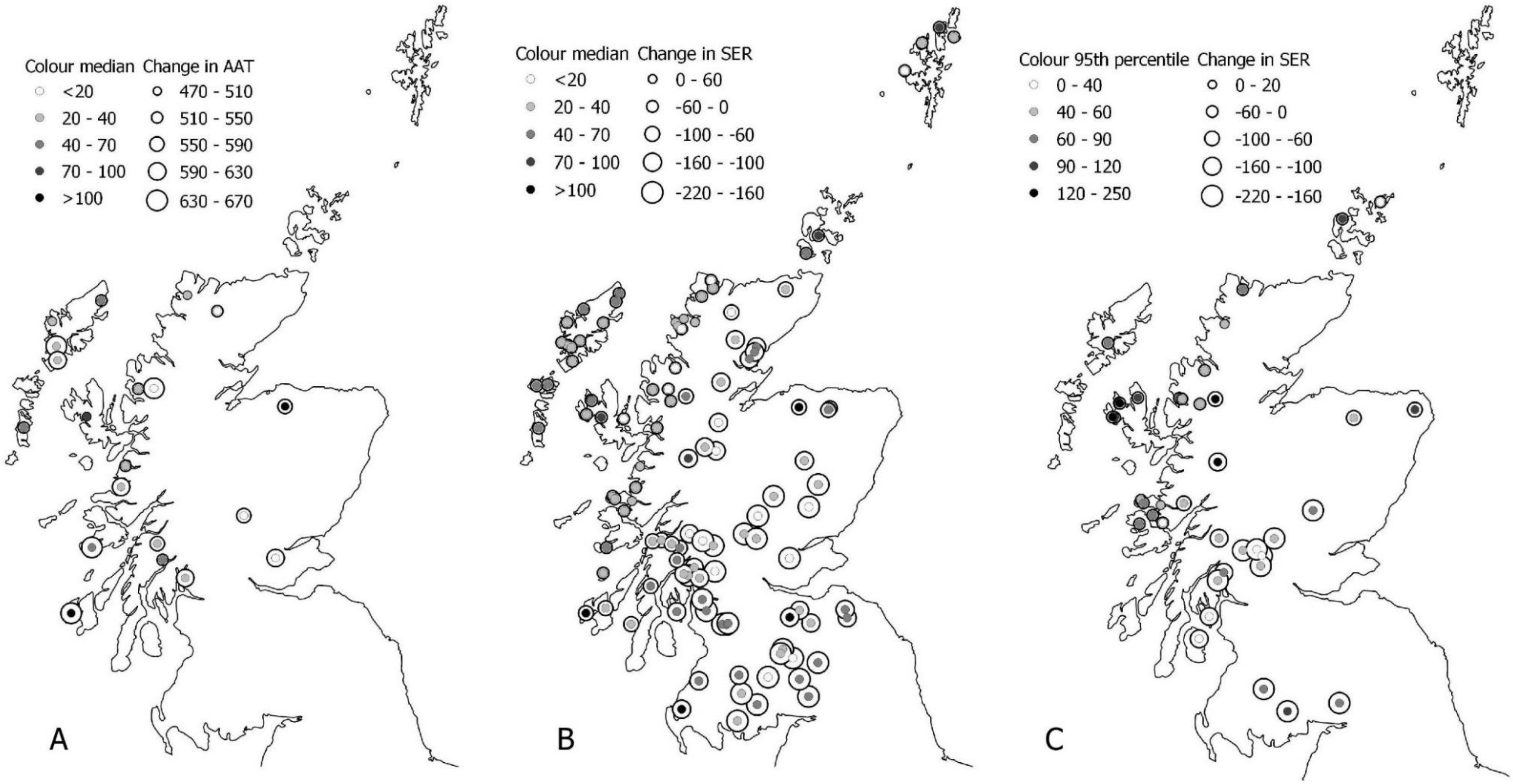


Figure 5.16: TOC risk maps, based on projected changes in climate indicators combined with identified vulnerabilities of catchments to climate pressures. Point size indicates magnitude of projected (UKCP18) change of the identified hazard between the baseline period 1981-2000 and the projection period 2041-2060 – for AAT, the bigger the circle the higher the projected increase; for SER, the bigger the circle the higher the projected decrease. Shade indicates current (2011-2016) colour concentration median values (when assessing risk to overall concentrations) or 95th percentile concentrations (when assessing risk to variability). A. Catchments where increasing overall TOC concentrations are hypothesised due to increases in AAT ("Temperature" category), B. catchments where increasing overall TOC concentrations are hypothesised due to reduction in SER ("Temperature", "Rainfall+Temperature", "Wetup", and "None" categories), C. catchments where increases in variability of concentrations are hypothesised due to reduction in SER, with consequences especially for increasing peak concentrations ("Rainfall" and "Rainfall+Temperature" categories).

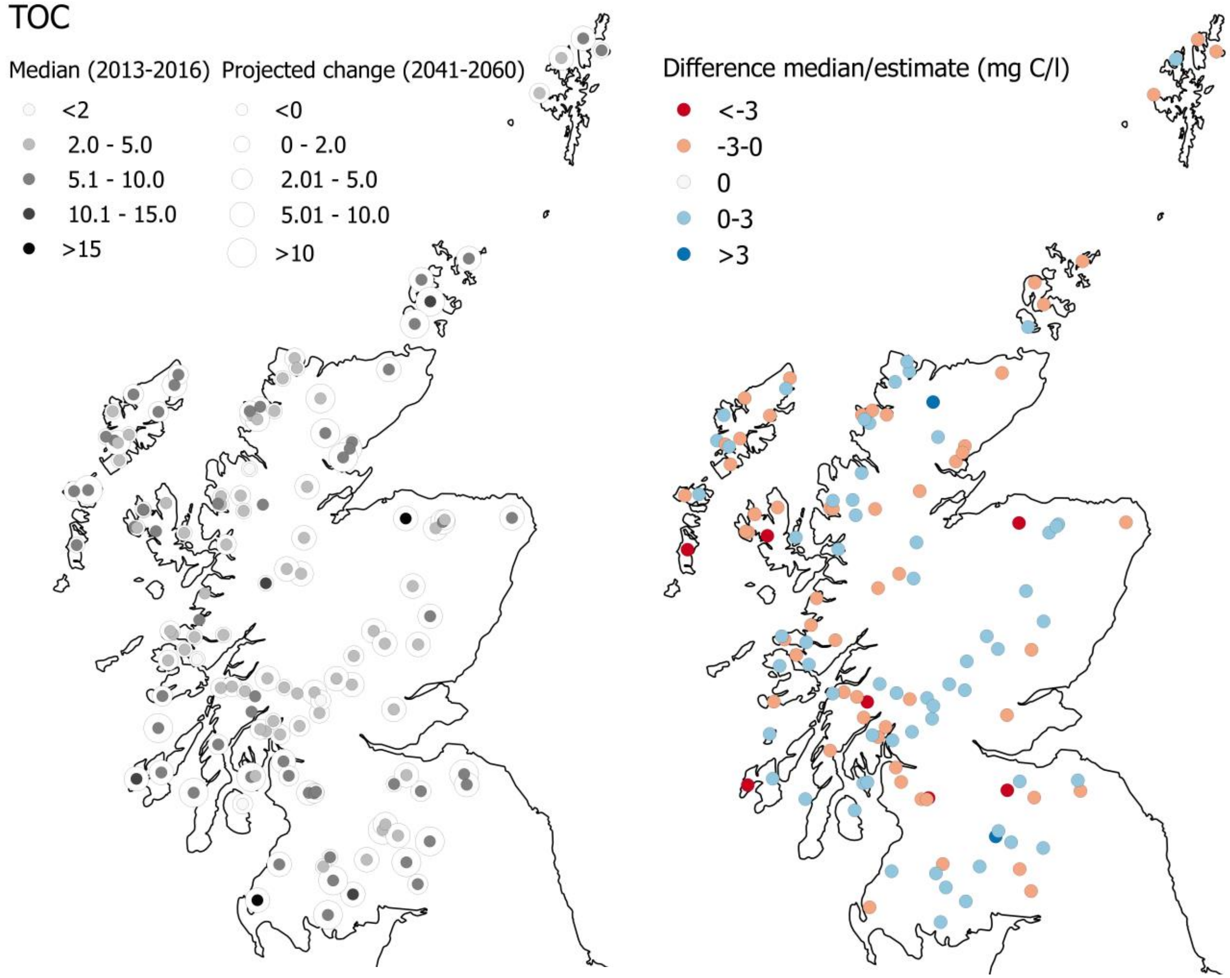


Figure 5.17: A. TOC risk map based on projections of the selected multiple linear regression model, using AAT and SER projections (UKCP18) for each catchment for 2041-2060. Point size reflects projected increase in TOC median concentration. Shade reflects current (2013-2016) TOC median concentration. B. Model performance – indicates if the model is over- or underestimating the median by reflecting the difference between actual TOC median concentration and estimated TOC median concentration of the model.

This risk map similarly predominantly highlights catchments in the South and Southeast of Scotland, with some additions in the North, especially for Orkney. The model is mainly driven by changes in SER, especially where SER becomes very small as the negative influence of the interaction terms reduces drastically. While the model strongly overestimates medians for some catchments and is therefore more likely to also overestimate future median, there is no spatial pattern in over- or underestimation observable. There is also no observable relation between degree of estimated change and over- or underprediction from the model, indicating that the map indeed reflects change rather than model error.

Having identified risk levels according to projected change, these can be combined with whatever metric is deemed appropriate to assess risk for individual catchments and the overall national supply system. For example, risk might not be deemed highest where concentrations are currently low, even if increases are projected, or vice versa, even if projected changes are small, risk might be high if the treatment work is already challenged by current concentrations. Risk might also be perceived high where sources serve a high number of people, or where there is no alternative or emergency supply. These considerations can then serve to arrive at a risk score, to inform and prioritise subsequent risk assessments, research, or investment.

5.4 Limitations and further approach

There are some advantage of the risk map based on model outcomes over the risk maps based on changes in identified drivers: interactions between pressures are included, which would be especially important for the interaction between SER and AAT; other catchment characteristics which may influence TOC median concentrations are included, and could also be changed according to projections if desired, to see what this would mean; and one map may be easier to interpret than several. However, statistical models are limited in their capability for forward projection, as they describe correlations rather than causal relationships. As discussed at several points above, it is possible to interpret the models based on our understanding of processes and infer causal relationships, but in many cases it was unclear if the variables and their coefficients represented more complex processes that were not easily described by the variables, either because important factors were missing from the model, or because the data weren't of high enough resolution and quality to adequately represent the effect, or the group was so diverse that different effects overlapped. Furthermore, these correlations, even if we can with some confidence attribute them to causes, may not necessarily be continued. The analysis builds on spatial differences. The "space for time" substitution assumes that different locations can represent different points in time, and looking at the different conditions over space allows to conclude how areas will look like when conditions change, so project spatial differences into temporal ones (Huang et al., 2019). For example, we could assume that some areas in Scotland in the "squeezed middle" between highland and lowland areas will increasingly exhibit properties associated with lowlands. However, changes in climate advance at an unprecedented rate and local conditions may not change at the same rate, upending the relationship. So, while using the multiple linear regression model to project future TOC median values gives very useful insight into possible effects of climate change and allows to identify potential "high risk" catchments, it needs a more detailed and individual follow-up risk assessment to unpick potential impacts more adequately. It would be worthwhile to investigate this further for selected catchments to confirm the formulated hypotheses, and to gain a better understanding of the processes, to estimate the extend of the increase and if there are thresholds for TOC/DOC release, or which other factors influence the processes. Catchments that have been identified as high risk in this analysis can be prioritised for further analysis (see Appendix D for suggestions). Currently, limited understanding of the system challenges evaluation of different scenarios, either with regard to changing climate or management options, but a model that allows to identify important parameters could help to get projections for water quality under climate and land management projections. The focus of further investigations would therefore lie on increasing the understanding of the underlying processes. Using and calibrating process-based models could be an effective way to test hypotheses around the processes and gain an understanding of these (Beven, 2012).

DOC data are available for only a few catchments, which is why the risk screening was based on TOC data. For further investigation, a catchment could be chosen where DOC data was available. The highest frequency data was weekly data, which could be coupled with rainfall and temperature data from the Met Office. Flow data are not routinely available, but flow measurements from further downstream might help to calibrate a rainfall runoff model. In terms of catchment data, finer scale data e.g., on soils may be available. A model could help to point out which kind of data would most help to improve process understanding. Most useful for decision-makers would be a projection of median, or average, and peak concentrations over time with confidence margins, which would enable an evaluation of best and worst case together with their related conditions. A coupling of magnitude of increases, or thresholds for DOC release, to specific catchment conditions could also help to transfer findings from one catchment to the other to further prove the identification of high-risk catchments.

Contrary to the risk screening map based on model projections, the risk map based on changes in identified drivers incorporated the hypotheses that were formulated resulting from the analysis regarding climate change impacts. It therefore addresses the discussed issue that current spatial differences may not well represent future temporal trends, but uses the characterisation of the different groups, together with current knowledge on processes, to conjecture potential impacts depending on sensitivities. Thus, the classification provides a starting point to understand differences in catchment responses to pressures and climate drivers, and provides a framework for viewing, understanding, and interpreting the results of more detailed, individual risk assessments in a broader context. This is crucial for risk assessing the whole supply system and making more strategic, pro-active decisions regardless of uncertainties about the future.

The developed risk screening maps allow identifying catchments for further in-depth analysis. The hypotheses formulated here are a starting point for further research that can be designed to verify, or disprove, assumptions made. There are gaps in our knowledge and understanding of the actual processes of DOC production and transfer in the catchments, which limits understanding of how these systems will respond to climate change, so it seems pertinent that data are collected that helps address these questions. These data can feed into more elaborate models. With increasing knowledge of the catchments systems, the classification can be further developed helping to put it into the broader context of the national supply system. A clearer picture of the groups and the characteristics of the catchments within these will also allow to risk screen catchments that currently have inadequate or no TOC data, and to allocate an initial risk group.

6. Risk screening: Estimating change in risk potential for *E. coli* contamination in raw water

Chapter 5 carried out Assessment 1.2 regarding raw water colour, this chapter now examines these aspects with regard to another common problem in Scottish Water catchments, contamination with *E. coli*. This chapter gives the relevant background to *E. coli* contamination (6.1), describes the methods for the analysis (6.2), results (6.3), and reviews limitations and further steps (6.4).

The ingestion of pathogens from animal and human faeces, such as some forms of *Escherichia coli*, streptococci, or *Cryptosporidium* spp., can lead to serious gastro-intestinal illness with symptoms including diarrhoea, fever, stomach pain, and vomiting, and can be especially problematic in children and elderly people (Hunter, 2003; Duhaime & Roberts, 2018). Infection can occur through direct contact with animal faeces, contaminated food (meat, dairy products, or vegetables), recreational or drinking water, or through human-to-human transmission (Rotariu et al., 2012). While transmission through food items seems to be the most important factor (Pennington, 2014), there are outbreaks linked to untreated or insufficiently treated drinking water (e.g., Licence, 2001; Saxena et al., 2015).

E. coli has high importance as an indicator organism for faecal contamination in water due to its prevalence in animal and human faeces and the development of quick and cost-effective tests (Odonkor & Ampofo, 2013). The statutory requirement for Scotland is that no *E. coli* is present in drinking water (The Water Supply (Water Quality) (Scotland) Regulations 2001). This is usually achieved via disinfection within the treatment process.

Managing water quality with regard to faecal pathogens can be challenging due to the possible sporadic nature of contamination events, depending on a number of factors such as climatic conditions (e.g., temperatures favourable to long survival, episodic rainfall events that induce flush-outs), management changes, or points of failure (e.g., direct access of livestock to surface water, sewer overflows). It is important to maintain a good awareness of risk factors and areas as well as efficient and reliable mitigation measures to reduce dependency on chemical end-point treatment. While there is usually awareness within water utilities of agricultural activity as a main point of concern, and an existing prioritisation of catchments for mitigation measures, a potential shift in risk factors needs to be considered in order to develop pro-active mitigation strategies for catchments that may have an increasing risk of contamination events.

6.1 Sources of *E. coli*

Main sources of *E. coli* are sheep and cattle, but they also come from other species such as deer (Rotariu et al., 2012). *Cryptosporidium* is mainly associated with grazing cattle, and their defecation close to surface water as a potential source of contamination (Duhaime & Roberts, 2018). Next to faeces from grazing animals, indirect application to land through slurry and manures is seen as a second major source (Vinten et al., 2004). Further input can be derived from farm steadings and human populations through sewer overflows or badly maintained septic tanks (Vinten et al., 2004; Tetzlaff et al., 2012).

While the major causes of faecal contamination in surface water and their relative importance have been well studied on field- or farm scale, it is more challenging to determine relationships on a catchment scale. There are a number of potential reasons for this, including the influence of farming practices, seasonal variations, climatic conditions influencing biological productivity and hydrological connectivity within a catchment (Tetzlaff et al., 2012), and the potential of pathogens to multiply and die-off depending on a number of biotic and abiotic factors (Oliver et al., 2016). At the same time, the ability to predict water quality outcomes depending on land management options is important at different scales from local decision making up to national policy formulation. Among the reasons for achieving a prediction of faecal indicator organism behaviour is the enabling of a screening to guide regulators in prioritising decisions and predict future scenarios due to climate and land-use changes (Oliver et al., 2016).

6.2 Methods

The risk screening aims to identify potential catchments where changes in climate and land use may lead to an increase in risk of faecal contamination of surface water, so that these catchments can be under increased attention to enable pro-active intervention. As a first step, the relationship between catchment characteristics and *E. coli* concentrations (as an indicator for contamination with faecal pathogens) is investigated for this set of catchments in order to highlight the overall most important risk factors. Then, it is considered how these risk factors may change within a future period, so that catchments can be identified for further investigation.

The analysis is making the following assumptions to guide the modelling:

- Catchments show very skewed concentration distributions with mostly very low concentrations and occasional high peaks due to contamination events.
- Livestock is widely seen as a predominant source of *E. coli*; therefore, it is expected that livestock numbers and improved grassland cover will relate to *E. coli* concentrations. Other sources of *E. coli* contamination could be slurry applications on arable areas, septic tanks, sewer overflows in urban areas, or contamination through wildlife.
- Relationships could be obscured by land management practices, such as buffer strips, fencing, overwintering habits, or livestock density; the location of the contamination source; and its hydrological connectivity to the water source.
- Some catchments may show a relationship to rainfall amounts, where source areas are connected in higher rainfall events, or where rainfall leads to a diluting effect.

6.2.1 Relationships between catchment characteristics and median *E. coli* concentrations

The complete set of 154 catchments (see 3.3.3) was used in the analysis, with sample frequencies varying per catchment from once per month or once every three months (leading to sample sizes spanning 20 to 70 per catchment). Median concentrations were chosen as a baseline value to represent an "average" contamination level and thus reflect baseline pressures in the catchment. Higher medians are expected to reflect a general higher concentration level and/or more frequent peaks. Although peak concentrations are, from a drinking water perspective, more worrying, they are less likely to show a relationship to pressures within the catchment (Neill et al., 2018) as they often relate to single contamination events.

Multiple linear regression was performed (B.9) with the log-transformed *E. coli* medians as a response variable, and catchment characteristics as independent variables: the type of source, area in km^2, elevation relief ratio, relief ratio, percentage of Southwest facing aspect, average amount of topsoil organic carbon, average baseflow index, percentages of arable, improved grassland, coniferous forest, urban and semi-natural land cover, numbers of deer, cattle and sheep densities, number of septic tanks, median pH, SER and AAT. Backwards stepwise regression based on BIC was used to find significant variables. This was performed with just the variables, but also including interactions of all independent variables with the amount of effective rainfall, AAT, and selected additional interactions (between relief ratio and numbers of cattle, sheep, deer, and aspect; between average topsoil organic carbon and numbers of cattle, sheep, deer, and improved grassland cover; between average baseflow index and improved grassland cover; and between coniferous forest cover and number of deer, and relief ratio). Interaction terms were included in batches to avoid running out of degrees of freedom. Significant variables and interactions from the batch-wise models were then combined in a last step to arrive at the final model. Plots of residuals, adjusted R^2, and normalised RSME from a ten times repeated 10-fold cross validation were used to evaluate the goodness of fit and potential overfitting issues.

6.2.2 Sensitivity to rainfall

The relationship between concentrations and different amounts of rainfall (daily total rainfall, and total rainfall of the preceding 3, 10, and 30 days) was tested per catchment and per season (March to October – summer; November to February – winter) using Spearman's rank correlation tests (B.10). Rainfall data were obtained from the UK Met Office. This resulted in 8 Spearman's rho values and corresponding p-values per catchment. A relationship was judged to be moderate if Spearman's rho >0.4 and as strong if Spearman's rho > 0.6 if p>0.05. The catchments were then manually sorted into "None" (if no or only one moderate correlation was found), "Short" (if at least one strong or two moderate correlations to a shorter rainfall period was found, with at least one in the summer season), "Medium" (if at least one strong or two moderate correlations to a medium rainfall period was found, with at least one in the summer season), or "Long" (if at least one strong or two moderate correlations to a long rainfall period was found. with at least one in the summer season).

These models as described above were rerun using interaction terms with these categories to see if they added to explaining variation.

6.2.3 Estimating land use change

While land use can serve to estimate current risk, in order estimate how risk may change, it is necessary to get an estimate of future land use. Predicting land use changes and land use patterns for a future period is extremely difficult due to numerous factors and complex relationships that lead to land use decisions. Physical properties that make specific land uses possible can be described by using land capability assessments (Bibby et al., 1991, Brown et al., 2008). Climate change, through higher temperatures and different rainfall amounts, will influence how land can be used, for example by the capability of soils to produce crops. It is envisaged that land capability will change in Scotland with some areas becoming more capable of crop production, and some areas more capable of grass production sustaining livestock (Brown et al., 2008). Where this happens, these catchments could see an intensification of agricultural practices, potentially putting them at higher risk for degrading water quality. Changes in land capability could thus be used as a first approximation to potential future land uses and projecting changes in land capability identifies where catchment have an increased risk through climate change and related land use changes.

Current and projected land capability data were provided, derived with the methods described in Brown et al. (2008). Future land capability projections were derived using projections in soil moisture deficit (inverse of SER) and AAT. The same procedure outlined in 3.4 was used for future UKCP18 climate data to develop a future bioclimate and land capability ‘reasonable worst case’ scenario, assuming that irrigation and drainage infrastructure continues to be available so that land use is optimised according to the capability class for that location.

6.3 Results and discussion

E. coli concentrations show skewed distributions in almost all catchments (Figure 6.1 & Figure 6.2), with the vast majority of catchment showing no or very little presence of *E.*

coli with occasional contamination events. Some catchments show *E. coli* presence almost all year round with strong spikes. Very few catchments have generally elevated *E. coli* levels. There is no indication that catchments generally show seasonality for *E. coli* presence.

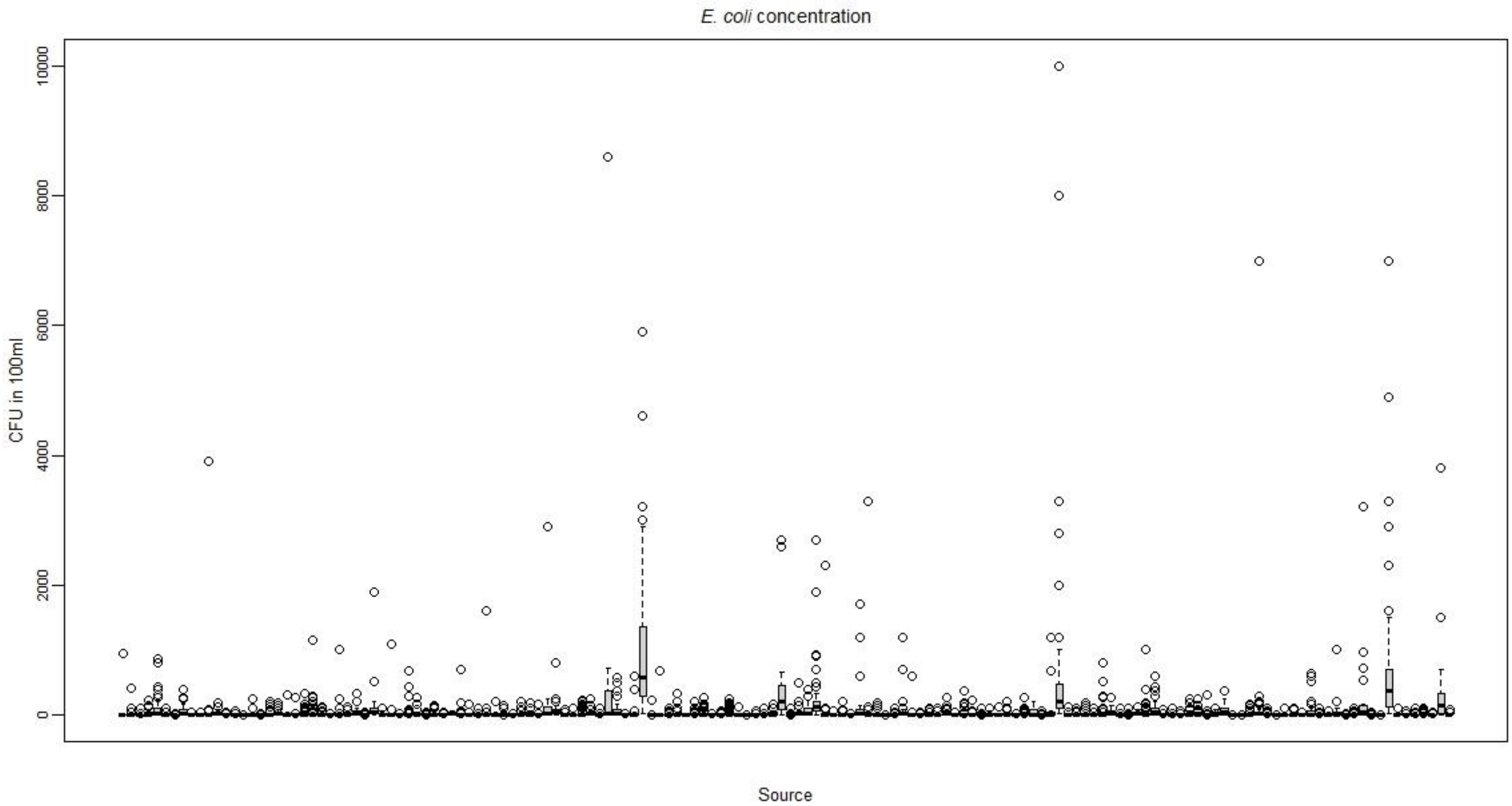


Figure 6.1: Boxplots showing the distribution of *E. coli* concentrations per water abstraction source. Due to a small number of exceptionally high concentrations, this plot only allows to distinguish that a few catchments have generally elevated *E. coli* concentrations and extremely high peak concentrations.

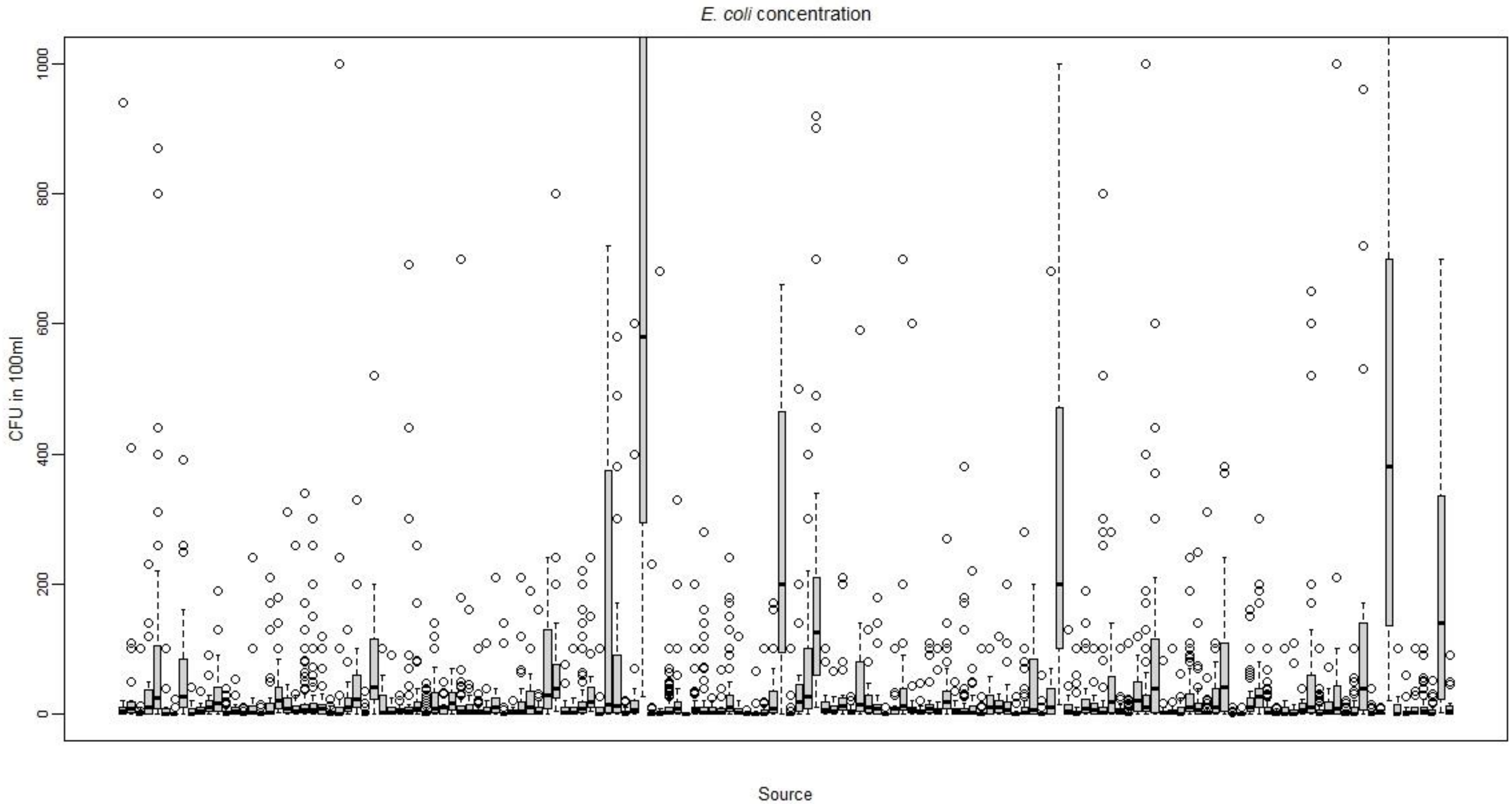


Figure 6.2: Boxplots showing the distribution of *E. coli* concentrations per water abstraction source, including all data but y-axis capped at 1000 CFU per 100ml for better comparison. It is distinguishable that many catchments normally have *E. coli* concentrations of close to 0, some have usually slightly elevated concentrations and a few have generally comparatively high concentrations of *E. coli*. Almost all catchments have high outliers and most show a right skewed distribution.

6.3.1 Regression models

The model was run without interactions (1A) and including interactions (1B) for comparison. The residual plots of these models (Figure 6.3) did not indicate major problems with the residuals and hence the models, but they show that it is likely that the few catchments with very high medians will influence the model to a high degree. The modelling was therefore repeated with these catchments removed (leaving catchments with *E. coli* median concentrations below 100 – 2A and 2B).

As the future projection of *E. coli* contamination risk has to be based on land capability, it is of interest to see how models perform that are based on land capability rather than actual land use. The model was therefore rerun on the complete data set substituting land use percentages and livestock numbers with percentage of land capability classes 1, 2, and 3 (suitable for arable agriculture) and percentage of land capability classes 4-5 (suitable for higher intensity livestock). Variables describing climate and soil characteristics were not included as these are used in the calculation of land capability. The model was therefore only run without interactions (3).

The various regression models are summarised in Table 6.1. Model 1A, without interactions terms, includes as independent variables the type of source (with median concentrations in rivers being higher), a negative correlation with relief ratio, which could be due to less agricultural activity in steeper areas), and positive correlations to all land cover variables (except urban area cover which was excluded from the model). These could indicate contamination from wildlife and from livestock. Model 1B, with interaction terms, retains some more variables, such as a positive correlation with area, indicating larger catchments as having higher *E. coli* medians. A negative correlation with average organic carbon content in the topsoil could be explained by a reflection of more natural, peaty catchments with less livestock, and a positive correlation with AAT could reflect a geographical pattern with catchments in the East being more intensively used, as well as an effect of temperature on microbial growth and reproduction. Percentage of urban area is included in this model with a positive correlation, unsurprising as urban areas offer contamination possibilities through e.g., sewer overflows (Barbosa et al., 2012). Number of septic tanks has a negative correlation, which is counterintuitive as septic tanks are a point source of pollution, however there are only few catchments with septic tanks in this dataset.

The average number of cattle is also, surprisingly, negatively correlated. Median pH and SER were positively correlated. Higher pH could indicate more agricultural activity, especially as grassland sustaining higher amounts of livestock tend to have higher soil pH. Increases in precipitation could be due to better connectivity of pollution sources to the water body (Tetzlaff et al., 2012).

Retained interactions were between SER and number of cattle and median pH, with the former being positively and the latter being negatively correlated. This means in catchments with a higher number of cattle, larger amounts on rainfall lead to higher *E. coli* medians, possibly due to a better connection of source to receptor. In catchments with a a more basic pH, rainfall has a more pronounced decreasing effect, which seems to contradict the first interaction, as higher numbers of cattle would be expected in areas with higher soil pH. Interactions were also retained for AAT and area with a negative correlation, and for AAT and septic tanks with a positive correlation, meaning higher AAT corresponds to lower *E. coli* medians especially in larger catchments, and to higher *E. coli* medians especially where there are septic tanks.

With the reduced dataset, without interactions, the model is very simple but only explains a low amount of variation (23%). The only retained variables are type of source, relief ratio, and percentage of improved grassland, with the coefficients having the same directions as in models 1A and 1B. In comparison to the model for the bigger dataset, percentage of semi-natural land cover, coniferous forest, and arable areas are not included. For arable areas, this is probably due to there being only a few catchments with only little arable area cover left in this dataset.

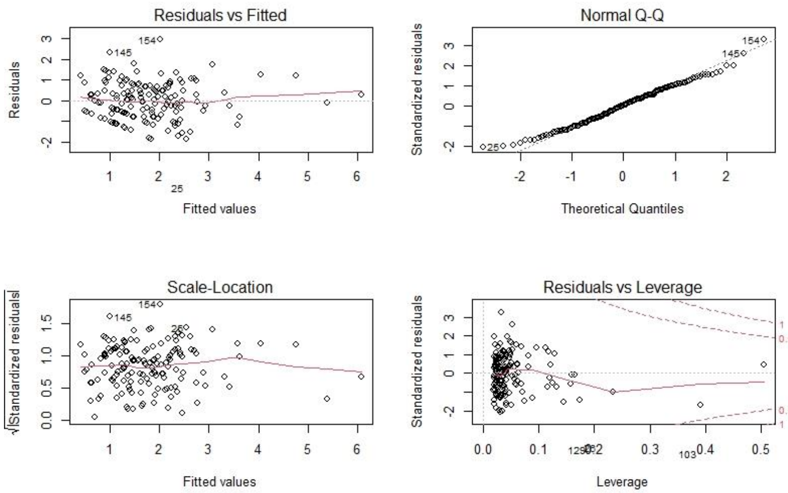
When including interactions, the model improves in terms of variability explained (31%), but there are still only a few variables left: type of source, relief ratio, average number of cattle, median pH, SER and the interactions of SER and average number of cattle and pH median, with the same coefficient directions as in model 1B.

Land capability does not work as well as actual land use in the models (with only 36% of variability explained). However, it results in a very simple model that uses few variables, with both categories of land capabilities included. Considering these only few explanatory variables, the variation explained is still quite high. The use of this model for an initial projection of *E. coli* concentrations for a future period for the purpose of a risk screening seems justified.

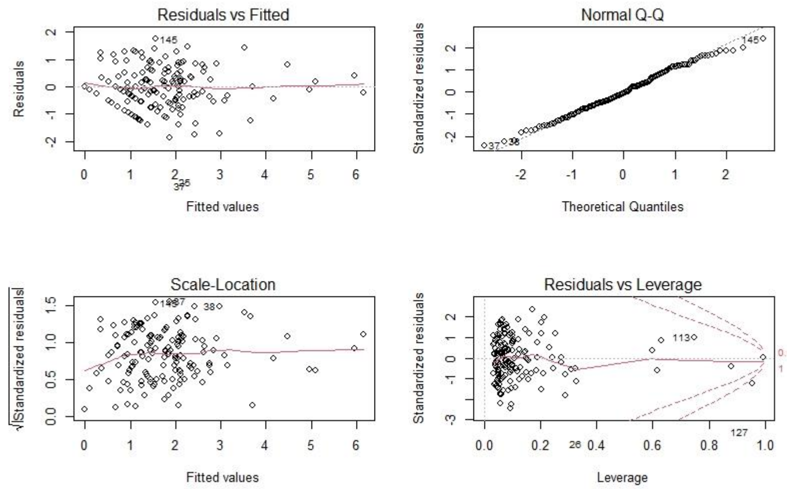
Table 6.1: Multiple linear regression models with log-transformed median *E. coli* concentration as response variable (n=154; *** p<0.001, ** p<0.01, * p<0.05; - = was eliminated during the modelling process, N/A = not included in the modelling process).

Variable	**Coefficients and significance**					**Comment**
	1A. Without interaction terms	1B. With interaction terms	2A. Without interaction terms, reduced dataset	2B. With interaction terms, reduced dataset	3. Without interaction terms, with land capability	
Variability explained	46%	58%	23%	31%	36%	
RSME/(max-min)	0.15	0.16	0.23	0.23	0.16	
Source: Loch River	-0.26 0.83***	-0.11 0.87***	-0.23 0.93***	-0.07 0.97***	-0.26 0.84***	This suggests that median concentrations in rivers tend to be higher than in reservoirs (lochs not significant).
Area	-	0.007**	-	-	-	*E. coli* median concentrations are higher in larger catchments.
Relief ratio	-6.37***	-4.77**	-3.76***	-4.94**	-4.17*	Steeper reliefs reduce *E. coli* concentrations. This could be due to less agricultural activity and less livestock kept on these lands.
Average topsoil organic carbon content	-	-0.02*	-	-	N/A	*E. coli* median concentrations are lower in catchments with more organic carbon in the soil.
AAT	-	0.0008	-	-	N/A	*E. coli* median concentrations are higher where AAT is higher.
% Semi-natural land cover	0.04***	0.04***	-	-	N/A	Positive correlations might be due to *E. coli* contamination through wildlife.
% Coniferous forest cover	0.03**	0.04***	-	-	N/A	Positive correlations might be due to *E. coli* contamination through wildlife.
% Arable area/ % area of prime land	0.08***	0.07**	-	-	0.075***	Arable areas can contribute to faecal pathogen presence through slurry application, or catchments with higher arable area cover could be prone to also have higher population densities.
% Improved grassland/ % area with land capability classes 3-	0.08***	0.06***	0.03***	-	0.01*	Livestock on grassland is a major source of faecal pathogen contamination.

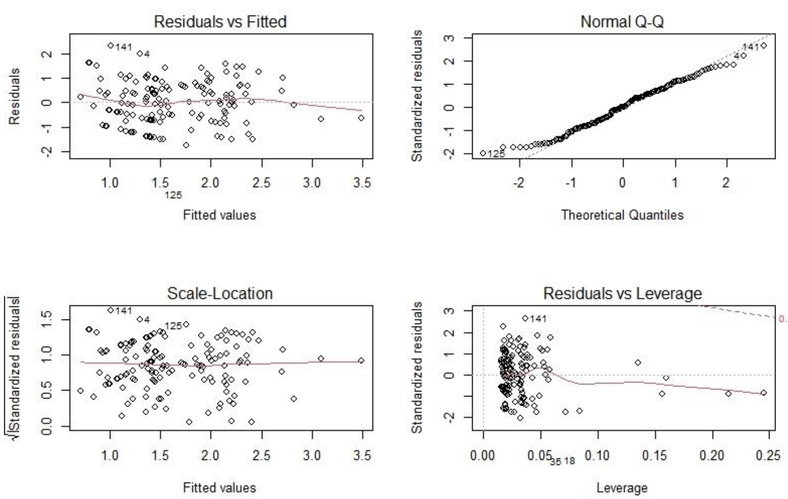
5						
% Urban area	-	0.89*	-	-	N/A	There could be incidents of e.g., sewer overflows in urban areas.
Number of septic tanks	-	-0.84*	-	-		Septic tanks are a point source of pollution, making a negative correlation counterintuitive.
Average no. of cattle	-	-0.009	-	-0.006	N/A	Cattle is a source of *E. coli* contamination, making a negative correlation counterintuitive.
Catchment pH Median value	-	0.8*	-	1.51***	N/A	Catchments with higher pH values tend to have higher *E. coli* concentrations. This could be an indicator for agricultural activity, especially grassland and the number of livestock.
SER	-	0.02**	-	0.03***	N/A	Higher rainfall amounts could keep sources connected to the water body.
Interaction SER and: Average no. of cattle Catchment pH median value	N/A	 0.00009*** -0.003**	N/A	 0.00007** -0.004**	N/A	In catchments with a higher number of cattle, more rainfall would correspond to higher *E. coli* concentrations. This could be due to an increased transport from the field to the water source. In catchments with lower median pH, higher rainfall corresponds to an increased *E. coli* median concentration.
Interaction AAT and: Area Septic tanks	N/A	 -0.000006** 0.0009*	N/A	-	N/A	In larger catchments, higher AAT corresponds to lower *E. coli* median concentrations. In catchments with more septic tanks, higher AAT corresponds to higher *E. coli* median concentrations.


Residuals vs Fitted
Residuals
Fitted values
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Residuals vs Leverage
Standardized residuals
Leverage

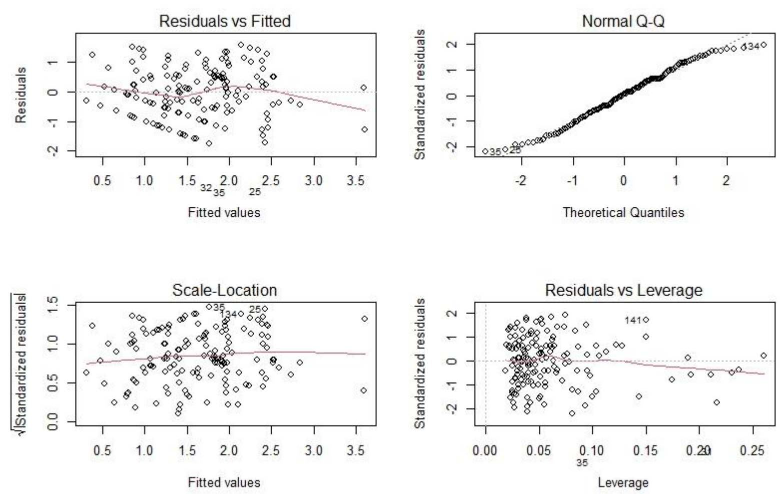
1A


Residuals vs Fitted
Residuals
Fitted values
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Residuals vs Leverage
Standardized residuals
Leverage

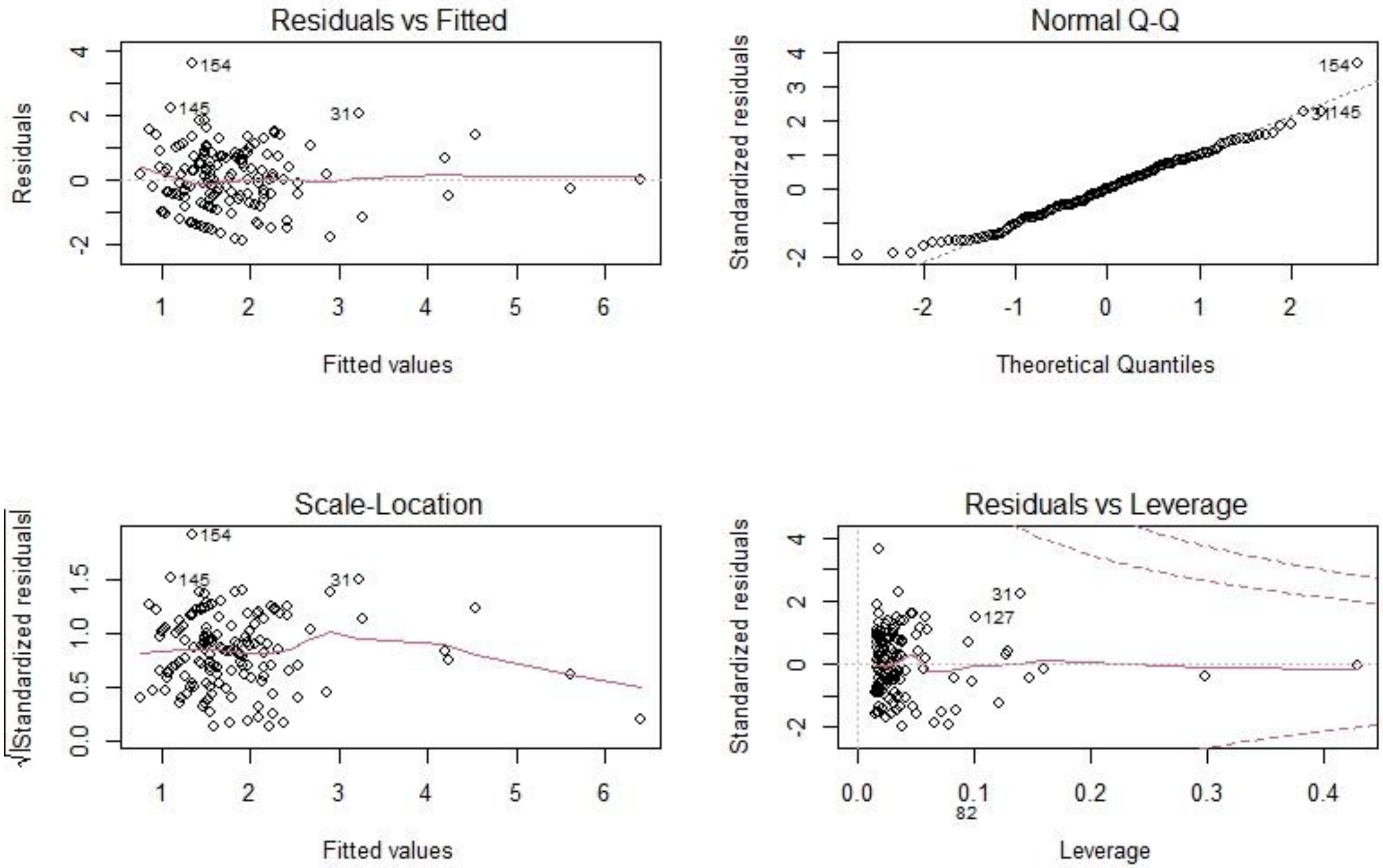
1B


Residuals vs Fitted
Residuals
Fitted values
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Residuals vs Leverage
Standardized residuals
Leverage

2A


Residuals vs Fitted
Residuals
Fitted values
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Residuals vs Leverage
Standardized residuals
Leverage

2B



3

Figure 6.3: Residual plots for multiple linear regression models with log-transformed median *E. coli* concentration as response variable and catchment characteristics as independent variables (n=154). Numbers correspond with models in Table 6.1. For each model, four plots are presented: in the upper left corner, residuals are plotted against fitted values, to assess if residuals have non-linear pattern that could indicate non-linear relationships, with distinguishable patterns indicating that no-linear relationships are present that the model cannot account for. In the upper right corner, a quantile-quantile plot for residuals allows evaluation of normal distribution of residuals (points should follow the straight line), as error terms must be normally distributed in a regression model. In the lower left corner, a scale-location plot shows if residuals are spread equally along the range of predictors, to check assumption of equal variance (homoscedasticity), with a straight line and randomly spread points indicating homoscedasticity. In the lower right corner, standardized residuals are plotted against leverage to find influential outliers, with points outside the red dashed lines influencing the regression line.

6.3.2 Sensitivity testing

For most catchments (107), no significant correlation with any of the rainfall periods could be observed (Table C.8). Only three catchments were classified as “Short”, 34 as “Medium”, and 10 as “Long” (Figure 6.4).

In the regression modelling, these categories were not retained in the model as a variable or in interaction terms. While some catchments may experience an effect of rainfall amounts (increasing concentrations, but no dilution effects could be detected as there were no significant negative correlations), overall, these hydrological differences seem to have no major bearing on median concentrations or how other factors such as land uses affect concentrations. In terms of catchment management, it might however be of importance to understand if source areas only get connected after a certain amount of rainfall, as it may offer opportunities to disconnect pathways.

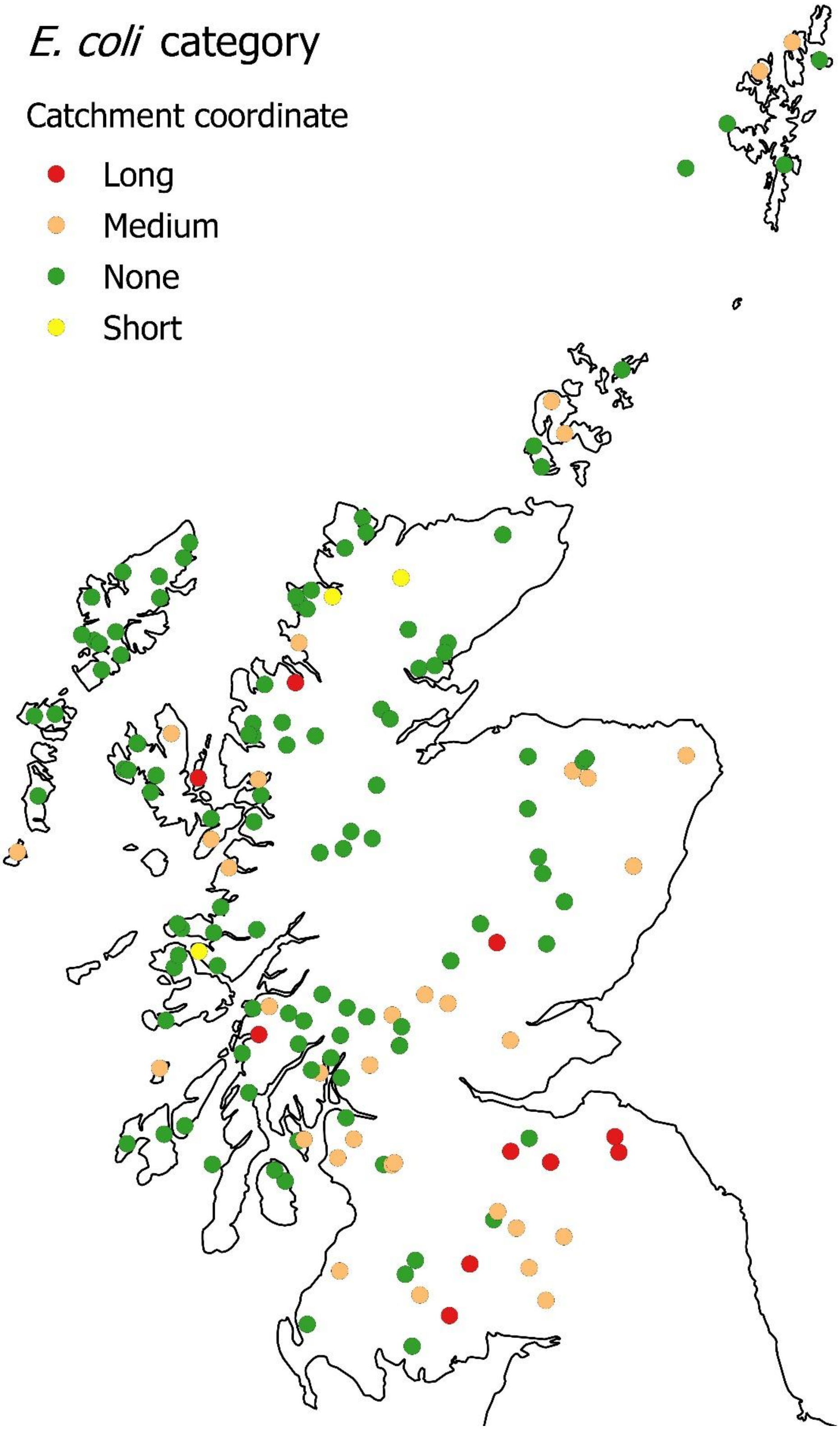


Figure 6.4: Spatial distribution of catchment categories allocated through the results of the Spearman's rank correlation tests run per catchment on sampled E. coli concentrations from 2011 – 2016 and different corresponding amounts of rainfall (daily total rainfa rainfall (daily total rainfall, and total rainfall of the preceding 3, 10, and 30 days) split by season (March to October – summer; November to February – winter). “Long” = at least one strong or two moderate correlations to a long rainfall period; “Medium” = at least one strong or two moderate correlations to a medium rainfall period; “None” = no or only one moderate correlation; “Short” - at least one strong or two moderate correlations to a shorter rainfall period.

6.3.3 Land use change

The future land capability assessment was based on projected climate data for the period 2041-2060 from the UKCP18, from a model driven by a higher-end scenario (see 3.4). It assumes that water is available for irrigation, so that drought is no limiting factor to land use. It thus represents a ‘worst case’ scenario. Under this scenario, biggest changes in land use can be observed in the South and East of Scotland (Figure 6.5), especially with shifts towards land that is more capable of sustaining arable agriculture (LC 1-3.2), and to a smaller extend towards land capable of sustaining the keeping of livestock (LC 4-5). Looking at how the distribution of land capability changes in the catchments (Figure 6.6), for many catchments especially in the Northeast, land classed as 4 or 5 shifts into the upper classes 1-3, so could potentially develop from a livestock-dominated landscape to one dominated by arable agriculture. Other catchments see a shift from land formerly classed as not capable of sustaining intensive agriculture to land classed as 4 or 5, and have a significantly higher percentage of these classes as a consequence, especially in the South of Scotland and on Shetland. Some catchments see increases in both these groups of LCs.

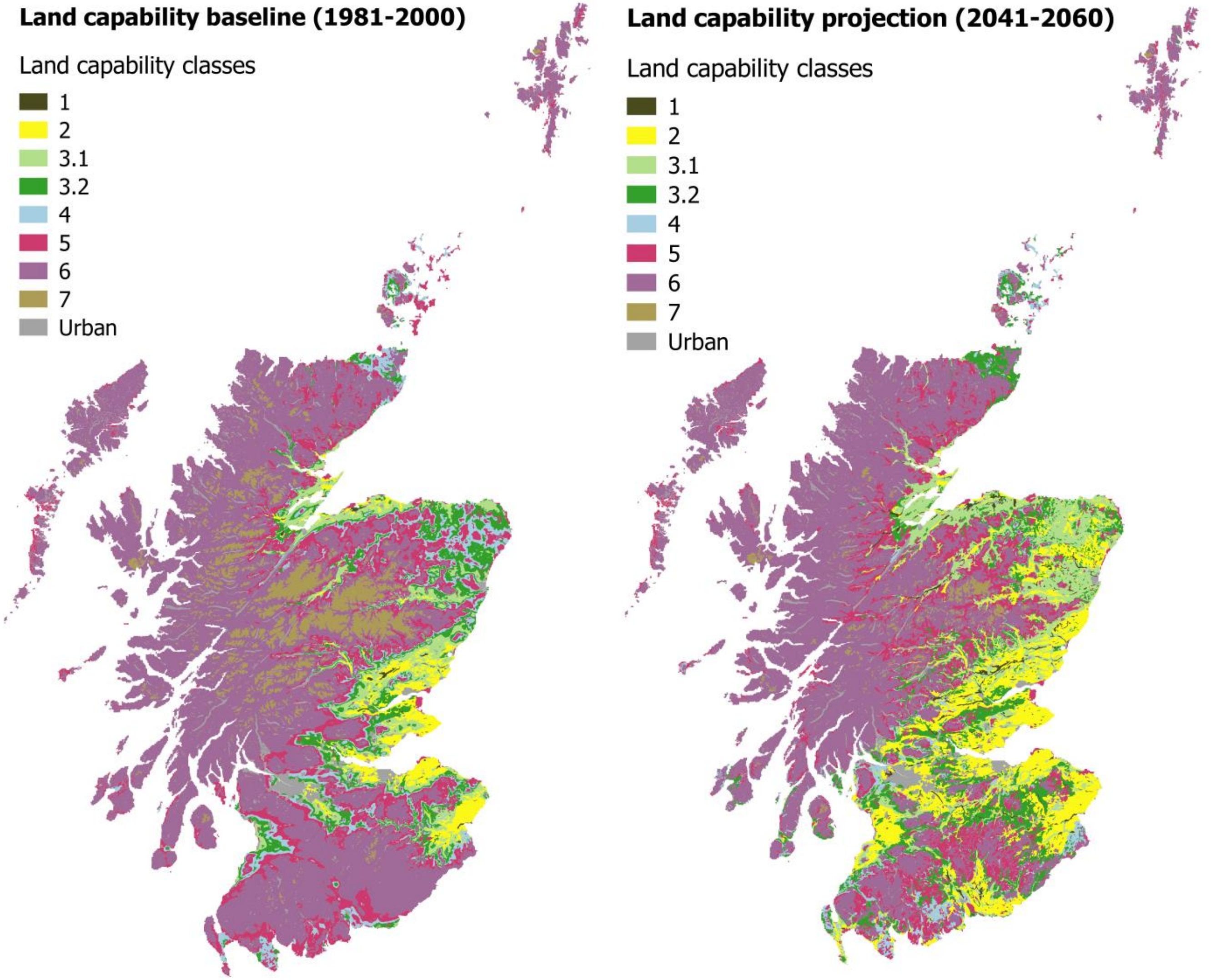


Figure 6.5: LCs calculated as in Brown et al. (2008), using A. climate indicator data from the baseline period 1981-2000 and B. projected climate indicator data from the UKCP18 for the period 2041-2060. Biggest changes especially in increases in LCs 2 and 3 can be observed in the South, along the Central Belt, and along the Northeast coast of Scotland.

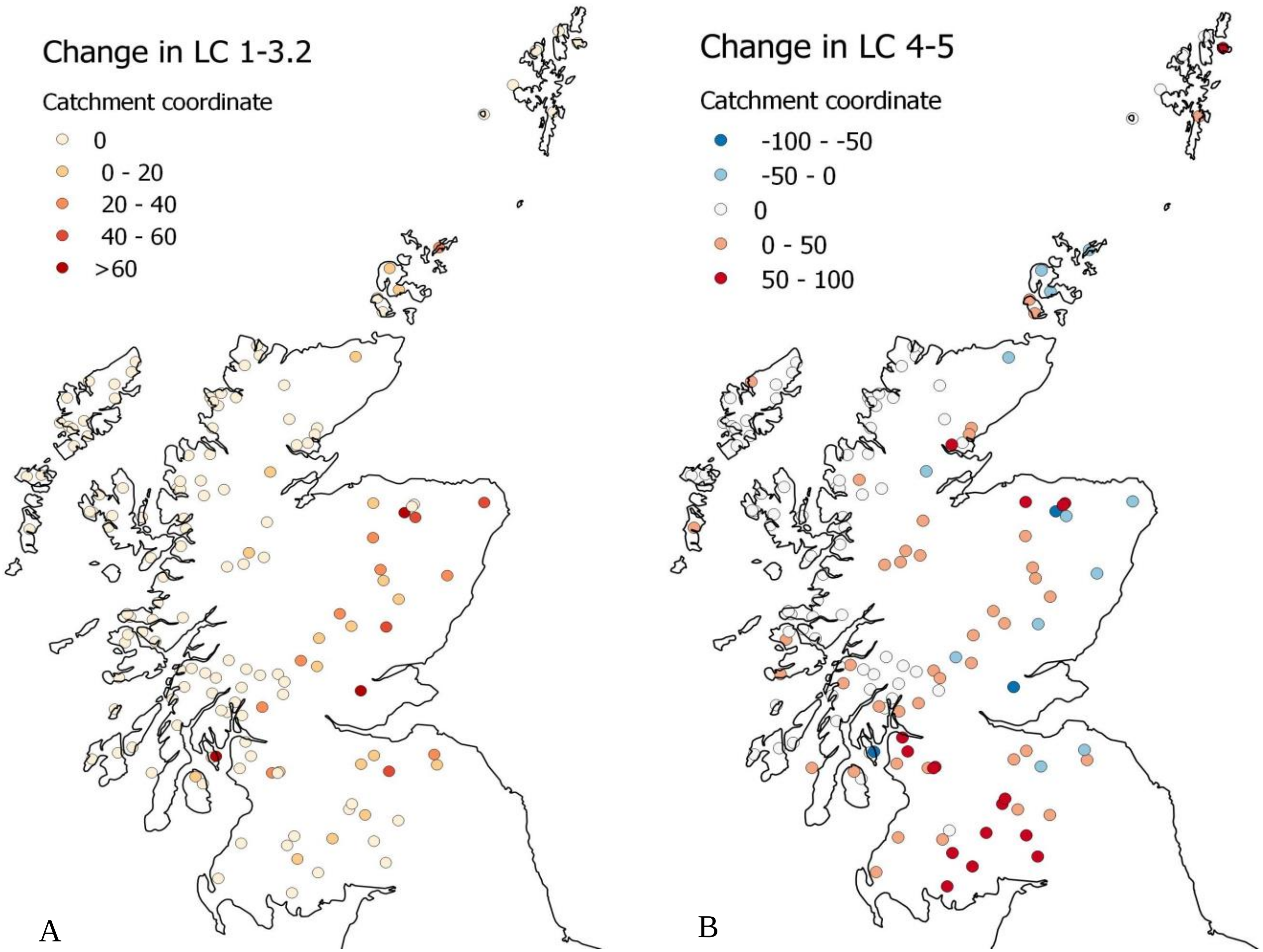


Figure 6.6: Changes in percentage of catchment area classified as A. LC 1-3, or B. LC 4 – 5, according to land capability projections derived with UKCP18 projections for a future period of 2041-2060 for climate indicators used in bioclimate and land capability classifications (Brown et al., 2008).

6.3.4 Conclusions for the risk screening

Model 3 was used to project future *E. coli* medians (Table C.7 & Figure 6.7A). This model is mainly driven by the change in LC 1-3.2, so the catchments showing the highest increases in *E. coli* median are those that show the biggest increases in these LCs. Catchments in the South are showing more moderate increases. Orkney catchments also show increases, whereas most catchments in the Northwest are unaffected. Comparing model estimates for the baseline period with the actual values on which it was modelled (Figure 6.7B), there are some strong over- or underestimations, which is to be expected since the variability explained by the model is less than 40%.

The risk map identifies as high risk some of the catchments that are already of concern in the Northeast of Scotland, but also highlights some catchments that are currently not problematic in terms of *E. coli*, especially in the South. As with the TOC maps, these screenings and allocated risk levels can now be combined with metrics that would represent the vulnerability of individual source-treatment systems and the national supply system.

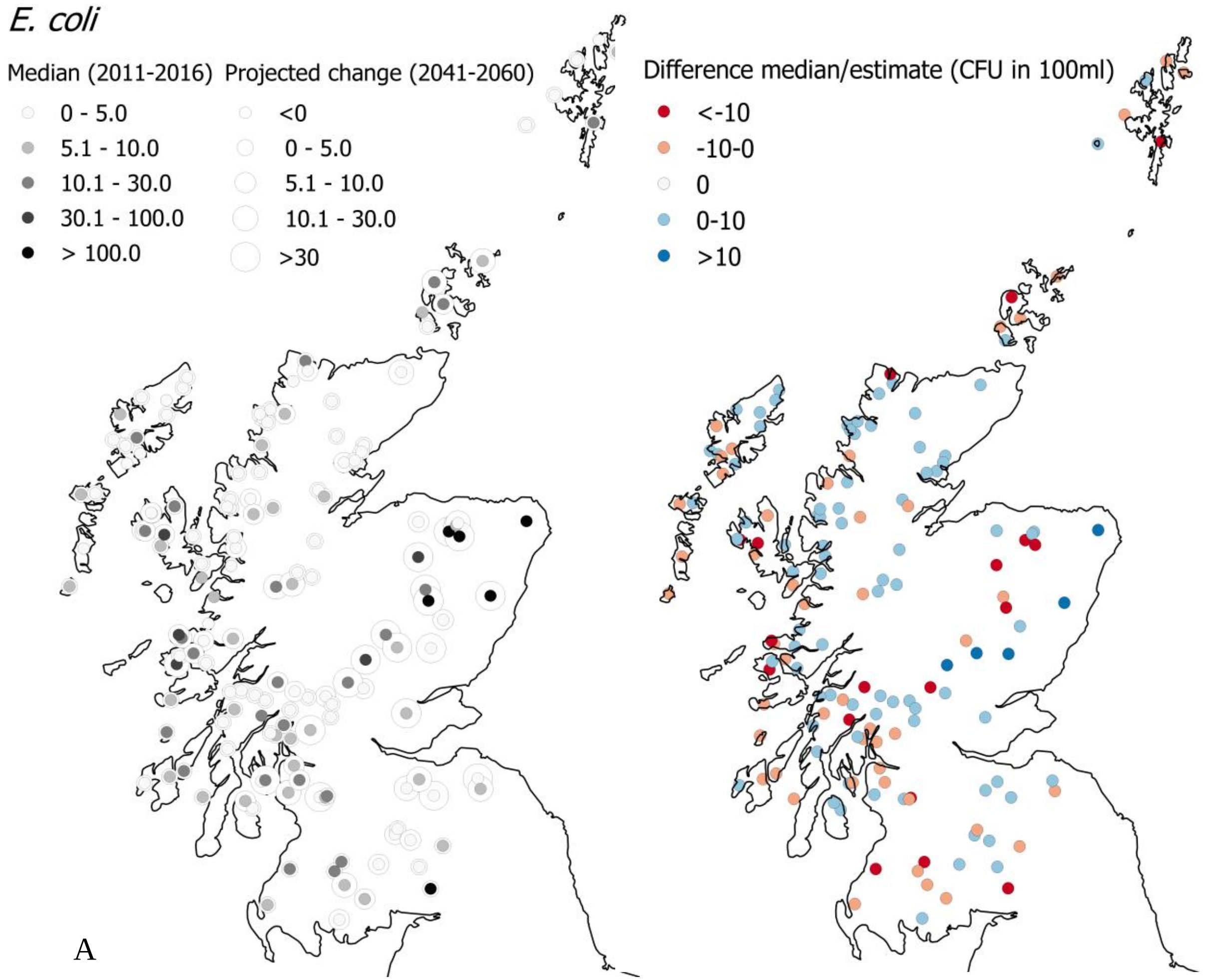


Figure 6.7: A. *E coli* risk map, based on projections for *E. coli* median concentrations from model 3, using projected percentages of catchment area with LC 1-3 and 4-5 (derived with UKCP18 climate data for 2041-2060 following procedures as in Brown et al. (2008). Point size reflects projected change in *E. coli* median concentration. Shade reflects current (2011-2016) *E. coli* median concentration. B. Model performance – indicates if the model is over- or underestimating the median by reflecting the difference between actual *E. coli* median concentration and estimated *E. coli* median concentration by the model.

6.4 Limitations and further approach

While land use in the form of improved grassland and arable areas are significant factors to explain median *E. coli* concentrations, they still only explain a rather low amount of variation. This could be due to several reasons, for example related to data issues, a mismatch between the scale of the data used to produce explanatory variables and the size of the catchment, or local conditions influencing the impact of overarching pressures. For example, if and how faecal indicator organisms (FIO) arrive at a surface water body will much depend on hydrological connectivity (Donnison et al., 2004, Tetzlaff et al., 2012, Neill et al., 2018). Nevertheless, land use can highlight where there is a potential for contamination and hence where catchments are at higher risk of elevated *E. coli* concentrations. These catchments can be targeted for further analysis, looking at critical source areas so a better estimation of risk can be obtained next to a starting point for mitigation approaches.

Projections in land capability to derive *E. coli* median concentration estimates for a future period are used as a basis for screening future risk for *E. coli* contamination. Contamination with faecal pathogens has been linked to land management, especially the number of livestock kept on a field, the application of slurry onto fields, and sources like farmyards and septic tanks (Hooda et al., 2000; Vinten et al., 2004). Land capability is one factor that contributes to determining how the land is used. Other factors, such as artificial change of natural conditions (e.g., drainage or irrigation), or farming traditions and practices, can lead to the land not being used according to its capacity. Furthermore, there may be local conditions not captured in the data used to calculate land capability that result in misclassification. Therefore, land capability (current or future) does not necessarily mirror land use.

Furthermore, specific land use does not translate into specific amounts of contamination. If and how much faecal pathogens reach the water body depends on a large proportion on local conditions and management. For example, the area where livestock is kept will determine how well, and when, it is connected to the water. Farming practices like the timing of slurry application, buffer strips, or preventing direct access of livestock to the water course also make a difference (Newell-Price, 2011).

In comparison to the multiple linear regression model on TOC medians, the model for *E. coli* medians is easier to interpret and it is easier to understand what is projected and where the limitations are. The screening map should only be used as an indication of general trends, rather than relying on the absolute values estimated for specific catchments for the future period. Catchments identified as high risk with this first risk screening can be examined further to integrate local conditions, to determine risk more accurately for these individual catchments (see Appendix D for suggestions). Further investigation could also explore which areas would be most detrimental for intensification in terms of water quality, what the impact of different forms of land use and practices might be on water quality, what management options could reduce the pressure, and what alternative land uses might be mutually beneficial alternatives. Direct effects of climate change, for example temperature dependent bacterial growth and reproduction, and hydrological connectivity of flow paths depending on weather events, should also be included in a further investigation. This could be achieved by more detailed, process-based models, and/or through the use of scenarios.

E. coli and coliform data are available for a wide range of catchments and tends to be at best weekly, but usually monthly or even three-monthly data. The models therefore need to be able to cope with low frequency data. Relevant catchment characteristics data will also be needed and might be available for the local and field scale, such as number of livestock, especially if landowners are involved in the process. The model should achieve a good representation of hydrological connectivity and pathways within the catchment to identify areas of high importance for water quality outcomes, and hence for intervention. An added challenge for models dealing with pathogens is their ability to multiply and die-off, and the complex, and partly poorly understood and predictable, relationships of the rates with the characteristics of the medium, e.g., soil/water temperature (Oliver et al., 2016). For decision-makers, a projection of *E. coli* concentration change with confidence margins would be useful for a first evaluation if engagement in the catchment is worthwhile. An evaluation of a range of different scenarios (climate, land use and management) would be most useful to establish potential management options and to engage stakeholders in a pro-active conversation about the future of the catchment. A model should therefore be simple enough to generate results in a reasonable time for a good number of scenarios and be understandable to a variety of stakeholders in terms of changing impacts and their consequences. At the same time, it needs to be reliable enough, as well as transparent enough, to gain the trust of those involved (Oliver et al., 2009).

7. Scotland as a ‘Hydro Nation’ and national exemplar to anticipate and respond to risks to raw water quality arising from climate change: lessons learned

The study suggested a staged approach to support a systematic inclusion of risks from climate change and building capacity to anticipate, respond and change within the water utility (Figure 1.1). In this approach, hazards, exposure, and vulnerabilities are identified to assess current patterns of risk to raw water quality in stage 1. Changes across space and time are examined in stage 2, enabling identification of changes in risk patterns. This requires understanding about relationships between different water quality parameters and their controls, as well as patterns of exposure and similarities between catchments in terms of vulnerability to pressures. The results support evaluation of risk control options for supply sites and the supply network at stage 3, and planning and decision-making on the programme and strategic level of the water company at stage 4.

This approach was tested for the public water supply system in Scotland, focusing on stages 1 and 2. This chapter discusses the findings from trialling the approach in terms of impacts of climate change on water quality, attempting to answer RQ 1 (7.1), implications for risk assessment and management (RQ 2, 7.2), the role of a catchment approach to manage risks (RQ 3, 7.3), gives a summary and some concluding remarks (7.4), and finishes with suggestions for further research and policy recommendations (7.5).

7.1 Climate change and drinking water quality

Water utilities, such as Scottish Water, are providers of public water supplies, wastewater treatment, and have wider responsibilities regarding the health of underpinning aquatic ecosystems. Utilities are bound by compliance responsibilities for multiple water quality parameters, for which they have standards to meet. In developed economies, as exemplified by the family of nations in the European Union, respective regional and national utilities will have a good idea about which parameters are problematic, at which areas, under which conditions, and have an estimate about the frequency of occurrence. Climate change might upend these patterns by changing the frequency or magnitude, but also by adding new parameters of emerging concern, either known or completely unknown. While studying the effect of climate change of one parameter in an individual catchment might provide some answers to the former aspects, the latter will remain unaddressed. It is therefore important that these studies are carried out in a context that enables transferability – within the catchment, meaning there is some indication about effects on “overall” water quality to understand viability of the individual water supply system in terms of water quality; and between catchments, meaning inferences can be made for other catchments, reflecting on the viability of the complete supply system.

In the present national-scale investigation, current water quality patterns were initially explored and related to catchment characteristics (4). Scotland has diverse natural conditions that are reflected in water quality, and the individuality of catchments are also seen in water quality variability. It was still possible to discern overall patterns related to catchment characteristics as well as geographically. Very broadly, two ‘types’ of catchments were distinguished with “upland” (dominated by semi-natural land cover, acidic, organic soils types with high yields of organic matter and metals, found predominantly in the West of Scotland) and “lowland” (with more intensive agricultural use, higher pH, and higher baseline concentrations of *E. coli* and coliform bacteria, found predominantly in the Northeast of Scotland) catchments, but acknowledging that these are stereotypical with many catchments falling somewhere in between and reacting to further controls, which however could not clearly be defined.

Following on from this, two water quality indicators (TOC and *E. coli*) were then chosen as diagnostic to be explored in more detail with regard to shifting patterns in space and time (5 and 6). Drivers for changes in water quality and areas of high risk were identified for each of the water quality indicators. For TOC, it was concluded that climate change is likely to have a direct effect on water quality through changes in precipitation patterns, especially due to prolonged periods of drought in summer followed by intense rainfall events. Depending on catchment characteristics, this could mean potential for increased production of DOC as wetness as an inhibiting factor decreases, coupled with increased temperatures; potential for increased mechanical disturbance during drought periods; or potential for increased peaks in concentrations after drought periods, where the buffering capacity of the system is low (especially river systems). For *E. coli*, it was concluded that climate change is likely to have a more indirect effect due to land use and management being a predominant factor in pollution risk, and changes in climate that favour more intensive agricultural practices carrying the risk of FIO pollution were projected for the Northeast as well as the “squeezed middle” of Scotland.

Rising trends in dissolved organic carbon have already raised concerns and required some investment into treatment within Scottish Water (Scottish Water, n. d.), and an assessment of the viability of continuing to upgrade treatment works to cope with rising concentrations is urgently needed. Similar trends and concerns are reported from other nations with large areas of peatland, such as in Northern Europe and North America. Similarly, pathogens, and especially *E. coli* are a focus for treatment around the world due to the potential disastrous consequences. A risk screening for these two water quality indicators is therefore in itself of interest for Scotland and many other countries. However, these parameters were also chosen in this case study as they were identified in the first analysis step as diagnostic and representative for a “type” of catchment. As colour correlated positively with aluminium, iron, manganese, and turbidity, and DOC is influenced and mobilised by the same processes as these parameters, it is likely that catchments experiencing increasing concentrations in TOC will also have increased metal and turbidity concentrations. *E. coli* has a long tradition as an indicator for faecal contamination, so projected increases in *E. coli* contamination risk also mean an increased risk for other FIO such as *Cryptosporidium*. As the increases are driven by increased agricultural, and especially arable use, it can also be hypothesised that issues around agricultural diffuse pollution, such as increasing nitrate or pesticide concentrations are probable. The typology is also a starting point to understand if emerging concerns are likely to become problematic at specific sites, although further fine-tuning would be needed especially if these emerging issues have different main drivers.

The hypotheses around the usefulness of diagnostic water quality indicators could be further tested either by enlarging the empirical analysis, for example looking at the datasets of two or more parameter within one or several catchments, and their relationships to rainfall/flow, or process-based models looking at more than one parameter. It would also be interesting to combine these two approaches to understand how strong the relationships are over space and time – do they vary e.g., during high flow and base flow conditions, and are there controls in catchments that could override them, leading to a decoupling? This could address some of the shortfalls of summarising data to median and 95th percentiles, help untangle some of the controls that could not be clearly distinguished in this study, and shape a more detailed typology of catchments.

Such a typology of catchments based on water quality outputs coupled with an understanding of controlling processes and underlying sensitivities is crucial in helping transfer findings of more detailed studies to other catchments. The initial attempt at an overarching typology (4) could identify predominant water quality concerns, relate them to certain catchment characteristics that were reflected to an extent in geographic distribution, and identify that local conditions must have an overriding influence in places. A typology based on response to climatic triggers as established for TOC during the risk screening (5.3.3) however constitutes a better step towards this goal, as it allowed to distinguish catchments based on their sensitivities, relate this to catchment characteristics and thus reflect on processes that dominate water quality responses. This simultaneously allowed to draw inferences regarding climate change and to understand how insights from one catchment might be transferred to others.

The analysis covered catchments across Scotland and reflected most of the natural and man-made conditions encountered in the country, except for exclusively urban areas and with an underrepresentation of arable agriculture. In terms of drinking water quality, these exceptions should not make transferability to either Scottish Water catchments that were not included in this study, or catchments used for private water supplies, invalid, as these would be similar catchments to the ones included in the study. Models are based on catchment characteristics which in turn were derived from freely available data, so characterisation of these catchments would be possible and model estimates could be produced for these. These estimates can be used in a similar way to this study, if the indicative nature of the numeric results with regard to risk is acknowledged, rather than using them as values to base a response strategy on. While the models itself could not be used for areas beyond the scope of the study, the methods could easily be adjusted to similar datasets for other regions to derive an understanding of relationships between water quality and catchment characteristics and a potential impact of climate change.

The strength of the dataset lay in its large spatial coverage together with a good number of sampled parameters. This opened up some exploration of the concept of space for time substitution, where spatial patterns and differences are used to infer how changing conditions shape the environment (Huang et al., 2019). This was especially relevant for the *E. coli* risk screening, where changes in climate would mean some catchments could see land use changes that would make them more closely resemble more intensively used catchments. Land use and management were identified as main drivers, and the model used to project risk in *E. coli* contamination was based on the assumption that catchments in the South and along the “lowland/highland divide” of Scotland would come to resemble Northeast areas of Scotland more closely. At the same time, at various points the analysis flagged up that catchments are incredibly complex with interactions between a multitude of characteristics. Many catchment conditions have evolved over thousands of years. Meanwhile, climate change advances at unprecedented rates, making it likely that comparable climate or land use conditions will be met by differing conditions in topography, geology, or soils, creating new and unique combinations, and potentially upending the potential for space for time substitution. In this case, an understanding of processes and their consequences is more appropriate as a base to formulate hypotheses about future outcomes.

Best insights were achieved when combining different methods and approaches. This was the case when analysing several water quality indicators at the same time (4), as well as when focusing on TOC (5). Different statistical methods allowed to unpick different aspects and patterns in the data, leading to a more comprehensive view of the catchments and their water quality. Exploring intra-annual patterns (5.3.2), sensitivities to climate variables (5.3.3), and linear relationships to catchment characteristics for baseline concentrations (5.3.4) led to a first formulation of catchment categories as well as to quantitative projections of future concentrations. This allowed risk level allocation and prioritisation and provided a transferability frame for any further in-depth analysis for priority areas, supporting strategic decision-making.

7.2 Assessing and managing risks from climate change

Risk assessment is a way for water utilities to achieve their overarching goal of securing reliable and safe drinking water at affordable cost. Climate change adds complexity that standard approaches to risk assessment cannot address. For example, changes can occur in the likelihood of occurrence and magnitude of hazards, the nature of hazards, patterns of exposure, and in vulnerabilities. Uncertainty enters through several routes, with uncertainty around climate change projections, potentially increased through downscaling to a catchment scale; uncertainty about processes within a catchment and how these will be influenced, and uncertainty about future changes in the catchment due to socio-economic drivers. The complexities and interrelatedness of systems often make predictions unachievable and assigning probabilities impossible.

Risk assessment is needed to understand how risks can be management and to support decisions on risk control. Taking consideration of the impacts of climate change in risk management means that part of the risk assessment is to identify where knowledge is limited, how this influences the ability to project what will happen in future, if more knowledge will support better decision-making, and if the cost of procuring more knowledge is justified. Especially the latter aspect resonates with the mandate of water companies to supply water with a quality of the required standard at minimum cost. Risk screening has been identified as a tool to support this through identifying priorities (for the nature of risks, for areas, or for systems), and outlining the nature and magnitude of uncertainties, and possible responses.

7.2.1 Risk screening

In the case study, the available data were initially used to screen the subset of national water supply catchments to better understand current patterns of hazards, exposure and vulnerability (4), effectively mirroring stage 1 of the proposed approach to integrate climate change impacts into risk assessment (Figure 1.1). The screening highlighted risks around colour and bacteriological parameters, associated with inherent catchment characteristics such as topography and soils, and land use and management. Also seen as diagnostic (see 7.1), TOC and *E. coli* were further analysed to understand main drivers for concentrations, how these might change over time and what the consequences would be. The analyses highlighted increases in risk especially for catchments in the South of Scotland due to decreases in SER, and in the Northeast and along the lowland/highland divide due to increased potential for agriculture, for TOC and *E. coli* respectively.

The risk screenings thus delivered several insights. First, they confirmed important drivers for concentrations and key processes upon which adaptive management should focus. For TOC, direct effects of climate change were discovered, as rainfall amounts as well as temperature were related to TOC concentrations, and these relationships could be explained by process knowledge. This identification of key drivers and processes supports a first indication of suitable adaptation and mitigation measures. For example, catchments identified as “temperature” driven (mainly wet, peaty catchments) are hypothesised to be especially vulnerable to a decrease in summer precipitation as wetness is an inhibiting factor in these catchments, meaning that a priority measure for these catchments should be to sustain water tables and wetness, to prevent a loss of sphagnum mosses, and support the integrity and health of associated ecosystems. For *E. coli*, indirect effects of climate change were identified as land use is a main control of contamination baselines and climate change is anticipated to remove constraints for associated land uses. While this doesn’t mean other potentially important drivers should be disregarded, the screening based on these premises supports a first risk scoring, prioritisation, and helps to focus potential further assessments to steer adaptive management. The screening suggests that management should focus on active engagement to incentivise water positive behaviour in catchments with increasing risk.

Second, they outlined the major limitations to our knowledge and sources of uncertainty. For TOC, uncertainty relates predominantly to the nature of risk factors, as natural processes and responses to climate triggers depending on different underlying catchment vulnerabilities are not well known. For *E. coli*, while risk factors, i.e., sources of contamination, are relatively well studied, the difficulty arises from forward projection of these factors and often overriding effects of local conditions. This also gives first indications of possible response options by understanding if and how uncertainties might be reducible (and if this is desirable). For TOC, it could be possible to reduce uncertainty through better understanding processes of TOC/DOC production, release, transport, and transformation. The risk screening identified catchments and catchment types where further monitoring and/or process modelling might be particularly beneficial, either because the catchment was identified as high risk, or because it might produce insights transferable to catchments of the same “category”, or both. The risk screening also highlighted the benefits of the natural integrity of ecosystems as they seem to be more resilient to future climate change. For *E. coli*, identification of critical source areas, scenario development, and ecosystem services assessment would be tools to reduce uncertainty, and/or provide a basis for decision-making under uncertainty, thus supporting further planning, engagement, and decision-making. Early stakeholder involvement could improve and steer modelling, encourage behaviour beneficial to water quality, and simultaneously generate other benefits to landowners, opening opportunities for ecosystem services based schemes.

Third, they identified areas of high risk and enabled risk mapping and ranking. Combining projections with details about the supply system, such as the treatment envelope of the plant, the number of people served by the plant, etc., identifies which systems are at risk of failing statutory requirements. These can be prioritised for the development of response options, which could include a more detailed investigation. Scores could also be allocated, to connect the outcomes and decisions made on programme level to planning on operational level and to established risk assessment procedures under Water Safety Plans.

7.2.2 Follow-on – priority catchments

The risk screening identified catchments that were at higher risk of seeing deterioration in water quality with regard to specific water quality aspects either because they appear to be more vulnerable to specific changes (in case of colour/TOC), or because they have a higher potential for increases in anthropogenic pressures (in case of *E. coli*). For some of these identified high-risk catchments, it can be indicated to follow up with further analyses to get a clearer picture of possible negative impacts and their consequences, and an evaluation of risk response options, supporting stage 3 of the process as laid out in Figure 1.1. As the risk screenings are based on statistical models, further investigation could address some of their limitations, and by focusing on an individual catchment, can better consider processes and individualities of the catchment, and focus on traits that have an impact on a local and field scale. For some individual sites, higher frequency sampling from real-time monitoring at water treatment works is stored for approximately three months (personal communication, Scottish Water), and could serve to look at fluctuations in concentrations over a shorter period of time.

Process-based models could improve understanding of processes around TOC and further develop the typology developed in 5.3.3. One of the challenges with models to help understand carbon concentrations in surface water is the need to include both carbon dynamics, and the biogeochemical processes of DOC production, in the terrestrial environment, and the hydrological flow-paths that route DOC into the aquatic environment, and processes therein (Birkel et al., 2014). Two specific developments of widely used process-based models are SWAT-DOC and INCA-C, which combine modelling carbon dynamics in terrestrial as well as aquatic environments. They have similar data requirements and have both shown to work reasonably well in temperate or boreal, mainly forested catchments (see 2.2.3 II). They offer the potential to evaluate the impact of changes in land use and management as well as shifts in climate. Both models are complex and highly parametrised. It is argued that the advantage of highly complex models is an increase in understanding of likely cause-effect relationships and system behaviour (Wade et al., 2008), and they are likely to be more reliable at reproducing the processes (McDonald & Urban, 2010). On the other hand, an addition of complexity and parameters makes the models susceptible to overparameterization and equifinality (Beven, 2012), meaning that differently parametrised models may perform equally well, so that it might become impossible to distinguish between the effects of parameters. This reduces process understanding and increases uncertainty.

Simpler models are argued to put the emphasis on processes that hold independently of spatial scales, and thus provide a stronger basis for projecting future system behaviour under changing conditions (Kirchner, 2006). However, such models don't work so well in exploring DOC dynamics in relation to land management (Birkel et al., 2014). Simpler models for soils DOC also don't necessarily include processes in the aquatic environment, so an additional model to simulate what happens to DOC in streams and reservoirs would be required. Again, there are multiple alternative ways forward to advance the science, as exemplified by the FREEDOM-BCCR project within selected water-supply reservoirs operated by Scottish Water, which specifically looked at DOC concentration trajectories in drinking water sources and the influence of climate change (Monteith et al., 2021). Within this project, the lake model PROTECH was used and newly parameterised to understand the production of DOC within reservoirs and the implications of changes in climate on these processes (Pickard et al., 2021).

Flow data are required for calibration and validation in all models, and the lack of flow data when looking at routine data from water utilities is likely to be a widespread issue, as data will mainly be only available for chemical parameters. A particular disadvantage of all of these models with regard to their usability for water utilities is that they seem to struggle with simulation of higher concentrations (Du et al., 2019; Futter et al., 2007; Meshesha et al., 2020), which is particularly problematic as the higher values are the most interesting in terms of water safety planning. Instead of aiming for an accurate simulation or projection of water quality, the focus of modelling could be on improving process understanding and identifying critical source areas. Calibration of a model and sensitivity and uncertainty analysis would contribute to this. Using several different models might help to understand effects of different parametrisation and process representation, although the effort involved in this might not be feasible for water suppliers, when compared to the outcome. Generally, any modelling will show what data requirements are for improving model accuracy and reducing uncertainties around model outputs.

In terms of more directly informing management decisions, a simpler model as developed by Birkel et al. (2014) could be particularly suited to be used in smaller headwater catchments with limited land use options to more quickly gauge the impacts of changes in temperature and precipitation, whereas more complex models such as INCA-C and SWAT-DOC would be best used for specific, high concern catchments with good data availability, that are a priority for mitigation and adaptation action. Here, careful modelling could help to understand carbon dynamics, how carbon concentrations might change and the associated circumstances, and how specific actions can influence these. While time consuming, the effort is justified by the priority given to the catchment due to prior identification as high concern and possible high benefit.

For *E. coli*, further analysis could identify for example critical source areas (and thus priority areas for management intervention) under a range of climate and land use scenarios. INCA-Pathogens, SWAT and HSPF are all established models that have been extensively tested for modelling of FIO, including for scenario analysis both in terms of management and of climate (see 2.2.3 II). They all have similar requirements in terms of data and studies have usually found them to perform acceptably to satisfactorily, and comparisons between them found no major differences in performance (Im et al., 2003; Nasr et al., 2007; Singh, Knapp & Demissie, 2004; Zeckoski et al., 2009). Generally, these models are aimed at larger, agricultural catchments, although SWAT has also been applied to very small catchments (Coffey et al., 2010). A specific challenge for modelling pathogens is their ability for growth and die-off, at rates which are influenced by environmental factors and highly complex (Oliver et al., 2009). As is the case for TOC, the models would require a certain amount of flow and water quality data for calibration. This would make them a better option for the larger catchments that are already of current concern (in the Northeast of Scotland), and where data are usually weekly or biweekly. Using them would be especially challenging for those catchments that have been identified as not yet of concern, but potentially increasing in risk, as it is likely that *E. coli* data will be very infrequent. Good calibration may not be achieved, and it may be difficult to distinguish dominant processes within the catchment. The initial risk screening can here be used to highlight catchments with low frequency data collection, where higher frequency monitoring should start, in order to gain a better basis for modelling due to potential negative future impacts.

For prioritising spatial targeting of mitigation measures, there might also be scope for less complex, risk-based models to determine relative risk of source areas (Oliver et al., 2016). Risk-based methods such as SCIMAP-FIO (Porter et al., 2017) can be quicker than process-based models, more accessible and communicable, and account for the many uncertainties from limited knowledge of pathogen growth and die-off and from limited data. They are especially useful when it comes to identifying and prioritising source areas, to support local management, but relative risk will not inform about changes in absolute risk and hence about impacts of future changes to drinking water quality risk. It would be possible to add complexity to the model that would allow an estimation of changes in relative risk due to changing climate, but calibration of this more complex model would be required, undermining the simplicity of the approach (Reaney et al., 2011).

The risk screenings elucidated drivers and specific pressures linked to water quality deterioration, areas of uncertainty, and highlighted particular catchments with higher risk of climate change related impacts on water quality. They highlight many questions that could be further explored to gain a better understanding of the nature and magnitude of change and of the associated aleatory and epistemic uncertainty, which will support assessment of suitable response options. While process-based and risk-based models can potentially deliver outputs to develop this understanding, most of these models require extensive data in order to achieve acceptable calibration and are time consuming to set up. Only if parameterisation and calibration is done with attention and care, these models will achieve their maximum potential and provide useful information for risk management. For water utilities, this means that the effort will have to be weight against the gain. Where the output from the insights process-based models can give is expected to be valuable, adequate data collection (e.g., higher frequency water quality indicator sampling, flow data collection, etc.) and model development can be carried out. Results from model runs, especially when used with scenarios, can also be used to communicate different options to stakeholders to find mutually beneficial management responses.

However, models still have limitations, including that they often struggle with uncertainty in their outputs even when data input is high, and that results obtained from them are not necessarily transferable to other catchments. Categorisation such as carried out within the TOC risk screening (see 5.3.3) could help to gain an insight which processes may be shared in other locations. The limitations again emphasise the need for robust adaptation actions, including the (re-)creation of resilient catchments that are able to buffer or adapt to future changes and keep providing ecosystem services.

7.2.3 Follow-on – the complete supply system

While the risk screening identified priority catchments for follow-up that could include further investigation and intervention, it also provided insights and a frame for consideration of the totality of the supply system that allows a more strategic approach to integrating climate change information, enabling stage 4 of the proposed approach. As discussed in 7.1, it supports transferability of insights gained from more detailed follow-on investigation in priority catchments, and to areas and catchments currently not included in the study or monitored, giving some ideas about long-term viability of areas and catchments as drinking water sources. Additionally, rather than considering climate change impacts in some priority catchments and evaluating response options for these catchments individually, climate change and its consequences can be considered for the complete supply system and business. This could include, or lead to, general decisions about preferences for dealing with uncertainty, hence moving from risk management (at the operational level) to uncertainty management (at the strategic and programme level).

Regionally, overarching scenarios could be built around scenarios used in climate change assessments, and consider further socio-economic aspects especially important to water utilities, such as changes in consumer behaviour and water consumption, use of alternative water sources, and water reuse schemes. Once these scenarios have been developed, they can be used in strategic planning for long-term viability of the utility (Luís et al., 2021), but they can also be used in the risk assessment for individual supplies, providing comparable assessments and a more holistic view. As they would consider wider business objectives, they give more consideration to consequences of risk control options and management decisions for aspects other than water quality, such as carbon offset, supporting other business objectives like achieving net zero. In turn, at the operational level, evaluating the effect of response options with regard to their outcome on water quality as well as other aspects ensures that measures, for example under the umbrella of net zero, are spatially well targeted and consider biophysical opportunities, rather than driven by socio-economic factors (Brown, 2020).

7.3 The role of catchment management and ecosystem-based adaptation

Experiences suggest that there is huge potential in catchment interventions to improve water quality in terms of economic, but also non-economic and societal benefits (see 2.3), and catchment approaches under the umbrella of EbA could become a crucial overarching strategy to increase water utilities' resilience to future changes. But there are also challenges associated with a catchment approach that may pose barriers for water utilities to explore this potential.

7.3.1 Challenges and opportunities of a catchment approach

Catchment-based initiatives need time to develop their potential and are therefore often only beneficial in the long-term. They still usually need an up-front investment and are thus financially disadvantageous in the short-term. They might also benefit some stakeholders, but mean a disadvantage to others. Therefore, catchment-based measures might need a finance plan or financial incentives for other stakeholders. In this context, Payment for Ecosystem Services schemes become especially interesting. These schemes constitute a concept to internalize environmental costs that are often not accounted for through the creation of markets (Schomers & Matzdorf, 2013). Compensating landowners for employing practices that are less harmful to the environment, or in this case more specifically cause less pollution to water sources, means that costs associated with water pollution are taken account for by paying for less impactful practices.

The case of Vittel shows that the lack of trust between stakeholders is an obstacle to successful negotiation, and that acceptance of farmers for mitigation measures is increased when taking part in determining which measures to adopt, and when these are tailored to their farms (Depres et al., 2008). Hence, a strong bottom-up element is necessary. This also helps to ensure that catchment schemes are advantageous to some stakeholders, which take an interest in the continued success of it. The necessary process can reduce conflict and enhance understanding of each other, as can be seen in the example of the Watershed Agricultural Council in New York (Appleton, 2002). It is also a means to engage the people that live in the catchment in formulating visions for the future of the region they live in. The bottom-up element however also brings along a specific management effort to make sure that views of stakeholders are heard throughout the processes and over the duration of the project. Often, the voluntary nature of the implementation of mitigation measures is a crucial aspect to engage key stakeholders (Appleton, 2002). Apart from being often regarded with suspicion by regulating bodies and non-governmental environmental organisations, voluntary initiatives can mean that measures are taken up that are most attractive to the land manager, rather than the most effective measures in terms of water quality (Willett & Porter, 2001).

Other challenges include the need to target mitigation measures well to the specific circumstances of the catchment, in terms of hydrology and socio-economic conditions, to make sure that the most effective measures to create win-win situations are chosen. This also means well conducted and comprehensive prior and accompanying monitoring. While some local interventions can reduce water pollution at the site scale, evidence showing performance at larger, e.g., whole catchment scales is sparse. For several reasons, it is difficult to assess the effectiveness of catchment-based mitigation measures such as changes to land management practices. Firstly, there are multiple, often highly interdependent, factors operating in a catchment such as economic, political, and climatic conditions that make it difficult to trace changes back to specific interventions. For example, mitigation measures for nitrate in agricultural catchments as implemented in NVZs in the UK have not led to the expected improvements (Worrall et al., 2009), which could be due to positive impacts being masked by opposing trends from other pressures. Secondly, there is a time lag between implementation of mitigation measures and observable results, e.g., due to buffering of soils or accumulation of nutrients in the soil, and temporal lags can also differ for different interventions (Dunn et al., 2014). To evidence effects of mitigation measures, it is necessary to entangle their effects from other factors such as climate variations, land use changes, and influences of other economic and policy drivers.

A particular challenge regarding climate change mitigation and adaptation actions stems from the fact that activities carried out under a climate change umbrella are often not connected to ecosystem services and a catchment approach, with climate change seen as a more cross-cutting issue whereas activities carried out under a catchment management approach are seen as commitments for nature conservation or ecosystem services planning (Wamsler et al., 2014). Consequently, there are several aspects that may be overlooked. Catchment approaches to improve water quality might fail to integrate information on climate change and thus to achieve long-term improvements in water quality, or have unwanted side effects. For example, Iacob et al. (2014) point out that afforestation could exacerbate existing drought problems if drier summers are projected. Measures to improve water quality, especially afforestation, land use change or ecosystem restoration, will have to consider future climate conditions to ensure long term viability. Catchment-based measures with a focus on one ecosystem service, such as water quality regulation, may fail to adequately consider wider benefits, especially climate mitigation and adaptation benefits, and appear a less attractive option – vice versa, carbon storage measures might not consider trade-offs, or additional benefits, missing out on opportunities.

7.3.2 Catchment approaches and risk management

The uncertainty over concrete outcomes of projects for improving water quality, and challenges for framing project and programme intervention costs against complex, harder to monetize, multiple benefits in more than one area of interest, means that the potential beneficiaries, water supply companies and consumers, often perceive catchment-based approaches as high risk. The difficulty of quantifying non-economic benefits makes it difficult to incorporate these benefits into economic valuations of risk control options, which is often the basis for decision-making. Therefore, current approaches, where risk control options are assessed within the framework of an individual supply site, require a shift towards ‘systemic thinking’, where trade-offs and co-benefits between ecosystem services are considered across space, time, and organisational structures to optimise representation of societal values (Everard & McInnes, 2013). While this may seem like a deviation from the “core” mission of water utilities to provide sufficient, safe, and affordable drinking water, it supports internalising all the costs associated with achieving this, and capitalising on all the benefits gained. This in turn allows a fuller, more comprehensive evaluation of these measures as a response option in risk assessment and management, especially with regard to assessing their robustness. This requires information flow between the different management levels of the drinking water provider (see 2.2.2 II) in the planning and decision-making process.

Response and control options for an individual water supply system, including catchment-based approaches, are made at the operational level and are likely to revolve around water quality. Wider benefits may be considered but will only become relevant if they are communicated and taken into account on the levels where these are considered: on the programme level where long-term viability of the complete supply system is considered, and the strategic level, where overall business aspects including moving towards net zero, energy efficiency, customer trust, etc. are considered. Vice versa, where EbA for supporting net zero is considered, this should be prioritized in catchments where most additional benefits can be achieved, especially for water quality. The challenge in expressing wider and additional benefits in monetary terms may then also be less pronounced. Achieving politically desirable outcomes such as contributing to carbon offset, building customer trust, achieve a positive perception with consumers as well as on a political level, can be better considered and valued at a strategic level.

There are several examples where drinking water suppliers have taken a catchment approach to complement their traditional treatment, and have found not only an economic benefit, but also increased other benefits that are expressed as ecosystem services enhancement. Lessons learned from these experiences include the need to engage landowners and other stakeholders to give them a certain degree of autonomy over the selected measures, the requirement of prior and accompanying monitoring to ensure maximum effectiveness of implemented measures, and a potential to use new market approaches to value water resources such as Payment for Ecosystem Services.

In view of mounting pressures, increasing catchments' resilience to change helps counterbalancing the negative impacts on water quality, while achieving wider benefits for the provider and a range of other stakeholder. ICM, EbA and NbS are key concepts that need to be incorporated into drinking water quality risk management and water safety planning. Assessing, and being able to capitalise on, all benefits of a measure are likely to move catchment-based initiatives towards low regret options, whereas the energy-intensive nature of treatment is likely to reduce reliance on treatment alone. In fact, those promoting NbS see them as part of a multi-measure solution, rather than an alternative to engineered approaches (Seddon, Chausson et al., 2020). This sits well with an approach of using different, and complimentary, types of response options as systemic interventions (Brown & Everard, 2015). While this typology of response options is formulated in a policy context, the typology could serve water utilities to characterise different risk control options, for example in their ability to control for specific water quality outcomes; in their potential to achieve additional benefits to either the company or other stakeholders; in the degree of necessary involvement of stakeholders; etc. This could lead to a mixture of options that achieve a maximum of additional benefits with a “back up” that ensures meeting statutory requirements in water quality.

Focusing on the catchment, its health and resilience, will not only support reaching the statutory requirements for drinking water quality, but also shifts the understanding of “good water quality” towards a more holistic perspective that resonates with other stakeholders, and policy and regulatory frameworks such as the Water Framework Directive. “Good ecological status” is linked to the type of catchment – different catchments will yield different water quality profiles. Understanding what kind of catchments will produce what kind of water quality (see a first attempt in 4.4.2) allows to identify drivers, pressures, and key processes that become the focus of climate change adaptation, by supporting the catchment to yield its “optimum” quality. Water treatment then becomes the tool to deliver specific requirements for drinking water where water quality naturally exceeds those.

7.4 Summary and conclusion

Climate change will have negative impacts on water quality, either directly through changes in hydrology, often associated with more frequent extreme events, or indirectly by increasing pressures on water resources and their catchments. Integrating climate change information into risk assessment to prepare for and address impacts through mitigation and adaptation actions is a current major challenge for water utilities. Current risk assessment and management in the water utility sector is usually built on traditional views of risk as consisting of the likelihood of occurrence of an event multiplied by its consequences, and Water Safety Plans are an accepted frame for risk assessments of single supply systems from source to tap. Climate change challenges this concept, with deep uncertainty making accurate predictions and assigning of likelihoods unattainable. With this in mind, the research proposed a framework to build capacity to integrate climate change considerations into risk assessment and management, and used a national assessment of Scotland to test the framework and answer three research questions:

RQ 1: How is climate change likely to impact raw water quality of Scotland's public water supplies?

In two assessment steps, different methods of empirical analysis were used to piece together a picture of current issues in public water supply sources in Scotland and their linkages to catchment properties, underlying vulnerabilities, and spatial patterns. Two water quality indicators, colour and *E. coli*, could be identified as of concern and diagnostic for a suite of water quality issues stereotypically associated with 'upland' and 'lowland' catchments. Analysing pressures and catchment vulnerabilities for them more closely was necessary to understand how changes in climate and land use could impact different catchments. Projections for climate and land use changes then allowed evaluating how risk of contamination might change in the future over the national supply system, and implementing a first risk scoring. Visualising the results in risk maps showed spatial patterns in risk development.

The TOC risk map highlighted catchments susceptible to decreases in rainfall over the summer, and to a lesser extend to increases in temperatures, leading to risk of overall increases in TOC concentrations and increases in peak concentrations. From a diagnostic point of view, it is likely that apart from TOC, other soil-derived water quality parameters mobilised by similar processes show similar risk patterns. Due to the projected pattern of change in the climate variables, this particularly highlighted catchments in the South and East of Scotland. The *E. coli* risk map focused on increases in agricultural land use potential as risk factor for contamination, and highlighted increasing risks in the Northeast as well as the "squeezed middle" of Scotland.

RQ 2: How can risks to water quality from climate change be managed in a drinking water context?

Rather than using a risk matrix of likelihood of occurrence and severity of consequence, derived from analysing past trends, the research focused on spatial and temporal patterns in water quality to understand dominant processes and connect this to vulnerability of catchments to pressures. This allowed to formulate hypothesis around impacts of projected changes, and visualising these by mapping change in exposure to specific hazards for identified vulnerable sites. Risk was thus assessed as a consequence of hazard, exposure and vulnerability, showing it to be a workable concept when attempting to integrate climate change information into risk assessment and management.

The assessment constituted a high-level risk screening, which provides a basis for risk-ranking and further investigation of individual sites, feeding into Water Safety Plans. It identified areas of uncertainty, pointing towards appropriate methods to either reduce uncertainty (e.g., through additional monitoring and modelling), or deal with it through robust approaches (underpinned e.g., through scenario modelling and analysis). It provides a frame for transferability, and for strategic evaluation of long-term viability of the water supply. Ongoing review of the risk screening including its different approaches (catchment-water quality relationships, typologies, scenarios of change) supports adaptive management.

RQ 3: What role does a catchment approach play in mitigating and adapting to climate change?

In this study, catchments were understood as crucial systems within the supply system onto which risk control should be focused, rather than a variable start point in a row of nodes for risk management. This puts the focus on building ecosystem resilience to changes and opens up potential for multiple benefits. Connecting overarching objectives of the water utility to risk assessment and management at the operational level further emphasises this. In this context, ecosystem services assessment, or similar, of the supply sources including their catchments become a crucial tool to understand wider implications of control options, rather than a “nice to have” add-on. Treatment works become the “last resort” safety net, catchments key assets, and EbA a core strategy. Moving towards and strengthening this concept resonates with Scotland’s mission to be a ‘Hydro Nation’, leading on innovative management of water resources and maximising their value.

7.5 Further research suggestions and policy recommendations

The analysis has highlighted several areas where further research could improve the findings, add further insights, or address new questions.

The risk screening could be complemented in different ways, for example through generating predictions from the multiple linear regression models using different climate projections, to test sensitivity of the predicted water quality outcomes. As more data become available over time, the possibility to analyse annul trends grows, or to compare different years especially when climate extremes (such as very dry or wet seasons) occur. This could add further approaches to risk screening and provide additional pieces of the puzzle. Improved data to represent catchment characteristics would improve this further. The developed models could be tested on catchments outside the scope of the study to check their performance, and to test transferability. The analysis could also be repeated for different water quality indicators, or a group of water quality indicators. Multi-target regression could be a promising approach to achieve a more holistic picture, especially to further test the diagnostic nature of the water quality indicators. It could also serve to identify “hot spots” for priority intervention. Especially for TOC, including further water quality indicators could help to sharpen the categorisation that has been developed and create a clearer picture of different “types” of catchments and their sensitivities and vulnerability. This would be especially useful as a basis for building climate change resilience as it combines empirical findings with process knowledge and would additionally allow a quicker estimation of where water quality concerns may appear when their emergence is first noted. There is also a strong link to water quantity through catchments having been identified as being sensitive to a reduction in SER for different reasons. A better understanding of relationships between water quality and water quantity should improve assessments of long-term viability of supply systems.

The study utilised a dataset from the routine monitoring of a water utility. This meant there were certain limitations (see 3.1), especially the lack of accompanying data for e.g., flow, temperature etc. While it is acknowledged that collecting data like flow and temperature would mean a considerable increase in monitoring effort, it might be possible to adjust sampling to better support similar studies, such as ensuring equally spaced sampling intervals. Based on this work, some catchments could be chosen where a large suite of parameters are sampled at a higher frequency. Real-time monitoring data could be stored for a longer period at selected treatment works, which would open up more possibilities for investigation especially for an analysis of peak concentrations. There should also be discussion about harmonising datasets from different institutions and to create a national database, to support combining data for better analysis.

As described in 7.2.2, there is a lot of potential for follow-on investigation in identified high risk catchments (see Appendix D for examples). This could span from trying to achieve more accurate water quality predictions (for one or more parameter), ecosystem services assessment, to scenario development. All of these are a good basis for stakeholder engagement. Again, this could range from mere information about water positive behaviour, to workshops to develop a shared vision for the catchment, or to the creation of incentive schemes such as payment for ecosystem services.

The risk screening should also regularly be updated with new data. Over time, identified options for the future will close, while others will emerge. Management needs to be sufficiently adaptable to deal with these changing circumstances and risk screening provides a tool to anticipate and respond to a range of possible futures. This can include attempts to make catchments resilient to changing conditions, so that water quality and quantity remains stable, adaptable treatment, and identification of supplies to supplement or replace existing supplies if required, etc.

There is a need to think and work across levels and hierarchies within a water utility to make sure that overall goals and visions are taken account of at the operational level, and that contributions of individual sites are incorporated into strategic planning. Additionally, working across technical groups and divisions can enhance information flow and holistic thinking, and bringing together expertise for example from those walking the catchments, running models, and carrying out risk assessments (Giffoni et al., 2022). As a way towards this, an ES or EbA framework could be developed where priority ecosystem services for the utility are identified, and their delivery assessed over the utility and/or at individual supply systems. This would provide a basis to develop strategies to create win-win situations where risk control and management enhances a maximum of services, and thus increase efficiency to meet business priorities (Gärtner et al., 2022). Finally, while providing safe and reliable drinking water is the overarching goal for a water utility, it is one of many uses we make of our water resources. Scotland, with its ambition to be a ‘Hydro Nation’, could be an international example for and leader towards breaking up established structures of compartmentalising water usages, by working across organisations to create, and compensate for, ecosystem services provisions to accurately reflect societal needs and priorities.

References

Legal documents

Action Programme for Nitrate Vulnerable Zones (Scotland) Regulations 2008

Bathing Waters (Scotland) Regulations 2008

Council Directive 91/676/EEC of 12 December 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources

Council Directive 91/271/EEC of 21 May 1991 concerning urban waste water treatment

Cross-Border River Basin Districts (Scotland) Directions 2014

Cryptosporidium (Scotland) Directions 2003

Directive (EU) 2020/2184 of the European Parliament and of the Council of 16 December 2020 on the quality of water intended for human consumption (recast)

Directive 2000/60/EC of the European Parliament and of the Council establishing a framework for the Community action in the field of water policy.

Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC

Directive 2006/113/EC of the European Parliament and of the Council of 12 December 2006 on the quality required of shellfish waters

Directive 2013/39/EU of the European Parliament and of the Council of 12 August 2013 amending Directives 2000/60/EC and 2008/105/EC as regards priority substances in the field of water policy

Private Water Supplies (Scotland) Regulations 2006

Public Water Supplies (Scotland) Regulations 2014

Scotland River Basin District (Standards) Directions 2014

Surface Waters (Shellfish) (Classification) (Scotland) Direction 2012

Surface Waters (Shellfish) (Classification) (Scotland) Regulations 1997

The Water (Scotland) Act 1980

The Water Environment (Controlled Activities) (Scotland) Regulations 2011

The Water Environment (River Basin Management Planning: Further Provision) (Scotland) Regulations 2013

The Water Environment and Water Services (Scotland) Act (WEWS) 2003

The Water Industry (Scotland) Act 2002

Reports, papers, and books

Abesser, C., Robinson, R., & Soulsby, C. (2006). Iron and manganese cycling in the storm runoff of a Scottish upland catchment. *Journal of Hydrology*, **326**, 59-78.

Abbott, B. W., Gruau, G., Zarnetzke, J. P., Moatar, F., Barbe, L., Thomas, Z., Fovet, O., Kolbe, T., Gu, S., Peirson-Wickmann, A.-C., Davy, P., & Pinay, G. (2017). Unexpected spatial stability of water chemistry in headwater stream networks. *Ecology Letters*, **21**(2), 296-308.

Acreman, M. C., Fisher, J., Stratford, C. J., Mould, D. J. & Mountford, J. O. (2007). Hydrological science and wetland restoration: some case studies from Europe. *Hydrology and Earth System Sciences Discussions*, **11**, 158-169.

Adger, W. N., Brown, I. & Surminski, S. (2018). Advances in risk assessment for climate change adaptation policy. *Philosophical Transactions of the Royal Society A*, **376**, http://dx.doi.org/10.1098/rsta.2018.0106.

Adger, W.N., Arnell, N.W. & Tompkins, E.L. (2005). Successful adaptation to climate change across scales. *Global Environmental Change*, **15**, 77–86.

AECOM (2015). *Aggregate Assessment of Climate Change Impacts on the Goods and Benefits Provided by the UK's Natural Assets* [pdf]. https://www.theccc.org.uk/publication/aecom-assessment-of-climate-change-impacts-on-uk-natural-assets/ [06/05/2022].

Aherne, J., Futter, M.N. & Dillon, P. J. (2008). The impacts of future climate change and sulphur emission reductions on acidification recovery at Plastic Lake, Ontario. *Hydrology and Earth System Sciences*, **12**, 383–392.

Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*, **2**(3), 370-374.

Allcock, R. & Buchanan, D. (1994). Agriculture and fish farming. In P.S. Maitland, P.J. Boon & D.S. McLusky (Eds.), *The Fresh Waters of Scotland: A National Resource of International Significance* (pp. 365-384). Wiley.

Anbumozhi, V., Radhakrishnan, J. & Yamaji, E. (2005). Impact of riparian buffer zones on water quality and associated management considerations. *Ecological Engineering*, **24**, 517-523.

Appleton, A. (2002). *How New York City Used an Ecosystem Services Strategy Carried out Through an Urban-Rural Partnership to Preserve the Pristine Quality of Its Drinking Water and Save Billions of Dollars and What Lessons It Teaches about Using Ecosystem Services* [pdf]. https://vtechworks.lib.vt.edu/bitstream/handle/10919/66907/2413_pes_in_newyork.pdf?sequence=1&isAllowed=y [10/05/2022].

Armstrong, A., Holden, J., Kay, P., Francis, B., Foulger, M., Gledhill, S., McDonald, A. T. & Walker, A. (2010). The impact of peatland drain-blocking on dissolved organic carbon loss and discolouration of water; results from a national survey. *Journal of Hydrology,* **381**, 112-120.

Arnold, J. G., Srinivasan, R., Muttiah, R. S. & Williams, J. R. (1998). Large area hydrologic modelling and assessment part I: model development. *Journal of American Water Resources Association*, **34**(1), 73-89.

Arora, S. & Keshari, A. K. (2021). Pattern recognition of water quality variance in Yamuna River (India) using hierarchical agglomerative cluster and principal component analyses. *Environmental Monitoring and Assessment,* **193**, no. 494, https://doi.org/10.1007/s10661-021-09318-1.

ASC (2016). *UK Climate Change Risk Assessment 2017 Evidence Report – Summary for Scotland.* Adaptation Sub-Committee of the Committee on Climate Change [pdf]. https://www.theccc.org.uk/wp-content/uploads/2016/07/UK-CCRA-2017-Scotland-National-Summary.pdf [10/05/2022].

Barbosa, A. E., Fernandes, J. N. & David, L. M. (2012). Key issues for sustainable urban stormwater management. *Water Research,* **46**, 6787-6798.

Battarbee, R. W., Shilland, E. M., Kernan, M., Monteith, D. T. & Curtis, C. J. (2014). Recovery of acidified surface waters from acidification in the United Kingdom after twenty years of chemical and biological monitoring (1988-2008). *Ecological Indicators*, **37B**, 267-273.

Bengraïne, K. & Marhaba, T. F. (2003). Using principal component analysis to monitor spatial and temporal changes in water quality. *Journal of Hazardous Materials*, **B100**, 179–195.

Bertram, J., Corrales, L., Davison, A., Deere, D., Drury, D., Gordon, B., Howard, G., Rinehold, A. & Stevens, M. (2009). *Water safety plan manual: step-by-step risk* *management for drinking-water suppliers.* WHO [pdf]. https://www.who.int/publications/i/item/9789241562638?msclkid=36e9027ad06611ecbdcef356819e0ee0 [10/05/2022]

Beven, K. J. (2000). Uniqueness of place and process representations in hydrological modelling, *Hydrology and Earth System Sciences*, **4**, 203-213.

Beven, K. J. (2012). Causal models as multiple working hypotheses about environmental processes. *Comptes Rendue Geosciences*, **344**(2), 77-88.

Beven, K. J. (2016). Facets of hydrology—Epistemic error, non-stationarity, likelihood, hypothesis testing, and communication. *Hydrological Sciences Journal*, **61**(9), 1652–1665.

Bhat, S. A. & Pandit, A. K. (2014). Surface Water Quality Assessment of Wular Lake, A Ramsar Site in Kashmir Himalaya, Using Discriminant Analysis and WQI. *Journal of Ecosystems*, **2014**, 1-18.

Bibby, J. S., Douglas, H. A., Thomasson, A. J., & Robertson, J. S. (1991). *Land Capability Classification for Agriculture* [pdf]. https://www.hutton.ac.uk/sites/default/files/files/soils/LAND%20CAPABILITY%20CLASSIFICATION%20FOR%20AGRICULTURE.PDF [22/06/2022].

Birkel, C., Soulsby, C. & Tetzlaff, D. (2014). Integrating parsimonious models of hydrological connectivity and soil biogeochemistry to simulate stream DOC dynamics. *Journal of Geophysical Research: Biogeosciences*, **119**, 1030-1047.

Boorman, D. B., Hollis, J. M. & Lilly, A. (1995). *Hydrology of soil types: a hydrologically-based classification of the soils of the United Kingdom.* Report No. 126. Institute of Hydrology [pdf]. https://docslib.org/doc/5621763/hydrology-of-soil-types-a-hydrologically-based-classification-of-the-soils-of-the-united-kingdom#:~:text=Report%20No.%20126%20Hydrology%20of%20soil%20types%3A%20a,Boorman%2C%20J.M.%20Hollist%20%26%20A.%20Lilly%2A%20November%201995 [10/12/2022].

Borja, Á., Galparsoro, I., Solaun, O., Muxika, I., Tello, E. M., Uriarte, A. & Valencia, V. (2006). The European Water Framework Directive and the DPSIR, a methodological approach to assess the risk of failing to achieve good ecological status. *Estuarine, Coastal and Shelf Science,* **66**, 84-96.

Bouraoui, F. & Grizzetti, B. (2014). Modelling mitigation options to reduce diffuse nitrogen water pollution from agriculture. *Science of the Total Environment*, **468-469**, 1267-1277.

Boyd, C. E. (2015). *Water Quality. An introduction* (2nd Ed.). Springer International Publishing.

Bragg, O. M. (2002). Hydrology of peat-forming wetlands in Scotland. *Science of the Total Environment*, **294**, 111-129.

Breiman, L., Friedman, J. H. , Olshen, R. A. , & Stone, C. G. (1984). *Classification and regression trees.* Wadsworth International Group.

Broadmeadow, S. & Nisbet, T. R. (2004). The effects of riparian forest management on the freshwater environment: a literature review of best management practice. *Hydrology and Earth System Sciences Discussions,* **8**, 286-305.

Brontowiyono, W., Asmara, A. A., Jana, R., Yulianto, A. & Rahmawati, S. (2022). Land-Use Impact on Water Quality of the Opak Sub-Watershed, Yogyakarta, Indonesia. *Sustainability*, **14**(7), 4346.

Brown, I., Towers, W., Rivington, M. & Black, H. I. J. (2008). Influence of climate change on agricultural land-use potential: adapting and updating the land capability system for Scotland. *Climate Research,* 37**,** 43-57.

Brown, I., Poggio, L., Gimona, A., & Castellazzi, M. (2010). Climate change, drought risk and land capability for agriculture: implications for land use in Scotland. *Regional Environmental Change,* **11**, 503-518.

Brown, I., Berry, P., Everard, M., Firbank, L., Harrison, P., Lundy, L., Quine, C., Rowan, J., Wade, R. & Watts, K. (2015). Identifying robust response options to manage environmental change using an Ecosystem Approach: a stress-testing case study for the UK. *Environmental Science & Policy*, **52**, 74-88.

Brown, I. & Everard, M. (2015). A working typology of response options to manage environmental change and their scope for complementarity using an Ecosystem Approach. *Environmental Science and Policy*, **52**, 61-73.

Brown, I. (2017). Hierarchical bioclimate zonation to reference climate change across scales and its implications for nature conservation planning. *Applied Geography*, **85**, 126-138.

Brown, I. (2020). Challenges in delivering climate change policy through land use targets for afforestation and peatland restoration. *Environmental Science and Policy*, **107**, 36-45.

Buckley, C. & Carney, P. (2013). The potential to reduce the risk of diffuse pollution from agriculture while improving economic performance at farm level. E*nvironmental Science and Policy*, **25**, 118-126.

Buffam, I., Galloway, J. N., Blum, L. K. & McGlathery, K. J. (2001). A stormflow/baseflow comparison of dissolved organic matter concentrations and bioavailability in an Appalachian stream. *Biogeochemistry*, **53,** 269–306.

Burt, T. P. & Howden , N. J. K. (2013). North Atlantic Oscillation amplifies orographic precipitation and river flow in upland Britain. *Water Resources Research*, **49**, 3504–3515.

Calder, I. R. (2007). Forests and water—Ensuring forest benefits outweigh water costs. *Forest ecology and management*, **251**, 110-120.

Carluer, N. & De Marsily, G. (2004). Assessment and modelling of the influence of man-made networks on the hydrology of a small watershed: implications for fast flow components, water quality and landscape management. *Journal of Hydrology*, **285**, 76-95.

Cavanaugh, J.E. & Neath, A.A. (2018). The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. WIREs Computational Statistics, 11(3). 1460

CBD (2009). *Connecting Biodiversity and Climate Change Mitigation and Adaptation: Report of the Second Ad Hoc Technical Expert Group on Biodiversity and Climate Change.* Technical Series No. 41 [pdf]. https://www.cbd.int/doc/publications/cbd-ts-41-en.pdf [10/05/2022]

CEH (n. d.). National River Flow Archive. Derived Flow Statistics [html]. http://nrfa.ceh.ac.uk/derived-flow-statistics [02/05/2016]

Challinor, A. J., Adger, W. N., Benton, T. G., Conway, D., Joshi, M. & Frame, D. (2018). Transmission of climate risks across sectors and borders. *Philosophical Transactions of the Royal Society A*, **376**, http://dx.doi.org/10.1098/rsta.2017.0301.

Chang, K., Gao, J. L., Wu, W. Y. & Yuan, X. Y. (2011). Water quality comprehensive evaluation method for large water distribution network based on clustering analysis. *Journal of Hydroinformatics*, **13**(3), 390–400.

Chapman, P. J. & Palmer, S. M. (2016). A review of the factors controlling trends in water colour/dissolved organic carbon. Report to Yorkshire Water Services, Project S3918, University of Leeds, Leeds.

Chapman, P. J., Edwards, A. C. & Cresser, M. S. (2001). The nitrogen composition of streams in upland Scotland: some regional and seasonal differences. *Science of the Total Environment*, **265**, 65-83.

Chausson, A., Turner, B., Seddon, D., Chabaneix, N., Girardin, C. A. J., Kapos, V., Key, I., Roe, D., Smith, A., Woroniecki, S. & Seddon, N. (2020). Mapping the effectiveness of nature-based solutions for climate change adaptation. *Global Change Biology*, **26**, 6134-6155.

Chen, L., Zheng, X., Wang, T & Zhang, J. (2015). Influences of key factors on manganese release from soil of a reservoir shore. *Environmental Science and Pollution Research*, **22**, 11801-11812.

Chen, Q., Mei, K., Dahlgren, R. A., Wang, T., Gong, J. & Zhang, M. (2016). Impacts of land use and population density on seasonal surface water quality using a modified geographically weighted regression. *Science of The Total Environment*, **572**, 450-466.

Chowdhury, U. K., Biswas, B. K., Chowdhury, T. R., Samata, G., Mandal, B. K., Basu, G. C., Chanda, C. R., Lodh, D., Saha, K. C., Mukherjee, S. K., Roy, S., Kabir, S., Quamruzzaman, Q. & Chakraborti, D. (2000). Groundwater arsenic contamination in Bangladesh and West Bengal, India. *Environmental Health Perspectives*, **108**, 393–397.

Clark, J. M., Chapman, P. J., Adamson, J. K. & Lane, S. N. (2005). Influence of drought-induced acidification on the mobility of dissolved organic carbon in peat soils. *Global Change Biology*, **11**(5), 791-809.

Clark, J. M, Lane, .N., Chapman, P. J. & Adamson, J. K. (2007). Export of Dissolved organic carbon from an upland peatland during storm events: Implications for flux estimates. *Journal of Hydrology*, 347, 438-447.

Clark, J. M., Ashley, D., Wagner, M., Chapman, P. J., Lane, S. N., Evans, C. D. & Heathwaite, A. L. (2009). Increased temperature sensitivity of net DOC production from ombrotrophic peat due to water table draw-down. *Global Change Biology*, 15(4), 794-807.

Coffey, R., Cummins, E., O'Flaherty, V. & Cormican, M. (2010). Pathogen Source Estimation and Scenario Analysis Using the Soil and Water Assessment Tool (SWAT). *Human and Ecological Risk Assessment*, **16**, 913-933.

Cole, L., Bardgett, R. D., Ineson, P. & Adamson, J. K. (2002). Relationships between enchytraeid worms (Oligochaeta), climate change, and the release of dissolved organic carbon from blanket peat in northern England. *Soil Biology and Biochemistry*, **34**(5), 599-607.

Conrad, V. (1946). Usual formulas of continentality and their limits of validity. *Eos, Transactions American Geophysical Union*, **27**(5), 663–664.

Cool, G., Lebel, A., Sadiq, R. & Rodriguez, M. J. (2014). Impact of catchment geophysical characteristics and climate on the regional variability of dissolved organic carbon (DOC) in surface water. *Science of the Total Environment*, **490**, 947-956.

Cooper, R., Thoss, V. & Watson, H. (2006). Factors influencing the release of dissolved organic carbon and dissolved forms of nitrogen from a small upland headwater during autumn runoff events. *Hydrological Processes*, **21**(5), 622-633.

Critchlow-Watton, N., Dobbie, K. E., Bell, R., Campbell, S. D. G., Hinze, D., Motion, A., Robertson, K., Russell, M., Simpson, J., Thomson, D. and Towers, W. (Eds.) (2014). *Scotland's State of the Environment Report, 2014.* Scotland's Environment Web [pdf]. https://www.environment.gov.scot/media/1170/state-of-environment-report-2014.pdf [10-05-2022].

Davies, H. & Neal, C. (2004). GIS-based methodologies for assessing nitrate, nitrite, ammonium distributions across a major UK river basin, the Humber. *Hydrology and Earth System Sciences.*, **8**(4), 822-833.

Dawson, J. J. C., Billett, M. F., Neal, C. & Hill, S. (2002). A comparison of particulate, dissolved and gaseous carbon in two contrasting upland streams in the UK. *Journal of Hydrology*, **257**(1–4), 226-246.

Dawson, J. J. C., Soulsby, C., Tetzlaff, D., Hrachowitz, M., Dunn, S. M. & Malcolm, I. A. (2008). Influence of hydrology and seasonality on DOC exports from three contrasting upland catchments. *Biogeochemistry*, **90**, 93–113.

De Brauwere, A., Ouattara, N. K. & Servais, P. (2014). Modelling Faecal Indicator Bacteria Concentrations in Natural Surface Waters: A Review. *Critical Reviews in Environmental Science and Technology*, **44**(21), 2380-2453.

De Wit, H.A., Stoddard, J.L., Monteith, D.T., Sample, J.E., Austnes, K., Couture, S., Fölster, J., Higgins, S.N., Houle, D. & Hruška, J. (2021). Cleaner air reveals growing influence of climate on dissolved organic carbon trends in northern headwaters. *Environmental Research Letters*, **16**, 104009

Del Río-Gamero, B., Ramos-Martín, A., Melián-Martel, N., Pérez-Báez, S. (2020). Water-Energy Nexus: A Pathway of Reaching the Zero Net Carbon in Wastewater Treatment Plants. *Sustainability*, **12**(22), 9377.

Delpla, I., Jung, A. V., Baures, E., Clement, M. & Thomas, O. (2009). Impacts of climate change on surface water quality in relation to drinking water production. *Environment International*, **35,** 1225-1233.

Demšar, D., Džeroski, S., Larsen, T., Struyf, J., Axelsen, J., Pedersen, M. B., & Krogh, P. H. (2006). Using multi-objective classification to model communities of soil microarthropods. *Ecological Modelling*, **191**(1), 131-143.

Depres, C., Grolleau, G. & Mzoughi, N. (2008). Contracting for environmental property rights: the case of Vittel. *Economica*, **75**, 412-434.

Der Kiureghian, A. & Ditlevsen, O. (2017). Aleatory or epistemic? Does it matter? *Structural Safety*, **31**(2), 105-112.

Desai, A., Rifai, H. S., Petersen, T. N., Stein, R. (2011). Mass Balance and Water Quality Modelling for Load Allocation of *Escherichia coli* in an Urban Watershed. *Journal of Water Resources Planning and Management*, **137**(5), 412-427.

Dessai, S., Hulme, M., Lempert, R. & Pielke, R. (2009). Do We Need Better Predictions to Adapt to a Changing Climate? *Eos*, **90**(13), 111-112.

Deutch, J. (2020). Is Net Zero Carbon 2050 Possible? *Joule*, **4**, 2237-2243.

Dick, J. J., Tetzlaff, D., Birkel, C. & Soulsby, C. (2015). Modelling landscape controls on dissolved organic carbon sources and fluxes to streams. *Biogeochemistry*, **122**, 361-374.

Ding, J., Jiang, Y., Liu, Q., Hou, Z., Liao, J., Fu, L., & Peng, Q. (2016). Influences of the land use pattern on water quality in low-order streams of the Dongjiang River basin, China: A multi-scale analysis. *Science of the Total Environment*, 551-552, 205-216.

Donnison, A., Ross, C. & Thorrold, B. (2004). Impact of land use on the faecal microbial quality of hill-country streams. *New Zealand Journal of Marine and Freshwater Research*, **38**(5), 845-855.

Donohue, I., McGarrigle, M. L. & Mills, P. (2006). Linking catchment characteristics and water chemistry with the ecological status of Irish rivers. *Water Research*, **40**(1), 91-98.

Doody, D., Harrington, R., Johnson, M., Hofmann, O. & McEntee, D. (2009). Sewerage treatment in an integrated constructed wetland. *Municipal Engineer,* **162**, 199-205.

Doswald, N., Munroe, R., Roe, D., Giuliani, A., Castelli, I., Stephens, J., Möller, I., Spencer, T., Vitra, B. & Reid, H. (2014). Effectiveness of ecosystem-based approaches for adaptation: review of the evidence-base. *Climate and Development*, **6**(2), 185-201.

Du, X., Zhang, X., Mukundan, R., Hoang, L. & Owens, E. M. (2019). Integrating terrestrial and aquatic processes toward watershed scale modelling of dissolved organic carbon fluxes. *Environmental Pollution*, **249**, 125-135.

Dudley, N. & Stolton, S. (2003). *Running pure: the importance of forest protected areas to drinking water* [pdf]. assets.panda.org/downloads/runningpurereport.pdf [10/05/2022]

Duhaime, K. & Roberts, D. (2018). Theoretical implications of best management practices for reducing the risk of drinking water contamination with *Cryptosporidium* from grazing cattle. *Agriculture, Ecosystems & Environment*, **259**, 184-193.

Dunn, S. M., Sample, J., Potts, J., Abel, C., Cook, Y., Taylor, C. & Vinten, A. J. A. (2014). Recent trends in water quality in an agricultural catchment in Eastern Scotland: elucidating the roles of hydrology and land use. *Environmental Science: Processes & Impacts,* **16**, 1659, DOI: 10.1039/c3em00698k.

Dunn, S. M., Towers, W., Dawson, J. J. C., Sample, J. & McDonald, J. (2015). A pragmatic methodology for horizon scanning of water quality linked to future climate and land use scenarios. *Land Use Policy*, **44**, 131-144.

DWQR (2021a). *Drinking Water Quality in Scotland 2020 – Public Water Supply* [pdf]. https://www.dwqr.scot/media/hita5vuz/annual-report-public-supplies-main-report-2020.pdf [10/05/2022]

DWQR (2021b). *Drinking Water Quality in Scotland 2020 – Private Water Supplies* [pdf]. https://www.dwqr.scot/media/2q4etvrh/pws-annual-report-2020-final.pdf [10/05/2022]

Edwards, K. J. & Ralston, I. B. M. (1997). *Scotland: environment and archaeology, 8000 BC-AD 1000*. Wiley.

Evans, C. D., Monteith, D. T., & Cooper, D. M. (2005). Long-term increases in surface water dissolved organic carbon: Observations, possible causes and environmental impacts. *Environmental Pollution,* **137**, 55-71.

Evans, C. D., Chapman, P. J., Clark, J. M., Monteith, D. T. & Cresser, M. S. (2006). Alternative explanations for rising dissolved organic carbon export from organic soils. *Global Change Biology*, **12**(11), 2044-2053.

Evans, C., Freeman, C., Cork, L.G., Thomas, D.N., Reynolds, B., Norris, D. & Garnett, M.H. (2006). *DOC export to upland waters: old or new carbon?* [pdf]. https://www.researchgate.net/publication/258499947_DOC_export_to_upland_waters_old_or_new_carbon [22/06/2022].

Evans, C. D., Jones, T.G ., Burden, A., Ostle, N., Zielinski, P., Cooper, M. D. A., Peacock, M., Clark, J.M., Oulehle, F., Cooper, D. & Freeman, C. (2012). Acidity controls on dissolved organic carbon mobility in organic soils. *Global Change Biology*, **18**(11), 3317-333.

Evans, M. & Warburton, J. (2010a). *Geomorphology of Upland Peat*. Blackwell Publishing Ltd.

Evans, M.G. & Warburton, J. (2010b). Peatland Geomorphology and Carbon Cycling. *Geography Compass*, **4**(10), 1513-1531.

Everard, M. (2013). Safeguarding the provision of ecosystem services in catchment systems. *Integrated environmental assessment and management*, **9**, 252-259.

Everard, M. & McInnes, R. (2013). Systemic solutions for multi-benefit water and environmental management. *Science of the Total Environment*, **461-462**, 170-179.

Eze, S., Palmer, S. M. & Chapman, P. J. (2018). Soil organic carbon stock in grasslands: Effects of inorganic fertilizers, liming and grazing in different climate settings. *Journal of Environmental Management*, **223**, 74-84.

Fawell, J. & Nieuwenhuijsen, M. J. (2003). Contaminants in drinking water. Environmental pollution and health. *British Medical Bulletin,* **68**, 199-208.

Ferrier, R. C., Edwards, A. C., Hirst, D., Littlewood, I. G., Watts, C. D. & Morris, R. (2001). Water quality of Scottish rivers: spatial and temporal trends. *Science of the Total Environment*, **265**, 327-342.

Field, A., Miles, J & Field, Z. (2012). *Discovering Statistics Using R*. Sages Publications Ltd.

Fisher, J. & Acreman, M. C. (2004). Wetland nutrient removal: a review of the evidence. *Hydrology and Earth System Sciences Discussions,* **8**, 673-685.

Forestry Commission (2017*). The UK Forestry Standard. The governments' approach to sustainable forestry.* 4th edition [pdf]. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/687147/The_UK_Forestry_Standard.pdf [29/11/2022].

Freeman, C., Evans, C.D., Monteith, D. T., Reynolds, B. & Fenner, N. (2001). Export of organic carbon from peat soils. *Nature,* **412**, 785.

Fu, B., Merritt, W. S., Croke, B. F. W., Weber, T. R., & Jakeman, A. J. (2019). A review of catchment-scale water quality and erosion models and a synthesis of future prospects. *Environmental Modelling & Software,* **114**, 75-97.

Futter, M. N., Butterfield, D., Cosby, B. J., Dillon, P. J., Wade, A. J. & Whitehead, P. G. (2007). Modelling the mechanisms that control in-stream dissolved organic carbon dynamics in upland and forested catchments. *Water Resources Research*, **43**, W02424.

Futter, M. N. & de Wit, H. A. (2008). Testing seasonal and long-term controls of streamwater DOC using empirical and process-based models. *Science of the Total Environment*, **407**, 698-707.

Gadiwala, M. S., Burke. F., Alam, M. T., Nawaz-ul-Huda, S., Azam, M. (2013). Oceanity and continentality climate indices in Pakistan. *Malaysian Journal of Society and Space*, **9**(4), 57-66.

Gärtner, N., Lindhe, A., Wahtra, J., Söderquist, T., Lång, L.-O., Nordzell, H., Norrman, J. & Rosén, L. (2022). Integrating Ecosystem Services into Risk Assessments for Drinking Water Protection. *Water*, **14**(8), 1180, https://doi.org/10.3390/w14081180.

Gallopin, G. C. (2006). Linkages between vulnerability, resilience, and adaptive capacity. *Global Environmental Change*, **16**, 293–303.

García, R. R., Fraser, M. D., Celaya, R., Ferreira, L. M. M., García, U. & Osoro, K. (2012). Grazing land management and biodiversity in the Atlantic European heathlands: a review. *Agroforestry Systems*, **87**, 19-43.

Garnier, M. & Holman, I. (2019). Critical Review of Adaptation Measures to Reduce the Vulnerability of European Drinking Water Resources to the Pressures of Climate Change. *Environmental Management*, **64**, 138–153 .

Geissen, V., Mol, H., Klumpp, E., Umlauf, G., Nadal, M., van der Ploeg, M., van de Zee, S. E. and Ritsema, C. J. (2015). Emerging pollutants in the environment: a challenge for water resource management. *International soil and water conservation research*, **3**(1), 57-65.

Ghaffari, G., Keestra, S., Ghodousi, J. & Ahmadi, H. (2009). SWAT-simulated hydrological impact of land-use change in the Zanjanrood basin, Northwest Iran. *Hydrological Processes*, **24**(7), 892-903.

Giao, N. T. & Nhien, H. T. H (2021). Evaluating Water Quality Variation in the Vietnamese Mekong Delta Area Using Cluster and Discriminant Analysis. *Applied Environmental Research*, **43**(1), 14-27.

Giffoni, E., Jude, S., Smith, H. M. & Pollard, S. J. T. (2022). Real-life resilience: Exploring the organisational environment of international water utilities. *Utilities Policy*, **79**, https://doi.org/10.1016/j.jup.2022.101394.

Gillen, Con (2013. *The Geology and Landscapes of Scotland* (2nd ed.). Dunedin Academic Press Ltd.

Gillespie, M. R., Crane, E. J. & Barron, H. F. (2013). *Deep geothermal energy potential in Scotland.* British Geological Survey Commissioned Report, CR/12/131 [pdf]. http://www.gov.scot/Resource/0043/00437996.pdf [10/05/2022]

Gilvear, D. J., Heal, K. V. & Stephen, A. (2002). Hydrology and the ecological quality of Scottish river ecosystems. *Science of the Total Environment*, **294**, 131-159.

Glendell, M. & Brazier, R. E. (2014). Accelerated export of sediment and carbon from a landscape under intensive agriculture. *Science of the Total Environment*, **1**, 476-477.

Gnauck, A. (2004). Interpolation and approximation of water quality time series and process identification. *Analytical and Bioanalytical Chemistry*, **380**, 484–492.

Gooday, R. D., Anthony, S. G., Chadwick, D. R., Newell-Price, P., Harris, D., Duethmann, D., Fish, R., Collins, A. L. & Winter, M. (2014). Modelling the cost-effectiveness of mitigation methods for multiple pollutants at farm scale. *Science of the Total Environment*, **468-469**, 1198-1209.

Gormley, A., Pollard, S. & Rocks, S. (2011). *Guidelines for Environmental Risk Assessment and Management – Green Leaves III* [pdf]. https://www.gov.uk/government/publications/guidelines-for-environmental-risk-assessment-and-management-green-leaves-iii?msclkid=84da0a07d05a11ec8e8a25e5fbee3e62 [10/05/2022]

Gosling (2014). Assessing the impact of projected climate change on drought vulnerability in Scotland. *Hydrology Research*, **45**(6), 806-816.

Graffelman, J. & Aluja-Banet, T. (2003). Optimal Representation of Supplementary Variables in Biplots from Principal Component Analysis and Correspondence Analysis. *Biometrical Journal*, **45**(4), 491-509.

Graham, M. C., Gavon, K. G., Kirika, A. & Farmer, J. G. (2012). Processes controlling manganese distributions and associations in organic-rich freshwater aquatic systems: The example of Loch Bradan, Scotland. *Science of the Total Environment*, **424**, 239-250.

Grand-Clement, E., Anderson, K., Smith, D., Luscombe, D., Gatis, N., Ross, M. & Brazier, R. E. (2013). Evaluating ecosystem goods and services after restoration of marginal upland peatlands in South-West England. *Journal of Applied Ecology*, **50**, 324-334.

Grayston, S. J., Griffith, G. S., Mawdsley, J. L., Campbell, C. D. & Bardgett, R. D. (2001). Accounting for variability in soil microbial communities of temperate upland grassland ecosystems. *Soil Biology and Biochemistry*, **33**(4-5), 533-551.

Grizzetti, B., Lanzanova, D., Liquete, C., Reynaud, A. & Cardoso, A. C. (2016). Assessing water ecosystem services for water resource management. *Environmental Science and Policy*, **61**, 194-203.

Gustard, A., Bullock, A. & Dixon, J. M. (1992). *Low flow estimation in the United Kingdom.* Report No. 108. Institute of Hydrology. https://doi.org/10.1002/esp.3290190707

Haasnoot, M., Kwakkel, J. H., Walker, W. E. & ter Maat, J. (2013). Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world. *Global Environmental Change*, **23**, 485–498.

Haggarty, R. A., Miller, C. A., Scott, E. M., Wyllie, F. & Smith, M. (2012). Functional clustering of water quality data in Scotland. *Environmetrics*, **23**(8), 685-695.

Hallegatte, S. (2009). Strategies to adapt to an uncertain climate change. *Global Environmental Change*, **19**, 240-247.

Halliday, S. J., Wade, A. J., Skeffington, R. A., Neal, C., Reynolds, B., Rowland, P., Neal, M., Norris, D. (2012). An analysis of long-term trends, seasonality and short-term dynamics in water quality data from Plynlimon, Wales. *Science of the Total Environment*, **434**, 186-200.

Hamid, F. S., Bhatti, M. S., Anuar, N., Anuar, N., Mohan, P. & Periathamby, A. (2018). Worldwide distribution and abundance of microplastic: How dire is the situation? *Waste Management & Research*, **36**(10), 873–897.

Hamilton-Taylor, J., Davison, W. & Morfett, K. (1996). The biogeochemical cycling of Zn, Cu, Fe, Mn, and dissolved organic carbon in a seasonally anoxic lake. *Limnology and Oceanography*, **41**(3), 408-418.

Hanlon, H., Bernie, D., Carigi, G., Lowe, J. (2021). Future Changes to high impact weather in the UK, *Climatic Change*, **166**(50), https://doi.org/10.1007/s10584-021-03100-5.

Harmel, R. D., Karthikeyan, R., Gentry, T. & Srinivasan, R. (2010). Effects of Agricultural Management, Land Use, and Watershed Scale on *E. coli* Concentrations in Runoff and Streamflow. *American Society of Agricultural and Biological Engineers*, **53**(6), 1833-1841.

Hayes, N. M., Deemer, B. R., Corman, J. R., Razavi, N. R., Strock, K. E. (2017). Key differences between lakes and reservoirs modify climate signals: A case for a new conceptual model. *Limnology and Oceanography Letters*, **2**, 47-62.

Heathwaite, A. L., Göttlich, K., Burmeister, E. G., Kaule, G. & Grospiethsch, T. (1993). Mires: definitions and form. In A.L. Heathwaite & K. Göttlich (Eds.): *Mires: process, exploitation and conservation* (pp. 1-75). Wiley.

Hefting, M. M., Van den Heuvel, R. N. & Verhoeven, J. T. A. (2013). Wetlands in agricultural landscapes for nitrogen attenuation and biodiversity enhancement: Opportunities and limitations. *Ecological engineering,* **56**, 5-13.

Helliwell, R. C., Coull, M. C., Davies, J. J. L., Evans, C. D., Norris, D., Ferrier, R. C., Jenkins, A. & Reynolds, B. (2007). The role of catchment characteristics in determining surface water nitrogen in four upland regions in the UK. *Hydrology and Earth System Sciences,* **11**(1), 356-371.

Henderson, R. K., Baker, A., Parons, S. A. & Jefferson, B. (2008). Characterisation of algogenic organic matter extracted from cyanobacteria, green algae and diatoms. *Water Research*, **42**(13), 3435-3445

Herngren, L., Goonetilleke, A. & Ayoko, G. A. (2005). Understanding heavy metal and suspended solids relationships in urban stormwater using simulated rainfall. *Journal of Environmental Management,* **76**, 149-158.

Hevesi, J. A., Flint, L. E., Church, C. D. & Mendez, G. O. (2011). *Using a watershed model (HSPF) to evaluate sources and transport of pathogen indicator bacteria in the Chino Basin, San Bernardino County, California.* U.S. Geological Survey Scientific Investigations Report 2009–5219 [pdf]. https://pubs.usgs.gov/sir/2009/5219/pdf/sir20095219.pdf [16/12/2021]

Holden, J. (2005). Peatland hydrology and carbon release: why small-scale process matters. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* **363**, 2891-2913.

Holden, J., Chapman, P. J., Palmer, S. M., Kay, P. & Grayson, R. (2012). The impacts of prescribed moorland burning on water colour and dissolved organic carbon: A critical synthesis. *Journal of Environmental Management,* **101**, 92-103.

Homoncik, S. C., MacDonald, A. M., Heal, K. V., Ó Dochartaigh, B. E., Ngwenya, B. T. (2010). Manganese concentrations in Scottish groundwater. *Science of the Total Environment*, **408**, 2467–2473.

Hood, E., Gooseff, M. N., Johnson, S. L. (2006). Changes in the character of stream water dissolved organic carbon during flushing in three small watersheds, Oregon. *Journal of Geophysical Research*, **111**, G01007, DOI:10.1029/2005JG000082.

Hooda, P. S., Edwards, A. C., Anderson, H. A. & Miller, A. (2000). A review of water quality concerns in livestock farming areas. *Science of the Total Environment*, **250**, 143-167.

Howard, G., Bartram, J., Pedley, S., Schmoll, O., Chorus, I. & Berger, P. (2006). Groundwater and Public Health. In O. Schmoll, G. Howard, J. Chilton & I. Chorus, (eds). *Protecting Groundwater for Health* (pp. 3-19). IWA Publishing.

Huang, X., Tang, G., Zhu, T., Ding, H., Na, J. (2019). Space-for-time substitution in geomorphology: A critical review and conceptual framework. *Journal of Geographical Sciences*, **29**(10), 1670-1680.

Hunter, P. R. (2003). Drinking water and diarrhoeal disease due to Escherichia coli. *Journal of Water & Health*, **1** (2), 65-72.

Iacob, O., Rowan, J. S., Brown, I & Ellis, C. (2014). Evaluating wider benefits of natural flood management strategies: an ecosystem-based adaptation perspective. *Hydrology Research*, **45**(6), 774-787.

Im, S. Brannan, K., Mostaghimi, S. & Cho, J. (2003). *A Comparison of SWAT and HSPF Models for Simulating Hydrologic and Water Quality Responses from an Urbanizing Watershed.* 2003 ASAE Annual International Meeting, Las Vegas, Nevada, USA, 27- 30 July, Paper Number: 032175 [pdf]. https://swat.tamu.edu/media/1303/asae032175.pdf [10/12/2022].

IPCC (2014). Annex II: Glossary. In R. K. Pachauri & L.A. Meyer (Eds.). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 117-130) [pdf]. https://www.ipcc.ch/report/ar5/syr/ [10/05/2022].

IPCC (2021). Summary for Policymakers. In V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Chaud, Y. Chen., L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu & B. Zhou (Eds.). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 3-32). Cambridge University Press

[pdf]. https://www.ipcc.ch/report/sixth-assessment-report-working-group-i/ [10/05/2022].

IPCC (2022). Summary for Policymakers. In H.-O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (Eds.). *Climate Change 2022: Impacts, Adaptation and Vulnerability. Working Group II contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 3-33). Cambridge University Press [pdf]. https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii/ [10/05/2022].

IWA (2004). *The Bonn Charter for Safe Drinking Water – September 2004* [pdf]. https://iwa-network.org/wp-content/uploads/2016/06/Bonn-Charter-for-Safe-Drinking-Water.pdf [10/05/2022].

Johnson, R. C. & Thompson, D. B. (2002). Hydrology and the natural heritage of the Scottish mountains. *Science of the Total Environment,* **294**, 161-168.

Jurgilevich, A., Räsänen, A., Groundstroem, F. & Juhola, S. (2017). A systematic review of dynamics in climate risk and vulnerability assessments. *Environmental Research Letters*, **12**, 013002, doi:10.1088/1748-9326/aa5508.

Kanamori, H., Weber, D.J., Rutala, W.A. (2016). Healthcare Outbreaks Associated With a Water Reservoir and Infection Prevention Strategies. *Healthcare Epidemiology*, **62**(11), 1423-1435.

Kaufman, L. & Rousseeuw, P.J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis.* Wiley.

Kauffman, C. M. (2014). Financing watershed conservation: Lessons from Ecuador's evolving water trust funds. *Agricultural Water Management,* **145**, 39-49.

Kay, D., Wyer, M., Crowther, J., Stapleton, C., Bradford, M., McDonald, A., Greaves, J., Francis, C. & Watkins, J. (2005). Predicting faecal indicator fluxes using digital land use data in the UK’s sentinel Water Framework Directive catchment: The Ribble study. *Water Research*, **39**, 3967–3981.

Kay, D., Anthony, S., Crowther, J., Chambers, B. J., Nicholson, F. A., Chadwick, D., Stapleton, C. M. & Wyer, M. D. (2010). Microbial water pollution: A screening tool for initial catchment-scale assessment and source apportionment. *Science of the Total Environment*, **408**, 5649–5656.

Kay, P., Edwards, A. C. & Foulger, M. (2009). A review of the efficacy of contemporary agricultural stewardship measures for ameliorating water pollution problems of key concern to the UK water industry. *Agricultural Systems*, **99**, 67-75.

Kazi, T. G., Arain, M. B., Jamali, M. K., Jalbini, N., Afridi, H. I., Sarfraz, R. A., Baig, J. A. & Shah, A. Q. (2009). Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. *Ecotoxicology and Environmental Safety*, **72**(2), 301-309.

Keeler, B. L., Polasky, S., Brauman, K. A., Johnson, K. A., Finlay, J. C., O'Neill, A., Kovacs, K., Dalzell, B. (2012). Linking water quality and well-being for improved assessment and valuation of ecosystem services. *PNAS*, **109**(45), 18619–18624.

Khan, S. J., Deere, D., Leusch, F. D. L., Humpage, A., Jenkins, M., Cunliffe, D. (2015). Extreme weather events: Should drinking water quality management systems adapt to changing risk profiles? *Water Research*, **85**, 124-136.

Kim, M, Boithias, L., Cho, K. H., Sengtaheuanghoung, O. & Ribolzi, O. (2018). Modelling the Impact of Land Use Change on Basin-scale Transfer of Faecal Indicator Bacteria: SWAT Model Performance. *Journal of Environmental Quality*, **47**(5), 1115-1122.

Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research*, **42**, W03S04, DOI:10.1029/2005WR004362.

Kiros, G. Shetty, A. & Nandagiri, L. (2015). Performance Evaluation of SWAT Model for Land Use and Land Cover Changes in Semi-Arid Climatic Conditions: A Review. *Hydrology: Current Research*, **6**(3), 1000216, http://dx.doi.org/10.4172/2157-7587.1000216.

Kodinariya, T. M. & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal of Advance Research in Computer Science and Management Studies*, **1**(6), 90-95.

Köhler, S. J., Buffam, I., Seibert, J., Bishop, K. H. &Laudon, H. (2009). Dynamics of stream water DOC concentrations in a boreal headwater catchment: Controlling factors and implications for climate scenarios. *Journal of Hydrology*, **373**, 44-56.

Köhler, S. J., Kothawala, D., Futter, M.N., Liungman, O., Tranvik, M. (2013). In-Lake Processes Offset Increased Terrestrial Inputs of Dissolved Organic Carbon and Colour to Lakes. *PLoS ONE*, **8**(8): e70598, doi:10.1371/journal.pone.0070598.

Kousky, C. (2015). New York City's Watershed Agricultural Program. In M. Lago, J. Mysiak, M.C. Gómez, G. Delacámara & A. Maziotis (Eds.). *Use of Economic Instruments in Water Policy: Insights from International Experience* (pp. 351-363). Springer International Publishing.

Kowalkowski, T., Zbytniewski, R., Szpejna, J. & Buszewskia, B. (2006). Application of chemometrics in river water classification. *Water Research*, **40**, 744– 752.

Kritzberg, E. S. & Ekström, S. M. (2012). Increasing iron concentrations in surface waters – a factor behind brownification? *Biogeosciences*, **9**, 1465-1478.

Kumar, V. (2011). Elongation ratio. In V. P. Singh, P. Singh, U. K. Haritashya (Eds.). *Encyclopedia of Snow, Ice and Glaciers* (p. 257). Encyclopedia of Earth Sciences Series. Springer.

Kundzewicz, K. W. (2018). *Quo vadis*, hydrology?. *Hydrological Sciences Journal*, **63**(8), 1118-1132.

Langan, S. J. & Soulsby, C. (2001). The environmental context for water quality variation in Scotland. *Science of the Total Environment*, **265**, 7-14.

Lempert, R. J. & Collins, M. T. (2007). Managing the Risk of Uncertain Threshold Responses: Comparison of Robust, Optimum, and Precautionary Approaches. *Risk Analysis*, **27**(4), 1009-1026.

Lester, R. E., Close, P. G., Barton, J. L., Pope, A. J. & Brown, S. C. (2014). Predicting the likely response of data-poor ecosystems to climate change using space-for-time substitution across domains. *Global Change Biology*, **20**(11), 3471-3481.

Li, P., Holden, J. & Irvine, B. (2016). Prediction of blanket peat erosion across Great Britain under environmental change. *Climatic Change*, **134**(1-2), 177-191.

Li, S., Gu, S., Tan, X. & Zhang, Q. (2009). Water quality in the upper Han River basin, China: The impacts of land use/land cover in riparian buffer zone. *Journal of Hazardous Materials*, **165**, 317–324.

Licence, K., Oates, K. R., Synge, B. A. & Reid, T. M. (2001). An outbreak of *E. coli* O157 infection with evidence of spread from animals to man through contamination of a private water supply. *Epidemiology & Infection*, **126** (1), 135-138.

Lilly, A., Baggaley, N. & Donnelly, D. (2012). *Map of soil organic carbon in top soils of Scotland.* Map prepared for EU project GS-SOIL - Assessment and strategic development of INSPIRE compliant Geodata-Services for European Soil Data. ECP-2008-GEO-31800 [html]. http://ukso.org/static-maps/soils-of-scotland.html [03/12/2022].

Liu, R., Zhang, P., Wang, X., Chen, Y. & Shen, Z. (2013). Assessment of effects of best management practices on agricultural non-point source pollution in Xiangxi River watershed. *Agricultural Water Management,* **117**, 9-18.

Lowe, J. A., Bernie, D., Bett, P., Bricheno, L., Brown, S., Calvert, D., Clark, R., Eagle, K., Edwards, T., Fosser, G., Fung, F., Gohar, L., Good, P., Gregory, J., Harris, G., Howard, T., Kaye, N., Kendon, E., Krijnen, J., Maisey, P., McDonald, R., McInnes, R., McSweeney, C., Mitchell, J. F. B. , Murphy, J., Palmer, M., Roberts, C., Rostron, J., Sexton, D., Thornton, H., Tinker, J., Tucker, S., Yamazaki, K., & Belcher, S. (2018). *UKCP18 Science Overview report.* Met Office [pdf]. https://www.metoffice.gov.uk/pub/data/weather/uk/ukcp18/science-reports/UKCP18-Overview-report.pdf [16/12/2018].

Luís, A., Lickorish, F. & Pollard, S. (2016). Evolution of strategic risks under future scenarios for improved utility master plans. *Water Research*, **88**, 719-727.

Ma, J.-Y., Li, M.-J., Qi, Z.-Z., Fu, M., Sun, T.-F., Elsheikha, H. M. & Cong, W. (2022). Waterborne protozoan outbreaks: An update on the global, regional, and national prevalence from 2017 to 2020 and sources of contamination. *Science of The Total Environment*, **806**(2), 150562, https://doi.org/10.1016/j.scitotenv.2021.150562

Maier, H. R., Guillaume, J. H. A., van Delden, H., Riddell, G. A., Haasnot, M. & Kwakkel, J. H. (2016). An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together? *Environmental Modelling & Software*, **81**, 154-164.

Mandel, P., Maurel, M. & Chenu, D. (2015). Better understanding of water quality evolution in water distribution networks using data clustering. *Water Research*, **87**, 69-78.

Mander, Ü., Haykawa, Y. & Kuusemets, V. (2005). Purification processes, ecological functions, planning and design of riparian buffer zones in agricultural watersheds. *Ecological Engineering*, **24**, 421-432.

Martin-Ortega, J., Allott, T. E. H., Glenk, K. & Schaafsma, M. (2014). Valuing water quality improvements from peatland restoration: Evidence and challenges. *Ecosystem Services*, **9**, 34-43.

MacDonald, A. M., Ó Dochartaigh, B. É. & Smedley, P. L. (2017). Baseline groundwater chemistry in Scotland's aquifers. British Geological Survey Open Report, OR/17/030 [pdf]. https://nora.nerc.ac.uk/id/eprint/519084/1/OR17030.pdf [17/03/2013]

McDonald, C. P. & Urban, N. R. (2010). Using a model selection criterion to identify appropriate complexity in aquatic biogeochemical models. *Ecological Modelling*, **221**(3), 428-432.

McEvoy, D., Fünfgeld, H. & Bosomworth, K. (2013). Resilience and Climate Change Adaptation: The Importance of Framing. *Planning Practice & Research*, **28**(3), 280-293.

MacGillivray, B. H., Hamilton, P. D., Strutt, J. E. & Pollard, S. J. T. (2006). Risk analysis strategies in the water utility sector: an inventory of applications for better and more credible decision-making. *Critical Reviews in Environmental Science and Technology*, **36**(2), 85-139.

McGrane, S. J., Tetzlaff, D. & Soulsby, C. (2014). Application of a linear regression model to assess the influence of urbanised areas and grazing pastures on the microbiological quality of rural streams. *Environmental Monitoring and Assessment*, **186**, 7141–7155.

Meerhoff, M., Teixeira-de Mello, F., Kruk, C., Alonso, C., Gonzalez-Bergonzoni, I., Pacheco, J. P., Lacerot, G., Arim, M., Beklioğu, M., Brucet, S., Goyenol, G., Iglesias, C., Mazzeo, N., Kosten, S & Jeppesen, E. (2012). Environmental Warming in Shallow Lakes: A Review of Potential Changes in Community Structure as Evidenced from Space-for-Time Substitution Approaches. *Advances in Ecological Research*, **46**, 259-349.

Meshesha, T. W., Wang, J. & Melaku, N. D. (2020). Modelling spatiotemporal patterns of water quality and its impacts on aquatic ecosystem in the cold climate region of

Alberta, Canada. *Journal of Hydrology*, **587**, 124925, https://doi.org/10.1016/j.jhydrol.2020.124952

Met Office (2019). *UKCP18 Science Overview. Executive Summary. January 2019* [pdf]. https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/research/ukcp/ukcp18-overview-summary.pdf [10/05/2022].

Milledge, D. G., Lane, S. N., Heathwaite, A. L. & Reaney, S. M. (2012). A Monte Carlo approach to the inverse problem of diffuse pollution risk in agricultural catchments. *Science of The Total Environment*, **433**, 434-449.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P. & Stouffer, R. J. (2008). Stationarity is dead: Whither water management? *Science*, **319**, 573–574.

Monteith, D., Stoddard, J. L., Evans, C. D., De Wit, H. A., Forsius, M., Høgåsen, T., Wilander, A., Skjelkvåle, B. L., Jeffries, D. S., Vuorenmaa, J., Keller, B., Kopácek, J. & Vesely, J. (2007). Dissolved organic carbon trends resulting from changes in atmospheric deposition chemistry. *Nature*, **450**, 537-541.

Monteith, D., Henrys, P. A., Evans, C. D., Malcolm, I., Shilland, E. M. & Pereira, M. G. (2015). Spatial controls on dissolved organic carbon in upland waters inferred from a simple statistical model. *Biogeochemistry*, **123**(3), 363-377.

Monteith, D., Pickard, A.E., Spears, B.M. & Feuchtmayr, H. (2021). An introduction to the FREEDOM-BCCR project. FREEDOM-BCCR briefing note 1 to the water industry. UKRI SPF UK Climate Resilience programme – Project no. NE/S016937/2 [pdf]. https://www.ceh.ac.uk/sites/default/files/2021-09/FREEDOM_BCCR_Project_Introduction_01.pdf [27/10/2021].

Moody, C. S., Worrall, F., Evans, C. D. & Jones, T. G. (2013). The rate of loss of dissolved organic carbon (DOC) through a catchment. *Journal of Hydrology*, **492**, 139-150.

Morecroft, M. D., Crick, H. Q. P., Duffield, S. J. & Macgregor, N. A. (2012). Resilience to climate change: translating principles into practice. *Journal of Applied Ecology*, **49**, 547–551.

Morris, S. & Holstead, K. L. (2013). *Review of the economics of sustainable land management measures in drinking water catchments.* CREW report CD2012/34 [pdf].

https://www.broads-authority.gov.uk/looking-after/managing-land-and-water/conservation-publications-and-reports/water-conservation-reports/2.-Sustainable-Land-Management.pdf?msclkid=03181873d04611ecb9f6444098445447 [10/05/2022].

Moxey, A. & Moran, D. (2014). UK peatland restoration: Some economic arithmetic. *Science of the Total Environment*, **484**, 114-120.

Moxley, J. (2014). *Trends in Organic Carbon in Scottish Rivers and Lochs* [pdf]. https://www.researchgate.net/publication/237019414_Trends_in_Organic_Carbon_in_Scottish_Rivers_and_Lochs [22/06/2022].

Moyer, D. L. & Hyer, K. E. (2003). *Use of the Hydrological Simulation Program-FORTRAN and bacterial source tracking for development of the faecal coliform total maximum daily load (TMDL) for Christians Creek, Augusta County, Virginia.* Water-Resources Investigations Report [pdf]. https://pubs.usgs.gov/wri/wri034162/wrir03-4162.pdf [17/08/2022].

Munang, R., Thiaw, I., Alverson, K., Mumba, M., Liu, J. & Rivington, M. (2013). Climate change and Ecosystem-based Adaptation: a new pragmatic approach to buffering climate change impacts. *Current Opinion in Environmental Sustainability*, **5**, 67–71.

Muscatelli, A. McKee, E. & McGivern, S. (2020). Scotland: a world-leading Hydro Nation. *International Journal of Water Resources Development*, **36**(2-3), 239-244.

Musolff, A., Fleckenstein, J. H., Opitz, M., Büttner, O., Kumar, R. & Tittel, J. (2018). Spatio-temporal controls of dissolved organic carbon stream water concentrations. *Journal of Hydrology*, **566**, 205-215.

Nasr, A., Bruen, M., Jordan, P., Moles, R., Kiely, G. & Byrne, P. (2007). A comparison of SWAT, HSPF and SHETRAN/GOPC for modelling phosphorus export from three catchments in Ireland. *Water Research*, **41**, 1065– 1073.

Neal, C., Lofts, S., Evans, C. D., Reynolds, B., Tipping, E. , Neal, M. (2008). Increasing Iron Concentrations in UK Upland Waters. *Aquatic Geochemistry,* **14**, 263-288.

Neal, C., Rowland, P. Neal, M., Jarvie, H. P., Lawlor, A., Sleep, D., Scholefield, P. (2011). Aluminium in UK rivers: a need for integrated research related to kinetic factors, colloidal transport, carbon and habitat. *Journal of Environmental Monitoring*, **13**, 2153-2164.

Neill, A. J., Tetzlaff, D., Strachan, N. , J., C., Hough, R. L., Avery, L. M., Watson, H., & Soulsby, C. (2018). Using spatial-stream-network models and long-term data to understand and predict dynamics of faecal contamination in a mixed land-use catchment. *Science of the Total Environment*, **612**, 840-852.

Newell-Price, J. P., Harris, D., Taylor, M., Williams, J. R., Anthony, S. G., Duethmann, D., Gooday, R. D., Lord, E. I., Chambers, B. J., Chadwick, D. R. & Misselbrook, T. H. (2011). *An Inventory of Mitigation Methods and Guide to their Effects on Diffuse Water Pollution, Greenhouse Gas Emissions and Ammonia Emissions from Agriculture.* User Guide [pdf]. https://www.researchgate.net/publication/268200287_MITIGATION_METHODS_-_USER_GUIDE_An_Inventory_of_Mitigation_Methods_and_Guide_to_their_Effects_on_Diffuse_Water_Pollution_Greenhouse_Gas_Emissions_and_Ammonia_Emissions_from_Agriculture [13/09/2022].

Nikoloski, S., Kocev, D., Levtić, J., Wall, D. P. & Džeroski, S. (2021). Exploiting partially-labelled data in learning predictive clustering trees for multi-target regression: A case study of water quality assessment in Ireland. *Ecological Informatics*, **61**, 101161, https://doi.org/10.1016/j.ecoinf.2020.101161

Nisbet, T., Silgram, M., Shah, N., Morrow, K. & Broadmeadow, S. (2011). *Woodland for water: woodland measures for meeting water framework directive objectives* [pdf]. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/291522/scho0711btyr-e-e.pdf [13/09/2022].

NYC DEP (2022). *New York City Drinking Water Supply and Quality Report 2021* [pdf]. https://www1.nyc.gov/assets/dep/downloads/pdf/water/drinking-water/drinking-water-supply-quality-report/2021-drinking-water-supply-quality-report.pdf [13/09/2022].

Odonkor, S. T. & Ampofo, J. K. (2013). *Escherichia coli* as an indicator of bacteriological quality of water: an overview. *Microbiology Research*, **4**(1), 5-11.

OFWAT (n. d.). *From catchment to customer. Can upstream catchment management deliver a better deal for water customers and the environment?* [pdf] https://www.ofwat.gov.uk/wp-content/uploads/2015/11/prs_inf_catchment.pdf [13/09/2022].

Oliver, D. M., Heathwaite, A. L., Fish, R. D., Chadwick, D. R., Hodgson, C. J., Winter, M. & Butler, A. J. (2009). Scale appropriate modelling of diffuse microbial pollution from agriculture. *Progress in Physical Geography*, **33**(3), 358–377.

Oliver, D. M., Porter, K. D. H., Pacheosky, Y. A., Muirhead, R. W., Reaney, S. M., Coffey, R., Kay, D., Milledge, D. G., Hong, E., Anthony, S. G., Page, T., Bloodworth, J. W., Mellander, P. Carbonneau, P. E., McGrane, S. J. & Quilliam, R. S. (2016). Predicting microbial water quality with models: Over-arching questions for managing risk in agricultural catchments. *Science of the Total Environment*, **544**, 39–47.

Oni, S. K., Futter, M. N., Molot, L. A. & Dillon, P. J. (2012). Modelling the long term impact of climate change on the carbon budget of Lake Simcoe, Ontario using INCA-C. *Science of the Total Environment*, **414**, 387-403.

Ou, Y & Wang, X. (2011). GIS and ordination techniques for studying influence of watershed characteristics on river water quality. *Water Science and Technology*, **64**(4), 861-870.

Ouyang, Y. (2005). Evaluation of river water quality monitoring stations by principal component analysis. *Water Research*, **39**, 2621–2635.

Pärn, J. & Mander, U. (2012). Increased organic carbon concentrations in Estonian rivers in the period 1992-2007 as affected by deepening droughts. *Biogeochemistry*, **108**, pp. 351-358

Pakeman, R. J., Gimona, A., Glendell, M., Yeluripati, J., Addy, S., van Hulst, F., Matthews, K., Mitchell, R. & Rivington, M. (2018). *Climate Change, Natural Capital and Adaptation in Scotland's Marginal Lands* [pdf]. https://www.climatexchange.org.uk/media/3256/climate-change-and-marginal-land.pdf [10/05/2022].

Panagos, P., Ballabio, C., Meusburger, K., Spinoni, J., Alewell, C. & Borrelli, P. (2017). Towards estimates of future rainfall erosivity in Europe based on REDES and WorldClim datasets. *Journal of Hydrology*, **548**, 251-262.

Paparrizos, J. & Gravano, L. (2016). K-Shape: Efficient and Accurate Clustering of Time Series. *ACM SIGMOD*, **45**(1), 69-76.

Parinet, B, Lhote, A. & Legube, B. (2004). Principal component analysis: an appropriate tool for water quality evaluation and management—application to a tropical lake system. *Ecological Modelling*, **178**, 295–311.

Parry, L. E., Chapman, P. J., Palmer, S. M., Wallage, Z. E., Wynne, H. & Holden, J. (2015). The influence of slope and peatland vegetation type on riverine dissolved organic carbon and water colour at different scales. *Science of the Total Environment*, **527-528**, 530-539.

Pennington, T. H. (2014). *E. coli* O157 outbreaks in the United Kingdom: past , present, and future. *Infectio and Drug Resistance*, **7**, 211-222.

Perona, E., Bonilla, U. I. & Mateo, P. (1999). Spatial and temporal changes in water quality in a Spanish river. *Science of the Total Environment*, **241**, 75-90.

Peterson, C. M., Rifai, H. S., Villarreal, G. C. & Stein, R. (2011). Modelling *Escherichia Coli* and Its Sources in an Urban Bayou with Hydrologic Simulation Program—FORTRAN , *Journal of Environmental Engineering*, **137**(6), 487-503.

Pickard, A. E., Elliott, J. A., Feuchtmayr, H., Chapman, P. J., Williamson, J., Spears, B. M., Banks, J., Bullen, C., Leith, F., Gaston, L., Moody, C. S., Straiton, S., & Monteith, D. (2021). *Is there potential to manage dissolved organic matter concentrations within upland reservoirs? FREEDOM-BCCR briefing note 3 to the water industry.* UKRI SPF UK Climate Resilience programme – Project no. NE/S016937/2. 2021 [pdf]. https://www.ceh.ac.uk/sites/default/files/2021-09/FREEDOM_BCCR_Reservoir_Interventions_03.pdf [27/10/2021].

Pike, R. J. & Wilson, S. E. (1971). Elevation-Relief Ratio, Hypsometric Integral, and Geomorphic Area-Altitude Analysis. *Geological Society of America Bulletin*, **82**(4), 1079-1084.

Poggio, L., Simonetti, E. & Gimona, A. (2018). Enhancing the WorldClim data set for national and regional applications. *Science of the Total Environment*, **625**, 1628-1643.

Poor, C. J. & McDonnell, J. J. (2007). The effects of land use on stream nitrate dynamics. *Journal of Hydrology*, **332**, 54-68.

Porter, K. D. H., Reaney, S. M, Quilliam, R. S., Burgess, C. & Oliver, D. M. (2017). Predicting diffuse microbial pollution risk across catchments: The performance of SCIMAP and recommendations for future development. *Science of the Total Environment*, **609**, 456-465.

Postel, S. L. & Thompson, B. H. (2005). Watershed protection: Capturing the benefits of nature's water supply services. *Natural Resources Forum*, **29**, 98-108.

Pratt, B. & Chang, H. (2012). Effects of land cover, topography, and built structure on seasonal water quality at multiple spatial scales. *Journal of Hazardous Materials*, **209-210**, 48-58.

Prüss-Ustün, A., Wolf, J., Corvalán, C., Bos, R. & Neira, M. (2016). Preventing disease through healthy environments. A global assessment of the burden of disease from environmental risks [pdf]. https://www.who.int/publications/i/item/9789241565196 [17/03/2023]

Randall, N. P., Donnison, L. M., Lewis, P. J. & James, K. L. (2015). How effective are on-farm mitigation measures for delivering an improved water environment? A systematic map. *Environmental Evidence*, **4**(18), DOI: 10.1186/s13750-015-0044-5

Reaney, S. M., Lane, S. N., Heathwaite, A. L. & Dugdale, L. J. (2011). Risk-based modelling of diffuse land use impacts from rural landscapes upon salmonid fry abundance. *Ecological Modelling*, **222**, 1016–1029.

Reichenberger, S., Bach, M., Skitschak, A. & Frede, H.-G. (2007). Mitigation strategies to reduce pesticide inputs into ground- and surface water and their effectiveness; A review. *Science of the Total Environment*, **384**, 1-35.

Reid, H., Bourne, A., Muller, H., Podvin, K., Scorgie, S. & Orindi, V. (2018). A Framework for Assessing the Effectiveness of Ecosystem-Based Approaches to Adaptation. In Z. Zommers & K. Alverson (Eds). *Resilience* (pp. 207-216). Elsevier.

Reynolds, B. (2007). Implications of changing from grazed or semi-natural vegetation to forestry for carbon stores and fluxes in upland organo-mineral soils in the UK. *Hydrology and Earth System Sciences*, **11**(1), 61-76.

Richardson, C. J., Flanagan, N. E., Ho, M. & Pahl, J. W. (2011). Integrated stream and wetland restoration: A watershed approach to improved water quality on the landscape. *Ecological engineering*, **37**, 25-39.

Ritson, J. P., Graham, N. J. D., Templeton, M. R., Clark, J. M., Gough, R., & Freeman, C. (2014). The impact of climate change on the treatability of dissolved organic matter (DOM) in upland water supplies: A UK perspective. *Science of the Total Environment,* **473-474**, 714-730.

Ritson, J. P., Croft, J. K., Clark, J. M., Brazier, R. E., Templeton, M. R., Smith, D. & Graham, N. J. D. (2019). Sources of dissolved organic carbon (DOC) in a mixed land use catchment (Exe, UK). *Science of the Total Environment*, **666**, 165–175.

Robins, N. S. (2002). Groundwater quality in Scotland: major ion chemistry of the key groundwater bodies. *Science of the Total Environment*, **294**, 41-56.

Rosenberg, D. M., Berkes, F., Bodaly, R. A., Hecky, R. E., Kelly, C. A. & Rudd, J. W. (1997). Large-scale impacts of hydroelectric development. *Environmental Reviews,* **5**, 27-54.

Rotariu, O., Ogden, I. D., MacRitchie, L., Forbes, K. J., Williams, A. P., Cross, P., Hunter, C. J., Teunis, P. F. M. & Strachan, N. J. C. (2012). Combining risk assessment and epidemiological risk factors to elucidate the sources of human *E. coli* O157 infection. *Epidemiology & Infection*, **140**, 1414–1429.

Rothwell, J. J., Evans, M. G., Daniels, S. M. & Allott, T. E. H. (2007). Baseflow and stormflow metal concentrations in streams draining contaminated peat moorlands in the Peak District National Park (UK). *Journal of Hydrology,* **341**, 90-104.

Rothwell, J. J., Dise, N. B., Taylor, K. G., Allott, T. E. H., Scholefield, P., Davies, H., & Neal, C. (2010). Predicting river water quality across North West England using catchment characteristics. *Journal of Hydrology,* **395**, 153-162.

Rounsevell, M. D. A., Reginster, I., Araujo, M. B., Carter, T. R., Dendoncker, N., Ewert, F., House, J. I., Kankaanpää, S., Leemans, R. & Metzger, M. J. M. (2006). A coherent set of future land use change scenarios for Europe. *Agriculture, Ecosystems & Environment,* **114**, 57-68.

Rowland, A. P., Neal, C., Reynolds, B., Neal, M., Lawlor, A. J. & Sleep, D. (2012). Manganese in the upper Severn mid-Wales. *Journal of Environmental Monitoring*, **14**, 155-164.

Sadeghi, A. M., & Arnold, J. G. (2002). A SWAT/Microbial Sub-Model for Predicting Pathogen Loadings in Surface and Groundwater at Watershed and Basin Scales. In *Total Maximum Daily Load (TMDL): Environmental Regulations, Proceedings of the March 11-13, 2002 Conference* (pp. 56-63). American Society of Agricultural and Biological Engineers [pdf]. https://www.researchgate.net/publication/238079474_A_SWATMicrobial_Sub-Model_for_Predicting_Pathogen_Loadings_in_Surface_and_Groundwater_at_Watershed_and_Basin_Scales [10/12/2022].

Sample, J. E., Baber, I. & Badger, R. (2016). A spatially distributed risk screening tool to assess climate and land use change impacts on water-related ecosystem services. *Environmental Modelling & Software,* **83**, 12-26.

Sawicka, K., Rowe, E. C., Evans, C. D., Monteith, D. T., Vanguelova, E. I., Wade, A. J., & Clark, J. M. (2017). Modelling impacts of atmospheric deposition and temperature on long-term DOC trends. *Science of the Total Environment,* **578**, 323-336.

Saxena, T., Kaushik, P. & Mohan, M. K. (2015). Prevalence of *E. coli* O157:H7 in water sources: an overview on associated diseases, outbreaks and detection methods. *Diagnostic Microbiology and Infectious Disease*, **82**(3), 249-264.

Schomers, S. & Matzdorf, B. (2013). Payments for ecosystem services: A review and comparison of developing and industrialized countries. *Ecosystem Services,* **6**, 16-30.

Schoonover, J. E. & Lockaby, B. G. (2006). Land cover impacts on stream nutrients and faecal coliform in the lower Piedmont of West Georgia. *Journal of Hydrology,* **331**, 371-382.

Schoumans, O. F., Chardon, W. J., Bechmann, M. E., Gascuel-Odoux, C., Hofman, G., Kronvang, B., Rubæk, G. H., Ulén, B. & Dorioz, J. M. (2014). Mitigation options to reduce phosphorus losses from the agricultural sector and improve surface water quality: A review. *Science of the Total Environment*, **468-469**, 1255-1266.

Scotlandguides (2020). *Scotland's rocks, landforms and soils* [html]. https://www.scotlandguides.org/scotlands-rocks-landforms-and-soils.html#The_main_soil_types_in_Scotland [10/05/2022]

Scottish Government (2011). *Getting the Best from Our Land. A land Use Strategy for Scotland.* Scottish Government [pdf]. https://www.gov.scot/publications/getting-best-land-land-use-strategy-scotland/ [10/05/2022]

Scottish Water (2017). *Drinking Water Protection Scheme (DWPS)* [pdf]. https://www.scottishwater.co.uk/-/media/ScottishWater/Document-Hub/Key-Publications/Energy-and-Sustainability/Sustainable-Land-Management/DrinkingWaterProtectionSchemeBookletUpdateOct2017.pdf [10/05/2022].

Scottish Water (2020). *Factsheet 3: Water treatment explained* [pdf]. https://www.scottishwater.co.uk/Help-and-Resources/Document-Hub/Factsheets-and-Leaflets/Factsheets [10/05/2022].

Scottish Water (2021). *Annual Report & Accounts 2020/21: Performance and prospects* [pdf]. https://www.scottishwater.co.uk/Help-and-Resources/Document-Hub/Key-Publications/Annual-Reports [10/05/2022].

Scottish Water (2022). *The Water Industry in Scotland* [html]. https://www.scottishwater.co.uk/about-us/what-we-do/the-water-industry-in-scotland?msclkid=14dc5340d0a211ec94e3bde4c6e87e55 [10/05/2022]

Scottish Water (n. d.). *A Sustainable Future Together* [pdf]. https://www.scottishwater.co.uk/-/media/ScottishWater/Document-Hub/Key-Publications/Strategic-Plan/030220StrategicPlanASustainableFutureTogether.pdf [17/08/2022].

Seddon, N., Daniels, E., Davis, R., Chausson, A., Harris, R., Hou-Jones, X., Huq, S., Kapos, V., Mace, G. M., Rizvi, A. R., Reid, H., Roe, D., Turner, B. & Wicander, S. (2020). Global recognition of the importance of nature-based solutions to the impacts of climate change. *Global Sustainability*, **3**, e15, 1-12, https://doi.org/10.1017/sus.2020.8.

Seddon, N., Chausson, A., Berry, P., Giardin, C. A. J., Smith, A. & Turner, B. (2020). Understanding the value and limits of nature-based solutions to climate change and other global challenges. *Philosophical Transactions of the Royal Society B*, **375**, 20190120, http://dx.doi.org/10.1098/rstb.2019.0120.

Selle, B., Schwientek, M., & Lischeid, G. (2013). Understanding processes governing water quality in catchments using principal component scores. *Journal of Hydrology,* **486**, 31-38.

SEPA (n. d.). *Rural diffuse pollution plan for Scotland (2015-2021)* [pdf]. https://www.sepa.org.uk/media/330130/rural-diffuse-pollution-plan-for-scotland-2015-2021.pdf [10/05/2022]

SEPA (2015). *Scottish bathing waters 2014-2015* [pdf]. http://www.sepa.org.uk/media/143136/sepa_bathing_waters_2014-15_web.pdf [10/05/2022].

Shareef, M.A., Toumi, A. & Khenchaf, A. (2014). Estimation of Water Quality Parameters Using the Regression Model with Fuzzy K-Means Clustering. *International Journal of Advanced Computer Science and Applications*, **5**(6), 151-157.

Sharp, E. L., Parsons, S. A. & Jefferson, B. (2006). Seasonal variations in natural organic matter and its impact on coagulation in water treatment. *Science of the Total Environment*, **363**, 183-194.

Sharpley, M., Matlock, M., Heathwaite, L., Simpson, T. (2009). Managing Agricultural Catchments to Sustain Production and Water Quality. In R.C. Ferrier & A. Jenkins (2009). *Handbook of Catchment Management* (pp. 107-134). Wiley-Blackwell.

Shen, Y. N., Lü, J., Chen, D. J., & Shi, Y. M. (2011). Response of stream pollution characteristics to catchment land cover in Cao-E River basin, China. *Pedosphere*, **21**(1), 115–123.

Shen, Z., Hou, X., Li, W., Aini, G., Chen, L. & Gong, Y. (2015). Impact of landscape pattern at multiple spatial scales on water quality: A case study in a typical urbanised watershed in China. *Ecological Indicators*, **48**, 417–427.

Shi, P., Zhang, Y., Zhanbin, L., Li, P., & Xu, G. (2017). Influence of land use and land cover patterns on seasonal water quality at multi-spatial scales. *Catena*, **151**, 182-190.

Siakeu, J., Oguchib, T., Aoki, T., Esakid, J. & Jarvie, H. P. (2002). Change in riverine suspended sediment concentration in central Japan in response to late 20th century human activities. *Catena*, **55**, 231-254.

Sillanpää, M., Ncibi, M. C., Matilainen, A. & Vepsäläinen, M. (2018). Removal of natural organic matter in drinking water treatment by coagulation: A comprehensive review. *Chemosphere*, **190**, 54-71.

Singh, O., Sarangi, A. & Sharma, M. C. (2008). Hypsometric Integral Estimation Methods and its Relevance on Erosion Status of North-Western Lesser Himalayan Watersheds. *Water Resource Management*, **22**, 1545–1560.

Singh, K. P., Malik, A., Mohan, D., & Sinha, S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. *Water Research*, **38**, 3980-3992.

Singh, J., Knapp, V. & Demissie, M. (2004*). Hydrologic Modelling of the Iroquois River Watershed Using HSPF and SWAT.* Illinois State Water Survey Contract Report 2004-08, Watershed Science Section [pdf]. https://swat.tamu.edu/media/90101/singh.pdf [19/12/2021].

Snow, R. (2005). *Continental Climate and Continentality* [html]. http://link.springer.com/referenceworkentry/10.1007/1-4020-3266-8_58/fulltext.html#copyrightInformation [06/10/2022].

Soulsby, C., Gibbins, C., Wade, A. J., Smart, R. & Helliwell, R. (2002). Water quality in the Scottish uplands: a hydrological perspective on catchment hydrochemistry. *Science of the Total Environment,* **294**, 73-94.

Struyf, J., Ženko, B., Blockeel, H., Vens, C., & Džeroski, D. (2011). *CLUS: User's Manual.* [pdf]. https://www.researchgate.net/publication/265243854_Clus_User%27s_Manual [16/12/2018].

Struyf, J. & Džeroski, S. (2006). Constraint Based Induction of Multi-objective Regression Trees. In F. Bonchi & J.F. Boulicaut (Eds). *Knowledge Discovery in Inductive Databases*. Springer.

Stutter, M. I., Chardon, W. J. & Kronvang, B. (2012). Riparian buffer strips as a multifunctional management tool in agricultural landscapes: introduction. *Journal of Environmental Quality,* **41**, 297-303.

Stutter, M. I., Dunn, S. M. & Lumsdon, D. G. (2012). Dissolved organic carbon dynamics in a UK podzolic moorland catchment: linking storm hydrogeochemistry, flow path analysis and sorption experiments. *Biogeosciences*, **9**, 2159-2175.

Sunstein, C. R. (2005). *Laws of Fear: Beyond the Precautionary Principle.* Cambridge University Press.

SWW (n. d.). *Upstream Thinking 2015-2020. An overview of progress contributing to 10 years of Upstream Thinking in the South West* [html]. https://www.southwestwater.co.uk/siteassets/document-repository/environment/j121-sww-ust-v7-290920.pdf [13/09/2022].

Tang, X., Zhu, B. & Katou, H. (2012). A review of rapid transport of pesticides from sloping farmland to surface waters: Processes and mitigation strategies. *Journal of Environmental Sciences,* **24**, 351-361.

Tetzlaff, D., Capell, R., & Soulsby, C. (2012). Land use and hydroclimatic influences on Faecal Indicator Organisms in two large Scottish catchments: Towards land use-based models as screening tools. *Science of the Total Environment,* **434**, 110-122.

Teunis, P. F. M., & Strachan, N. J. C. (2012). Combining risk assessment and epidemiological risk factors to elucidate the sources of human *E. coli* O157 infection. *Epidemiology & Infection*, **140**, 1414-1429.

Thompson, J., Cassidy, R., Dody, D. G. & Flynn, R. (2013). Predicting critical source areas of sediment in headwater catchments. *Agriculture, Ecosystems and the Environment*, 179, 41-52.

Tipping, E., Smith, E. J., Bryant, C. L. & Adamson, J. K. (2007). The organic carbon dynamics of a moorland catchment in N. W. England. *Biogeochemistry*, **84**, 171-189.

Tiwari, T., Laudon, H., Beven, K. & Ågren, A. M. (2014). Downstream changes in DOC: Inferring contributions in the face of model uncertainties. *Water Resources Research*, **50**, 514-525.

Tompkins, E. L. & Adger, W. N. (2004). Does Adaptive Management of Natural Resources Enhance Resilience to Climate Change? *Ecology and Society*, **9**(2), 10, http://www.ecologyandsociety.org/vol9/iss2/art10/

Tranvik, L.J. & Jansson, M. (2002). Terrestrial export of organic carbon. *Nature*, **415**, 861–862.

Trudgill, S. T. (1986). Introduction. In S.T. Trudgill (Ed.). *Solute Processes*. Wiley.

Tu, J. & Xia, Z.-G. (2008). Examining spatially varying relationships between land use and water quality using geographically weighted regression I: Model design and evaluation. *Science of the Total Environment*, **407**, 358-378.

Tu, J. (2011). Spatially varying relationships between land use and water quality across an urbanization gradient explored by geographically weighted regression. *Applied Geography*, **31**, 376-392.

United Nations General Assembly (2015). *Resolution adopted by the General Assembly on 27 July 2012*, A/RES/66/288 [pdf]. https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/66/288&Lang=E [10/05/2022].

United Utilities (2022). *History of Catchment Systems Thinking* [html]. https://www.unitedutilities.com/corporate/responsibility/stakeholders/catchment-systems-thinking/catchment-management/ [13/09/2022].

Ussiri, D. A. N. & Johnson, C. E. (2004). Sorption of organic carbon fractions by spodosol mineral horizons. *Soils Science Society of America Journal*, **68**(1), 253-262.

Vadde, K. K., Wang, J., Cao, L., Yuan, T., McCarthy, A. J. & Sekar, R. (2018). Assessment of Water Quality and Identification of Pollution Risk Locations in Tiaoxi River (Taihu Watershed), China. *Water*, **10**(2), 183, https://doi.org/10.3390/w10020183.

Vairavamoorthy, K. (2021). Adaptive Planning: a tool for times of uncertainty. In K. Hayward (Ed.): *The Source – The Magazine of the International Water Association. December 2021, Issue 25* (pp. 20-23). International Water Association.

Van Dijk, A. I. J. M. & Keenan, R. J. (2007). Planted forests and water in perspective. *Forest ecology and management*, **251**, 1-9.

Varol, M. (2020). Spatio-temporal changes in surface water quality and sediment phosphorus content of a large reservoir in Turkey. *Environmental Pollution*, **259**, 113860, https://doi.org/10.1016/j.envpol.2019.113860.

Verstraeten, G., Poesen, J., Gillijns, K. & Govers, G. (2006). The use of riparian vegetated filter strips to reduce river sediment loads: an overestimated control measure? *Hydrological Processes*, **20**, 4259-4267.

Vidon, P, Carleton, W. & Mitchell, M. J. (2014). Spatial and temporal variability in stream dissolved organic carbon quantity and quality in an Adirondack forested catchment. *Applied Geochemistry*, **46**, 10-18.

Villanueva, C. M., Cordier, S., Font-Ribera, L. Salas, L. A. & Levallois, P. (2015). Overview of Disinfection By-products and Associated Health Effects. *Current Environmental Health Reports*, **2**(1), 107–115.

Viner, D., Ekstrom, M., Hulbert, M., Warner, N. K., Wreford, A. & Zommers, Z. (2020). Understanding the dynamic nature of risk in climate change assessments—A new starting point for discussion. *Atmospheric Science Letters*, **21**, e958, https://doi.org/10.1002/asl.958.

Vinten, A. J. A., Douglas, J. T., Lewis, D. R., Aitken, M. N. & Fenlon, D. R. (2004). Relative risk of surface water pollution by *E. coli* derived from faeces of grazing animals compared to slurry application. *Soil Use and Management*, **20**, 13-22.

Wade, A. J., Jackson, B. M & Butterfield, D. (2008). Over-parameterised, uncertain ‘mathematical marionettes’ - How can we best use catchment water quality models? An example of an 80-year catchment-scale nutrient balance. *Science of the Total Environment*, **400**, 52-74.

Wagener, T., Sivapalan, M., Troch, P., & Woods, R. (2007). Catchment classification and hydrologic similarity. *Geography Compass*, **1**(4), 901-931.

Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P., & Krayer von Krauss, M. P. (2003). Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment*, **4**(1), 5–17.

Walker, W. E., Haasnoot, M. & Kwakkel, J. H. (2013). Adapt or Perish: A Review of Planning Approaches for Adaptation under Deep Uncertainty. *Sustainability*, **5**, 955-979.

Wallage, Z. E., Holden, J. & McDonald, A. T. (2006). Drain blocking: An effective treatment for reducing dissolved organic carbon loss and water discolouration in a drained peatland. *Science of the Total Environment*, **367**, 811-821.

Wamsler, C., Luederitz, C & Brink, E. (2014). Local levers for change: Mainstreaming ecosystem-based adaptation into municipal planning to foster sustainability transition. *Global Environmental Change*, **29**, 189-201.

Warrington, B. M., Aust, W. M., Barrett, S. M., Ford, W. M., Dolloff, C. W., Schilling, E. B., Wigley, T. B., & Bolding, M. C. (2017). Forestry Best Management Practices Relationships with Aquatic and Riparian Fauna: A Review. *Forests*, **8**(9), 331, https://doi.org/10.3390/f8090331

Wassénius, E. & Crona, B. I. (2022). Adapting risk assessments for a complex future. *One Earth*, **5**, 35-43.

Watts, G., Battarbee, R. W., Bloomfield, J. P., Crossman, J., Daccache, A., Durance, I., Elliot, J. A., Garner, G., Hannaford, J., Hannah, D. M., Hess, T., Jackson, C. R., Kay, A. L., Kernan, M., Knox, J., Mackay, J., Monteith, D., Ormerod, S. J., Rance, J., Stuart, M. E., Wade, A. J., Wade, S. D., Weatherhead, K., Whitehead, P. Wilby, R. L. (2015). Climate change and water in the UK – past changes and future prospects. *Progress in Physical Geography: Earth and Environment*, **39**(1), 6-28.

Weatherhead, E. K. & Howden, N. J. K. (2009). The relationship between land use and surface water resources in the UK. *Land Use Policy*, **26S**, S243–S250.

Weaver, C. P., Lempert, R. J., Brown, C., Hall, J. A., Revell, D. & Sarewitz, D. (2013). Improving the contribution of climate model information to decision making: the value and demands of robust decision frameworks. *WIREs Climate Change*, **4**, 39-60.

Wen, H., Perdrial, J., Bernal, S., Abbott, B. W., Dupas, R., Godsey, S. E., Harpold, A., Rizzo, D., Underwood, K., Adler, T., Hale, R., Sterle, G. & Li, L. (2020). Temperature controls production but hydrology controls export of dissolved organic carbon at the catchment scale. *Hydrology and Earth System Sciences Discussions*, **24**, 1-22.

Whitehead, P. G., Butterfield, D. & Wade, A. J. (2009). Simulating metals and mine discharges in river basins using a new integrated catchment model for metals: pollution impacts and restoration strategies in the Aries–Mures river system in Transylvania, Romania. *Hydrology Research*, **40**(2-3), 323–346.

Whitehead, P. G., Wilby, R. L., Battarbee, R. W., Kernan, M. & Wade, A. J. (2009). A review of the potential impacts of climate change on surface water quality. *Hydrological Sciences Journal*, **54**(1), 101-123.

Whitehead, P. G., Leckie, H., Rankinen, K., Butterfield, D., Futter, M. N. & Bussi, G. (2016). An INCA model for pathogens in rivers and catchments: Model structure, sensitivity analysis and application to the River Thames catchment, UK. *Science of the Total Environment*, **572**, 1601-1610.

WHO (2011). *Guidelines for drinking-water quality*. 4th edition [pdf]. https://www.who.int/publications/i/item/9789241549950?msclkid=7d75cff3d03f11ec8faf4ad0d6daeaa7 [10/05/2022].

WHO (2017a). S*afely managed drinking water – thematic report on drinking water 2017* [pdf]. https://data.unicef.org/resources/safely-managed-drinking-water/ [10/05/2022].

WHO (2017b). *Climate-resilient water safety plans: managing health risks associated with climate variability and change* [pdf]. https://www.who.int/publications/i/item/9789241512794?msclkid=b08059afd03f11ec87df97a12026e7f2 [10/05/2022].

WHO (2017c). *Guidelines for Drinking-water Quality. Fourth Edition Incorporating the First Addendum* [pdf]. https://www.who.int/publications-detail-redirect/9789241549950?msclkid=43b33444d06711ec881e4ef71f9fe986 [10/05/2022].

Wilkinson, M. E., Quinn, P. F., Barber, N. J. & Jonczyk, J. (2014). A framework for managing runoff and pollution in the rural landscape using a Catchment Systems Engineering approach. *Science of the Total Environment*, **468-469**, 1245-1254.

Willett, I. R. & Porter, K. (2001). *Watershed Management for Water Quality Improvement: the role of agricultural research* [pdf]. https://www.aciar.gov.au/publication/watershed-management-water-quality-improvement-role-agricultural-research [13/09/2022].

Wilson, L., Wilson, J., Holden, J., Johnstone, I., Armstrong, A. & Morris, M. (2011). Ditch blocking, water chemistry and organic carbon flux: Evidence that blanket bog restoration reduces erosion and fluvial carbon loss. *Science of the Total Environment*, **409**, 2010-2018.

Winterdahl, M. Erlandsson, M., Futter, M. N., Weyhenmeyer, G. A. & Bishop, K. (2014). Intra-annual variability of organic carbon concentrations in running waters: Drivers along a climatic gradient. *Global Biogeochemical Cycles*, **28**, 451-464.

Winterdahl, M., Laudon, H., Lyon, S. W., Pers, C. & Bishop, K. (2016). Sensitivity of stream dissolved organic carbon to temperature and discharge: Implications of future climates. *Journal of Geophysical Research: Biogeosciences*, **121**, 126-144.

Worrall, F., Armstrong, A. & Adamson, J. K. (2007). The effects of burning and sheep-grazing on water table depth and soil water quality in an upland peat. *Journal of Hydrology*, **339**, 1-14.

Worrall, F., Armstrong, A. & Holden, J. (2007). Short-term impact of peat drain-blocking on water colour, dissolved organic carbon concentration, and water table depth. *Journal of Hydrology*, **337**, 315-325.

Worrall, F., Spencer, E. & Burt, T. P. (2009). The effectiveness of nitrate vulnerable zones for limiting surface water nitrate concentrations. *Journal of Hydrology*, **370**, 21-28.

WWAP (United Nations World Water Assessment Programme)/UN-Water (2018). *The United Nations World Water Development Report 2018: Nature-based Solutions for Water*. UNESCO [html]. https://unesdoc.unesco.org/ark:/48223/pf0000261424 [10/12/2022].

Wyatt, K. H., Turetsky, M. R., Rober, A. R., Giroldo, D., Kane, E. S., Stevenson, R. J. (2012). Contributions of algae to GPP and DOC production in an Alaskan fen: effects of historical water table manipulations on ecosystem responses to a natural flood. *Oecologia*, **169**, 821–832.

Xenopoulos, M. A., Lodge, D. M., Frentress, J., Kreps, T. A., Bridgham, S. D., Grossman, E. & Jackson, C. J. (2003). Regional comparisons of watershed determinants of dissolved organic carbon in temperate lakes from the Upper Great Lakes region and selected regions globally. *Limnology and Oceanography*, **48**, 2321-2334.

Yallop, A. R. & Clutterbuck, B. (2009). Land management as a factor controlling dissolved organic carbon release from upland peat soils 1: spatial variation in DOC productivity. *Science of the Total Environment*, **407**, 3803–3813.

Young, I., Smith, B. A. & Fazil, A. (2015). A systematic review and meta-analysis of the effects of extreme weather events and other weather-related variables on Cryptosporidium and Giardia in fresh surface waters. *Journal of Water and Health*, **13**(1), 1-17.

Zeckoski, R.W., Lazar, A.N., Liang, D. & Wade, A.J. (2009). *Comparison of the HSPF and HBV-INCA Models: Crossing the Atlantic Divide.* American Society of Agricultural and Biological Engineers, St. Joseph, Michigan. Reno, Nevada, June 21 - June 24 (html).
https://elibrary.asabe.org/abstract.asp?JID=5&AID=27169&CID=reno2009&T=1 [17/08/2022]

Appendices

A. Available data sets

A.1 Climate change risk to raw water quality data set 1: Catchment characteristics

DOI 10.15132/10000197, “DS1_SWCatchmentdataall.csv”

A.2 Climate change risk to raw water quality data set 2: Subset catchment characteristics & water quality summary

DOI 10.15132/10000198, “DS2_SWCatchmentWaterQualitySubset.csv”

A.3 Climate change risk to raw water quality data set 3: Water quality (TOC, Iron, Manganese, Turbidity & Colour)

DOI 10.15132/10000199, “DS3_TOCColourSeries.csv”

A.4 Climate change risk to raw water quality data set 4: Colour time series

DOI 10.15132/10000200, “DS4_ColourTimeSeries20102016.csv”

A.5 Climate change risk to raw water quality data set 5: TOC & climate data

DOI 10.15132/10000201, “DS5_TOCClimateSeries20132016.csv”

A.6 Climate change risk to raw water quality data set 6: *E. coli* & rainfall data

DOI 10.15132/10000202, “DS6_EcoliRainfallSeries20102016.csv”

B. Example R codes

B.1 PCA

```
# Load packages and dataset
library("FactoMineR")
library("factoextra")
data <-read.csv("DS2_SWCatchmentWaterQualitySubset.csv")
# Choose required variables
tmp <- data[,c("Alu50","Col50","Iron50","Mang50","pH50","Turb50","Coli50","Ecoli50","Area","ReliefRatio","SlopeLittle","SlopeSteep",
"GeologyLimestone","GeologySandstone","TOCAverage","BFIAvergae","SPRAverage","Heather15","Decid15","Urban15",
"Arable15","Imprgrass15","Cattle","Sheep","SepticTank","TempMeanAnnual","PrecTotAnnual","PrecdaysAnnual")]
tmp[,c(7,8)]<-tmp[,c(7,8)]+1 # add 1 to Coliform and E.coli medians for log transformation
tmp[,c(1,2,3,4,6,7,8)] <- log10(tmp[,c(1,2,3,4,6,7,8)]) # log transform water quality medians except pH
# run PCA
res.pca <- PCA(tmp, scale.unit=TRUE, quanti.sup =9:28, graph = TRUE)
# quanti.sup indicates columns of tmp corresponding to the supplementary variables
# The remaining columns/variables are used for the PCA
res.pca$eig
res.pca$var
res.pca$quanti.sup
scores <- res.pca$ind
sweep(res.pca$var$coord,2,sqrt(res.pca$eig[1:ncol(res.pca$var$coord),1]),FUN="/") #calculates the loadings
dimdesc(res.pca) #caluclates correlations and significance
par(pty="m")
# Biplot
fviz_pca_biplot(res.pca, pointsize=2, label="var", labelsize=7, col.quanti.sup="grey40",
```

```
            col.var = "black",geom.var=c("arrow","text"), # This defines symbol and labels for the columns (variables)
            col.ind = "black",
            geom.ind=c("point","text"), # This defines symbol and labels for the rows
            repel=TRUE, # to avoid label overplotting
            title="") +
ylab("PC2") + xlab("PC1")+ theme(axis.text = element_text(size=18),
                       axis.title = element_text(size=20))
```

B.2 RDA

```
# load package and data
library(vegan)
library(corrplot)
data <- read.csv("DS2_SWCatchmentWaterQualitySubset.csv")
# Choose required variables
tmp <- data[,c("Alu50","Col50","Iron50","Mang50","pH50","Turb50","Coli50","Ecoli50","Area",
"ReliefRatio","SlopeLittle","SlopeSteep","GeologyLimestone","TOCAverage","BFIAvergae","SPRAverage","Heather15","Decid15","Conif15",
"Urban15","Arable15","Imprgrass15","Cattle","Sheep","SepticTank","TempMeanAnnual","PrecTotAnnual","PrecdaysAnnual")]
tmp[,c(7,8)]<-tmp[,c(7,8)]+1 # add 1 to Coliform and E.coli medians for log transformation
tmp[,c(1,2,3,4,6,7,8)] <- log10(tmp[,c(1,2,3,4,6,7,8)]) # log transform water quality medians except pH
water<- tmp[,c(1:8)]
pred<- tmp[,c(9:28)]
# run RDA
m1<-rda(water~., pred) # full model
m0<-rda(water~1, pred) # null model
mback <- step(m1, test="perm") # backward elimination of non-significant variables
mback # look at remaining variables
corrplot(cor(pred)) # check for potential mulitcollinearity
# remove slope variables and only retain ReliefRatio to avoid multicollinearity
mown<-rda(water~ReliefRatio + BFIAvergae + TOCAverage + Imprgrass15 + Arable15 + Urban15 + Sheep + Cattle + PrecdaysAnnual, data = pred)
mown
summary(mown)
anova(mown) # significant model
vif.cca(mown) # check variance inflation factors
# between 1.2 and 2.3, which is moderate so no concern
par(pty="s")
```

```
plot(mown, display="sites", xlim=c(-2,3), cex.lab=2, cex.axis=2)
points(mown, display="sites", pch=19, cex=1.5)
points(mown, display="species", pch=3, cex=2, scaling=1)
text(mown, display="species", cex=2.5, scaling=1.5)
points(mown, display = "cn", col="darkgrey", scaling=1)
text(mown, display="cn", col="darkgrey", cex=2.5, scaling=1)
```

B.3 Cluster analysis

```
# load packages and data
library(cluster)
data<-read.csv("DS2_SWCatchmentWaterQualitySubset.csv")
# Choose required variables
means<-data[,c("Alu50", "Col50", "Iron50", "Mang50", "Turb50", "pH50", "Coli50", "Ecoli50")]
means<-scale(means) # scaling to 0 means
# Run PAM
d2<-dist(means, method="euclidean")
means.pam<-pam(d2,5) # adjust number of clusters as required
plot(means.pam)
clusplot(means, means.pam$cluster, main='2D representation of the Cluster solution', color=TRUE, shade=TRUE,labels=2, lines=0)
kruskal.test(data$Alu50~data$Cluster) #test if difference between clusters is significant
```

B.4 Linear regression models colour-TOC/iron/manganese/turbidity

```
# load packages and data
library(dplyr)
library(broom)
library(reshape2)
options(scipen=999)
colTOC<-read.csv("DS3_TOCColourSeries.csv")
colTOC<-na.omit(colTOC) # remove entries where no TOC data
GROUPS <- unique(colTOC$Catchment.ID) # Specify groups, one unique model per group (per catchment)
# run model in a loop
for (i in 1:length(GROUPS)){
  CURRENT_GROUP <- GROUPS[i]
  df <- filter(colTOC, Catchment.ID==CURRENT_GROUP) # subset the dataframe
  fit <- lm(Colour ~ TOC+Iron+Manganese+Turbidity, data = df) # Build a model, include variables as required
  coeff <- tidy(fit) # Get a pretty data frame of the coefficients & p values
  coeff <- coeff[,c(1,2,5)] # Extract P.Value & Estimate
    # Rename (intercept) to INT for pretty column names
  coeff[coeff$term=="(Intercept)", ]$term <- "INT"
    # Make it into wide format with reshape2 package.
  coeff <- coeff %>% melt(id.vars=c("term"))
    # Defactor the resulting data.frame
  coeff <- mutate(coeff,
                  variable=as.character(variable),
                  term=as.character(term))
  # Rename for prettier column names later
  coeff[coeff$variable=="estimate", ]$variable <- "Beta"
  coeff[coeff$variable=="p.value", ]$variable <- "P"
  coeff <- dcast(coeff, .~term+variable)[,-1]
  rsquared <- summary(fit)$r.squared
```

```
  # Create a df
  row <- cbind(data.frame(group=CURRENT_GROUP, rsquared=rsquared), coeff)
    # If first iteration, create data.frame -- otherwise: rowbind
  if (i==1){
    RESULT_ROW = row
  }
  else{
    RESULT_ROW = rbind(RESULT_ROW, row)
  } # End if.
} # End for loop
RESULT_ROW[,c(2:12)]<-round(RESULT_ROW[,c(2:12)], 4)
write.table(RESULT_ROW, "TOCIronMangTurbColourmodel.csv") # specify file name as required
```

B.5 Shape-based clustering on colour time series

```
# load packages and data
library(reshape2)
library(imputeTS)
library(TSrepr)
library(dtwclust)
ts<-read.csv("DS4_ColourTimeSeries20102016.csv")
ts<-na_interpolation(ts, option = "linear") # Interpolate missing values
ts[,c(2:155)]<-log10(ts[,c(2:155)]) #log data
AllTS<-t(ts) # Switch rows and columns
colnames(AllTS)<-AllTS[1,]
AllTS<-AllTS[-1,]
#Extract seasonal profiles
class(AllTS)<-"numeric"
SP<-apply(AllTS, 1, function(x) repr_seas_profile(x, 365, meanC))
SPall<-t(SP)
#Clustering
spclustering<-tsclust(SPall, type="partitional", k=2, preproc=zscore, distance = "sbd", centroid = "shape")
#clustering on time series shapes for yearly means, specify number of clusters with k=...
spclustering
plot(spclustering)
SPClusters<-predict(spclustering)
write.csv(SPClusters, "SPCLusterList.csv")
```

B.6 Climate sensitivity testing for TOC

```
# Load packages and data
library("lubridate")
sens<-read.csv("DS5_TOCClimateSeries20132016.csv")
# Seperate by season
# Create a column with months first
sens$Date<-format(sens$Date)
sens$month<-month(sens$Date)
sens$season<-ifelse(sens$month==1, "winter", ifelse(sens$month==2, "winter", ifelse(sens$month==3, "spring", ifelse(sens$month==4, "spring",
ifelse(sens$month==5, "spring", ifelse(sens$month==6, "summer", ifelse(sens$month==7, "summer", ifelse(sens$month==8, "summer",
ifelse(sens$month==9, "autumn", ifelse(sens$month==10, "autumn", ifelse(sens$month==11, "autumn", ifelse(sens$month==12, "winter",
NA))))))))))))
# Seperate
sensWinter<-sens[which(sens$season=="winter"),]
sensSpring<-sens[which(sens$season=="spring"),]
sensAutumn<-sens[which(sens$season=="autumn"),]
sensSummer<-sens[which(sens$season=="summer"),]
# Spearman's correlation test per catchment
# Split catchments
catch<-split(sensSummer, sensSummer$Catchment.ID) # adjust dataset as required
#Test function
regr<-cor.test(~TOC+Rainfall03, data = catch[[i]], method="spearman", continuity=FALSE, conf.level=0.95)$estimate
#Run loop for Spearman's rank correlation coeffiencient, adjust variable as required
for (i in names(catch)){
  regr[i]<-cor.test(~TOC+Rainfall03, data = catch[[i]], method="spearman", continuity=FALSE, conf.level=0.95)$estimate # $p.value for significance
  print(regr[i])
}
```

B.7 Linear regression models on TOC medians per category

```
# load packages and data
library(foreign)
library(Hmisc)
library("PerformanceAnalytics")
library("corrplot")
library(caret)
data <- read.csv("DS2_SWCatchmentWaterQualitySubset.csv")
datacomp<-data[!is.na(data$TOCMedian),] # Remove rows with missing values
n=nrow(datacomp)
modelfullBIC<-with(datacomp, aov(log(TOCMedian) ~ Area+ElevationReliefRatio+
                    ReliefRatio+Aspect+TOCAverage+BFIAvergae
                   +Other15+Conif15+Arable15+Imprgrass15+Deer+
                   +Sheep+Cattle+pwsurplus81+aat55for810+pHMed))
modelfinalBIC=step(modelfullBIC,k=log(n))
summary.lm(modelfinalBIC)
### This model explains 47%.
### Let's check behaviour of residuals.
par(mfrow=c(2,2))
plot(modelfinalBIC)
### Validation
set.seed(0)
train_control <- trainControl(method = "repeatedcv", number = 10, repeats = 10)
modelfinalBIC <- train(log(TOCMedian) ~ Area+ReliefRatio+TOCAverage
               +Imprgrass15+Sheep+pwsurplus81+aat55for810+pHMed, data = datacomp, method = "lm",
               trControl = train_control)
print(modelfinalBIC)
```

```
### RSME 0.4 R2 0.45 MAE 0.32
mean(log(datacomp$TOCMedian)) # 1.59
0.4172285/(max(log(datacomp$TOCMedian))-min(log(datacomp$TOCMedian))) #0.13
# Model for category
subrain<-data[which(data$CategoryTOC=="Rainfall"),]
subraintemp<-data[which(data$CategoryTOC=="Rainfall+Temp"),]
subtemp<-data[which(data$CategoryTOC=="Temp"),]
subwetup<-data[which(data$CategoryTOC=="Wetup"),]
subnone<-data[which(data$CategoryTOC=="None"),]
# repeat above for each dataset
```

B.8 Linear regression models on TOC medians with interactions

```
# load packages and data
library(foreign)
library(Hmisc)
library("PerformanceAnalytics")
library("corrplot")
library(caret)
data <- read.csv("DS2_SWCatchmentWaterQualitySubset.csv")
datacomp<-data[!is.na(data$TOCMedian),] # Remove rows with missing values
n=nrow(datacomp)
# In this analysis we treat TOCMedian as the response and introduce all variables at the same time plus a two-way interaction with one variable at a time
# Group 1 = interactions with pwsurplus81
fullprec<-
with(datacomp,aov(log(TOCMedian)~CategoryTOC+Area+ElevationReliefRatio+ReliefRatio+Aspect+TOCAverage+BFIAvergae+Other15+Conif15+Arable15+Imprgrass15+Deer+Cattle+Sheep+pwsurplus81+aat55for810+pHMed+pwsurplus81*(CategoryTOC+Area+ElevationReliefRatio+ReliefRatio+Aspect+aat55for810+TOCAverage+BFIAvergae+Other15+Conif15+Arable15+Imprgrass15+Deer+Cattle+Sheep+pHMed)))
modelprecBIC=step(fullprec,k=log(n))
summary.lm(modelprecBIC)
# Group 2 = interactions with aat55for810
fulltemp<-
with(datacomp,aov(log(TOCMedian)~CategoryTOC+Area+ElevationReliefRatio+ReliefRatio+Aspect+TOCAverage+BFIAvergae+Other15+Conif15+Arable15+Imprgrass15+Deer+Cattle+Sheep+pwsurplus81+aat55for810+pHMed+aat55for810*(CategoryTOC+Area+ElevationReliefRatio+ReliefRatio+Aspect+pwsurplus81+TOCAverage+BFIAvergae+Other15+Conif15+Arable15+Imprgrass15+Deer+Cattle+Sheep+pHMed)))
modeltempBIC=step(fulltemp,k=log(n))
summary.lm(modeltempBIC)
# Group 3 = Category interactions
```

```
fullcat<-
with(datacomp,aov(log(TOCMedian)~CategoryTOC+Area+ElevationReliefRatio+ReliefRatio+Aspect+TOCAverage+BFIAvergae+Other15+Conif15
+Arable15+Imprgrass15+Deer+Cattle+Sheep+pwsurplus81+aat55for810+pHMed+CategoryTOC*(aat55for810+Area+ElevationReliefRatio+ReliefRa
tio+Aspect+pwsurplus81+TOCAverage+BFIAvergae+Other15+Conif15+Arable15+Imprgrass15+Deer+Cattle+Sheep+pHMed)))
modelcatBIC=step(fullcat,k=log(n))
summary.lm(modelcatBIC)
### Try this model on its own
### This model explains 54%.
### Let's check behaviour of residuals.
par(mfrow=c(2,2))
plot(modelcatBIC)
### Let's look at the predictive capability of the model
par(mfrow=c(1,1)) # reset plotting pane
plot(log(datacomp$TOCMedian),predict(modelcatBIC)) # Plot the response against the predicted values of the response
### Formalise this as a linear model
modelcatBICform <- with(datacomp,lm(predict(modelcatBIC)~log(datacomp$TOCMedian)))
abline(modelcatBICform) # This plots a straight line through the data by least-squares
### If the model predicted the response perfectly, then they would be linked by a straight line of slope 1
### If we insert such a line on the plot as a reference line we can make that comparison
abline(0,1)
### Validation
set.seed(0)
train_control <- trainControl(method = "repeatedcv", number = 10, repeats = 10)
modelcatBICVal <- train(log(TOCMedian) ~ CategoryTOC + Aspect + TOCAverage +
                Conif15 + Imprgrass15 + Deer + Sheep + pwsurplus81 + aat55for810 +
                pHMed + CategoryTOC:Conif15, data = datacomp, method = "lm", trControl = train_control)
print(modelcatBICVal)
### RSME 0.38 R2 50% MAE 0.3
```

```
0.3811012/(max(log(datacomp$TOCMedian))-min(log(datacomp$TOCMedian))) #0.125
confint(modelcatBIC)
# Group 4 = Other possibly important interactions
fullother<-
with(datacomp,aov(log(TOCMedian)~+CategoryTOC+Area+ElevationReliefRatio+ReliefRatio+Aspect+aat55for810+pwsurplus81+TOCAverage+BF
IAvergae+Other15+Conif15+Arable15+Imprgrass15+Deer+Cattle+Sheep++pHMed+TOCAverage:Conif15+TOCAverage:Imprgrass15+TOCAverage
:Other15+TOCAverage:Sheep+TOCAverage:Cattle+TOCAverage:Deer+ReliefRatio:TOCAverage+ReliefRatio:Aspect+ReliefRatio:Conif15+ReliefR
atio:Other15+ReliefRatio:Imprgrass15+ReliefRatio:Sheep+ReliefRatio:Cattle+ReliefRatio:Deer+TOCAverage:pHMed))
modelotherBIC=step(fullother,k=log(n))
summary.lm(modelotherBIC)
# Put this all together now.
modelprecBIC
modeltempBIC
modelcatBIC
modelotherBIC
modelfullBIC<-with(datacomp, aov(log(TOCMedian) ~ CategoryTOC + Area + ReliefRatio + Aspect
                    + TOCAverage + Other15 + Conif15 + Imprgrass15 + Deer +
                     Sheep + pwsurplus81 + aat55for810 + pHMed + pwsurplus81:aat55for810 +
                     Sheep:pwsurplus81 + pwsurplus81:pHMed + CategoryTOC:Conif15
                    +  TOCAverage:Deer + ReliefRatio:Conif15 + ReliefRatio:Deer))
modelfinalBIC=step(modelfullBIC,k=log(n))
summary.lm(modelfinalBIC)
### This model explains 57%.
### Let's check behaviour of residuals.
par(mfrow=c(2,2))
plot(modelfinalBIC)
### Let's look at the predictive capability of the model
par(mfrow=c(1,1)) # reset plotting pane
```

```
plot(log(datacomp$TOCMedian),predict(modelfinalBIC)) # Plot the response against the predicted values of the response
### Formalise this as a linear model
modelfinalBICform <- with(datacomp,lm(predict(modelfinalBIC)~log(datacomp$TOCMedian)))
abline(modelfinalBICform) # This plots a straight line through the data by least-squares
### If the model predicted the response perfectly, then they would be linked by a straight line of slope 1
### If we insert such a line on the plot as a reference line we can make that comparison
abline(0,1)
### Validation
library(caret)
set.seed(0)
train_control <- trainControl(method = "repeatedcv", number = 10, repeats = 10)
modelBICVal <- train(log(TOCMedian) ~ ReliefRatio
                    + TOCAverage + Conif15 + Deer +
                      Sheep + pwsurplus81 + aat55for810 + pHMed + pwsurplus81:aat55for810 +
                      Sheep:pwsurplus81 + pwsurplus81:pHMed
                    +  TOCAverage:Deer + ReliefRatio:Conif15 + ReliefRatio:Deer, data = datacomp, method = "lm", trControl = train_control)
print(modelBICVal)
### RSME 0.35 R2 56% MAE 0.28
0.3464227/(max(log(datacomp$TOCMedian))-min(log(datacomp$TOCMedian))) #0.11
confint(modelfinalBIC)
## Let's get the predictions for future climate
dataFutureRed<-datacomp[,-c(230,231)]
names(dataFutureRed)[names(dataFutureRed)=="pwsurplus50"] <- "pwsurplus81"
names(dataFutureRed)[names(dataFutureRed)=="aat55for205"] <- "aat55for810"
prediction2050TOC<-predict(modelfinalBIC, newdata=dataFutureRed, interval="confidence")
PredictTOCInteractionsAll<-cbind(datacomp,prediction2050TOC)
write.csv(PredictTOCInteractionsAll, "Predictions2050TOCBIC_Interactions_All.csv")
predictionsCurrent<-predict(modelfinalBIC, newdata=datacomp, interval = "confidence")
```

```
PredictTOCInteractionsAll<-cbind(datacomp,predictionsCurrent)
write.csv(PredictTOCInteractionsAll, "PredictionsCurrentTOCBIC_Interactions_All.csv")
```

B.9 Linear regression models on *E. coli* medians

```
# Load data
data <- read.csv("DS2_SWCatchmentWaterQualitySubset.csv")
# In this analysis we treat EcoliMed as the response and introduce all variables at the same time plus a two-way interaction with one variable at a time
data$Source <- as.factor(data$Source)
data$Category_Ecoli <- as.factor(data$CategoryEcoli)
options(scipen=999)# Disable scientific notation
n=nrow(data)
# Model without interactions
modelfull<-
with(data,aov(EcoliMed~Source+Area+ElevationReliefRatio+ReliefRatio+Aspect+aat55for810+TOCAverage+BFIAvergae+Other15+Conif15+Arabl
e15+Imprgrass15+Urban15+Deer+Cattle+Sheep+SepticTank+pHMed+pwsurplus81))
modelfinalBIC=step(modelfull,k=log(n))
summary.lm(modelfinalBIC)
### This model explains 71%.
### Let's check behaviour of residuals.
par(mfrow=c(2,2))
plot(modelfinalBIC)
### Try log the response
modelfull<-
with(data,aov(log(EcoliMed+1)~Source+Area+ElevationReliefRatio+ReliefRatio+Aspect+aat55for810+TOCAverage+BFIAvergae+Other15+Conif15
+Arable15+Imprgrass15+Urban15+Deer+Cattle+Sheep+SepticTank+pHMed+pwsurplus81))
modelfinalBIC=step(modelfull,k=log(n))
summary.lm(modelfinalBIC)
### This model explains 46%.
### Let's check behaviour of residuals.
par(mfrow=c(2,2))
```

```
plot(modelfinalBIC)
### Let's look at the predictive capability of the model
par(mfrow=c(1,1)) # reset plotting pane
plot(data$EcoliMed,exp(predict(modelfinalBIC)-1)) # Plot the response against the predicted values of the response
### Formalise this as a linear model
modelfinalBICform <- with(data,lm(exp(predict(modelfinalBIC)-1)~data$EcoliMed))
abline(modelfinalBICform) # This plots a straight line through the data by least-squares
### If the model predicted the response perfectly, then they would be linked by a straight line of slope 1
### If we insert such a line on the plot as a reference line we can make that comparison
abline(0,1)
### Validation
library(caret)
set.seed(0)
train_control <- trainControl(method = "repeatedcv", number = 10, repeats = 10)
modelBICVal <- train(log(EcoliMed+1) ~ Source + ReliefRatio + Other15 +
                    Conif15 + Arable15 + Imprgrass15, data = data, method = "lm", trControl = train_control)
print(modelBICVal)
### RSME 0.94 R2 41% MAE 0.78
0.9350434/(max(log(data$EcoliMed+1))-min(log(data$EcoliMed+1))) #0.15
confint(modelfinalBIC)
# Model with LCA
modelfullLCA<-
with(data,aov(log(EcoliMed+1)~Source+Area+ElevationReliefRatio+ReliefRatio+Aspect+SepticTank+PrimeLand8120_v2.4+LCA3_5_8120_v2.4))
modelfinalLCA=step(modelfullLCA,k=log(n))
summary.lm(modelfinalLCA)
### This model explains 32%.
### Run with PrimeLand including class 3
```

```
modelfullLCA<-
with(data,aov(log(EcoliMed+1)~Source+Area+ElevationReliefRatio+ReliefRatio+Aspect+SepticTank+PrimeLandIncl3_8120_v2.4+LCA4_5_8120_v
2.4))
modelfinalLCA=step(modelfullLCA,k=log(n))
summary.lm(modelfinalLCA)
### This model explains 36%.
confint(modelfinalLCA)
### Let's check behaviour of residuals.
par(mfrow=c(2,2))
plot(modelfinalLCA)
### Let's look at the predictive capability of the model
par(mfrow=c(1,1)) # reset plotting pane
plot(data$EcoliMed,exp(predict(modelfinalLCA)-1)) # Plot the response against the predicted values of the response
plot(log(data$EcoliMed+1),predict(modelfinalLCA)) # Plot the response against the predicted values of the response
### Formalise this as a linear model
modelfinalBICform <- with(data,lm(predict(modelfinalLCA)~log(data$EcoliMed+1)))
abline(modelfinalBICform) # This plots a straight line through the data by least-squares
### If the model predicted the response perfectly, then they would be linked by a straight line of slope 1
### If we insert such a line on the plot as a reference line we can make that comparison
abline(0,1)
### Validation
library(caret)
set.seed(0)
train_control <- trainControl(method = "repeatedcv", number = 10, repeats = 10)
modelBICVal <- train(log(EcoliMed + 1) ~ Source + ReliefRatio + PrimeLandIncl3_8120_v2.4 +
                     LCA4_5_8120_v2.4, data = data, method = "lm", trControl = train_control)
print(modelBICVal)
### RSME 0.997 R2 33% MAE 0.82
```

```
0.9971493/(max(log(data$EcoliMed+1))-min(log(data$EcoliMed+1))) #0.16
confint(modelfinalLCA)
dataFutureRed<-data[,-c(59,60)]
names(dataFutureRed)[names(dataFutureRed)=="PrimeLandIncl3_2050mean"] <- "PrimeLandIncl3_8120_v2.4"
names(dataFutureRed)[names(dataFutureRed)=="LCA4_5_2050mean"] <- "LCA4_5_8120_v2.4"
prediction2050Ecoli<-predict(modelfinalLCA, newdata=dataFutureRed, interval="confidence")
PredictEcoli<-cbind(data,prediction2050Ecoli)
write.csv(PredictEcoli, "Predictions2050EcoliBIC_LCA.csv")
predictionsCurrent<-predict(modelfinalLCA, newdata=data, interval = "confidence")
PredictEcoliAll<-cbind(data,predictionsCurrent)
write.csv(PredictEcoliAll, "PredictionsCurrentEcoliBIC_All.csv")
### Extract residuals
mod_summary<-summary.lm(modelfinalLCA)
write.csv(mod_summary$residuals, "EcoliModelResiduals.csv")
# Model with reduced dataset (E.coli medians <100)
datared<-data[data$EcoliMed<100,]
# Follow steps as above
# Model with interactions as in B.8
```

B.10 Rainfall sensitivity testing for *E. coli*

```
# load package and data
library("lubridate")
sens<-read.csv("DS6_EcoliRainfallSeries20102016.csv")
# Seperate by season
# Create a column with months first
sens$Date<-format(sens$Date)
sens$month<-month(sens$Date)
sens$season<-ifelse(sens$month==1, "winter", ifelse(sens$month==2, "winter", ifelse(sens$month==3, "summer", ifelse(sens$month==4, "summer",
ifelse(sens$month==5, "summer", ifelse(sens$month==6, "summer", ifelse(sens$month==7, "summer", ifelse(sens$month==8, "summer",
ifelse(sens$month==9, "summer", ifelse(sens$month==10, "summer", ifelse(sens$month==11, "winter", ifelse(sens$month==12, "winter",
NA))))))))))))
# Seperate
sensWinter<-sens[which(sens$season=="winter"),]
sensSummer<-sens[which(sens$season=="summer"),]
# Spearman's correlation test per catchment - all together
# Split catchments
catch<-split(sens, sens$Catchment.ID) # adjust dataset as required
# Rainfall
# Test function
regr<-cor.test(~Ecoli+Rainfall, data = catch[[i]], method="spearman", continuity=FALSE, conf.level=0.95)$estimate
#Run loop for Spearman's rank correlation coeffiencient, adjust variable as required
for (i in names(catch)){
  regr[i]<-cor.test(~Ecoli+Rainfall, data = catch[[i]], method="spearman", continuity=FALSE, conf.level=0.95)$estimate
  print(regr[i])
}
# $p.value for significance
```

C. Model outputs

C.1 Model outputs from the colour-TOC/iron/manganese/turbidity linear regression

Table C.1: Model output for linear regression on colour as dependent variable and TOC as independent variable for data from 2013-2016. Coloured cells include coefficients that are statistically significant ($p<0.05$).

Catchment ID	**R^2**	**Intercept**	**TOC**
388	0.99	-8.84	10.25
393	0.99	-5.63	10.39
425	0.97	-2.52	6.00
407	0.97	-5.87	10.22
276	0.96	-3.09	7.75
338	0.96	1.27	6.32
451	0.95	5.67	7.45
456	0.95	-7.46	10.43
411	0.95	-2.57	7.77
299	0.94	3.15	6.35
21	0.94	1.67	7.81
268	0.94	-2.04	8.06
33	0.94	-1.69	7.72
427	0.93	-4.31	9.06
412	0.92	-0.16	6.66
273	0.92	-0.12	7.84
446	0.92	-6.80	8.72
280	0.92	-2.36	7.44
398	0.91	0.25	7.51
277	0.91	-0.09	7.20
442	0.91	-14.38	10.25
16	0.91	-6.83	8.26
296	0.91	-0.72	5.12
416	0.90	1.33	6.78
278	0.89	2.18	6.77
376	0.89	-9.62	10.18
335	0.87	-0.31	7.19
334	0.87	-0.39	9.32
34	0.86	-3.09	8.38
384	0.86	-3.90	9.03
75	0.86	-0.48	6.42
343	0.86	1.36	7.48
432	0.85	-8.37	7.74
342	0.85	4.85	6.21
339	0.85	1.76	8.01
271	0.85	5.63	7.07
25	0.84	2.06	7.62
418	0.82	0.61	6.97
337	0.82	-1.42	7.69
297	0.81	6.91	6.52
365	0.81	-6.27	5.75
330	0.81	-34.27	8.58
445	0.81	-10.94	11.44
89	0.81	0.46	7.08
434	0.80	-11.35	8.80
450	0.80	1.10	6.64
433	0.80	-0.81	6.87
437	0.79	-5.10	7.00

401	0.79	-6.36	9.96
45	0.79	7.54	7.00
340	0.79	0.66	6.46
2	0.79	-9.06	8.86
448	0.78	-2.42	4.21
266	0.78	5.58	5.67
279	0.76	3.64	6.30
15	0.76	5.01	5.71
212	0.74	2.56	6.54
208	0.73	-1.16	8.01
20	0.73	1.70	5.38
270	0.73	2.79	6.60
211	0.73	-1.75	7.06
447	0.72	-23.70	10.65
346	0.71	-7.60	6.77
7	0.71	1.42	6.28
17	0.69	-11.95	6.58
422	0.66	8.94	5.90
9	0.66	11.86	4.94
222	0.65	0.84	6.58
436	0.64	-11.85	8.41
402	0.64	1.09	6.65
305	0.63	1.99	6.17
126	0.62	5.26	6.55
59	0.62	-0.27	6.64
357	0.61	-6.38	8.05
429	0.61	-5.80	7.39
80	0.61	11.78	6.81
136	0.61	-2.16	7.12
201	0.61	30.28	7.70
146	0.60	15.57	8.26
32	0.55	8.21	4.37
344	0.53	-2.59	5.15
103	0.52	8.22	6.23
193	0.51	14.10	5.42
192	0.49	8.40	5.28
440	0.47	11.07	4.11
43	0.47	8.36	4.53
6	0.46	-6.03	7.53
213	0.45	1.31	5.88
410	0.45	14.44	3.06
424	0.44	5.85	3.40
415	0.42	20.65	3.89
138	0.40	13.51	3.88
320	0.40	4.81	3.65
232	0.37	21.39	5.04
349	0.36	26.38	2.88
24	0.34	-2.41	4.35
227	0.34	23.05	4.30
368	0.33	1.35	3.47
311	0.31	33.43	2.49
426	0.31	6.66	2.56
350	0.28	44.84	4.04
203	0.27	32.28	3.40
42	0.26	14.48	2.98
319	0.25	3.77	3.91
373	0.19	45.81	4.06
116	0.18	12.48	2.14
22	0.18	21.16	3.16
356	0.17	8.47	0.12

215	0.15	29.63	1.53
23	0.14	5.54	1.53
312	0.13	107.56	2.56
310	0.13	37.29	0.65
452	0.12	20.42	2.20
321	0.11	19.04	1.48
30	0.10	18.20	1.38
341	0.10	13.89	1.43
327	0.10	24.26	2.31
333	0.08	18.33	1.29
167	0.07	24.69	0.57
325	0.07	48.18	2.13
131	0.07	17.69	0.33
242	0.07	17.33	0.41
127	0.05	59.26	1.03
414	0.03	16.39	0.40
214	0.01	62.33	0.42
439	0.01	61.10	-0.28
449	0.01	2.19	0.04

Table C.2: Model output for linear regression on colour as dependent variable and TOC, Iron and Manganese as independent variable for data from 2013-2016. Coloured cells include coefficients that are statistically significant ($p<0.05$).

Catchment ID	R^2	Intercept	Iron	Manganese	TOC
388	1.00	-1.34	0.20	-1.22	4.18
393	0.99	-4.99	0.09	-0.79	7.10
407	0.99	3.18	0.13	-1.95	4.69
451	0.98	4.49	0.09	-0.70	6.14
456	0.98	8.32	0.14	-2.61	3.72
425	0.97	-2.35	0.00	-0.06	5.93
276	0.97	1.33	0.07	-3.42	5.60
338	0.97	1.67	0.06	-0.53	5.49
299	0.96	-0.16	0.08	-3.07	5.18
296	0.96	-0.85	0.00	-0.26	6.37
268	0.96	-2.31	0.06	-1.41	6.75
21	0.95	-3.44	0.02	-0.16	7.47
33	0.95	-2.32	0.01	-0.05	7.31
411	0.95	-2.61	0.01	-0.03	7.62
418	0.95	-4.99	0.00	-0.23	9.80
273	0.94	-0.74	0.01	-0.28	7.34
335	0.94	-0.19	0.14	-0.48	4.04
446	0.94	-2.63	0.08	-0.25	6.17
427	0.93	-2.73	0.01	-0.24	8.42
412	0.93	-0.23	0.01	-0.05	6.20
278	0.93	-0.78	0.01	-0.05	6.06
280	0.93	-2.97	0.05	-0.65	6.49
6	0.92	9.95	0.12	-0.11	1.43
277	0.92	-0.75	0.01	-0.03	7.34
432	0.92	-4.72	0.12	-0.44	4.34
16	0.92	-8.12	0.01	-0.02	7.60
342	0.91	7.50	0.06	-0.56	5.10
442	0.91	-9.58	0.03	-0.13	8.72
398	0.91	0.02	0.01	-0.05	7.42
416	0.91	-0.11	0.01	0.01	6.15
376	0.91	-8.50	0.04	-0.23	8.75
266	0.91	0.38	0.17	-0.54	3.45
271	0.90	12.49	0.04	-0.27	5.15
340	0.89	4.76	0.11	-0.59	4.77
334	0.88	0.26	0.03	-0.52	8.03
434	0.88	1.02	0.15	-0.57	2.91
330	0.88	-21.63	0.04	-0.09	5.86
343	0.88	2.68	0.03	-1.03	7.14
337	0.87	-2.44	0.06	-0.22	5.61
339	0.87	0.32	0.03	-0.03	6.60
15	0.87	5.06	0.00	-0.18	7.10
384	0.86	-8.60	0.01	-0.01	8.51
34	0.86	-2.47	0.01	-0.01	7.78
365	0.86	-9.86	0.02	-0.03	5.71
75	0.86	-0.30	-0.01	-0.02	6.48
25	0.86	0.68	-0.01	0.03	8.11
445	0.85	21.44	0.07	-2.17	6.51
297	0.85	1.79	0.00	0.21	5.64
450	0.83	-1.88	0.05	-0.14	6.68
222	0.83	3.80	0.04	-0.18	5.38
270	0.82	-2.48	0.21	-0.37	4.31
437	0.82	-5.27	0.05	-0.13	6.09
89	0.81	0.08	0.00	-0.01	7.13
433	0.81	-1.36	0.02	-0.44	7.12
401	0.81	-3.31	0.09	-0.90	6.47

45	0.80	6.57	0.02	-0.01	6.04
402	0.79	0.52	0.08	-0.76	5.85
279	0.79	1.36	0.06	-0.19	5.10
448	0.79	-0.76	0.01	-0.06	3.48
429	0.79	-19.41	0.01	-0.14	9.23
2	0.79	-8.64	0.00	-0.01	8.80
346	0.78	-5.30	0.11	-0.66	5.11
208	0.76	-0.58	0.17	-0.08	3.94
320	0.75	6.03	0.14	-0.57	2.80
7	0.75	0.41	0.00	0.00	6.76
20	0.75	0.78	0.02	-0.03	5.12
212	0.74	1.89	0.01	-0.03	6.49
447	0.74	-14.71	0.04	-0.19	8.24
357	0.74	6.70	0.12	-0.35	5.30
211	0.73	-3.04	0.01	0.00	6.89
80	0.72	5.46	0.03	-0.05	6.31
17	0.71	-16.54	0.00	0.01	7.08
213	0.71	-2.17	0.02	-0.06	6.53
59	0.70	5.56	0.03	-0.08	5.24
136	0.70	-5.93	0.00	-0.03	8.35
422	0.70	7.74	0.01	-0.04	5.09
203	0.70	15.60	-0.02	-0.03	6.89
9	0.69	13.41	0.01	-0.10	4.62
344	0.68	0.52	0.19	-0.07	2.91
192	0.68	7.48	0.06	-0.26	4.67
126	0.68	3.79	0.00	-0.04	6.87
146	0.66	33.37	0.02	-0.33	6.68
311	0.64	27.92	0.10	-0.45	0.92
436	0.64	-5.14	0.06	-0.38	5.91
426	0.64	3.75	0.09	-0.98	1.56
305	0.64	2.22	0.00	0.00	6.30
201	0.62	32.34	0.01	0.00	6.83
410	0.62	9.40	0.00	-0.01	4.98
349	0.62	2.60	0.01	-0.24	6.98
193	0.58	12.48	0.05	-0.13	4.85
215	0.57	26.12	0.09	-0.09	0.72
319	0.57	10.35	0.05	-0.18	1.89
32	0.55	7.69	0.01	-0.02	4.04
24	0.55	5.06	0.02	-0.09	2.87
312	0.55	31.52	-0.01	0.00	8.11
415	0.54	15.82	0.02	-0.14	4.89
103	0.54	18.79	0.06	-0.18	2.33
440	0.54	1.49	-0.08	-0.08	7.78
232	0.52	17.24	0.01	-0.03	5.43
368	0.52	-1.54	0.13	-0.37	2.17
310	0.51	32.52	0.24	-0.43	0.27
43	0.49	7.41	0.01	-0.02	4.66
452	0.44	22.30	0.08	-0.57	1.21
424	0.44	5.80	0.00	0.00	3.46
350	0.43	8.10	-0.01	0.03	8.19
138	0.42	12.54	0.00	-0.01	4.13
449	0.42	0.51	0.19	0.14	-0.02
22	0.40	19.16	0.04	-0.04	1.65
23	0.38	3.44	0.01	-0.02	2.11
325	0.36	33.91	0.15	-0.04	0.60
227	0.35	22.84	0.00	-0.01	4.28
42	0.30	13.55	0.02	-0.02	2.91
321	0.30	18.70	0.09	-0.85	0.95
414	0.29	8.93	0.00	-0.01	2.89
30	0.29	16.79	0.04	-0.23	1.45

356	0.25	10.29	0.00	0.06	-0.36
116	0.23	11.92	0.01	-0.06	2.19
333	0.22	15.88	0.08	-0.13	1.03
373	0.20	44.07	0.03	-0.09	3.27
167	0.19	21.39	0.04	-0.10	0.49
341	0.19	11.50	0.11	-0.43	1.21
214	0.15	58.06	0.00	0.00	0.37
127	0.15	50.11	0.02	-0.10	1.22
327	0.12	23.61	0.02	-0.20	2.32
242	0.11	15.41	0.00	0.00	0.98
131	0.10	17.23	0.00	0.00	0.57
439	0.03	65.11	0.01	-0.15	-0.38

Table C.3: Model output for linear regression on colour as dependent variable and TOC, Iron, Manganese and Turbidity as independent variable for data from 2013-2016. Coloured cells include coefficients that are statistically significant ($p<0.05$).

Catchment ID	R^2	Intercept	Iron	Manganese	TOC	Turbidity
388	1.00	-0.03	0.20	-0.96	4.20	-6.77
393	0.99	-4.15	0.09	-0.75	7.09	-4.78
407	0.99	4.59	0.14	-0.69	4.20	-19.36
451	0.98	4.50	0.09	-0.70	6.15	-0.08
456	0.98	8.60	0.14	-2.61	3.69	-1.00
276	0.97	0.05	0.08	-1.40	5.45	-4.47
338	0.97	1.45	0.07	-0.42	5.47	-2.08
425	0.97	-2.35	0.00	-0.07	5.91	-0.10
296	0.97	-2.08	0.01	-0.18	6.07	-1.02
299	0.96	-0.06	0.09	-2.81	4.71	-3.83
278	0.96	-1.16	0.02	-0.04	6.72	-5.55
268	0.96	-2.47	0.06	-1.29	6.78	-0.80
21	0.95	-3.72	0.02	-0.17	7.41	1.30
411	0.95	-2.47	0.00	-0.07	7.42	4.97
418	0.95	-4.70	0.01	-0.34	9.77	-0.29
33	0.95	-2.50	0.01	-0.05	7.35	0.31
273	0.94	1.77	0.01	-0.23	7.56	-4.09
427	0.94	-2.01	0.01	-0.21	8.72	-5.04
335	0.94	-0.01	0.16	-0.44	4.02	-2.90
446	0.94	-5.76	0.07	-0.23	6.20	6.04
280	0.94	-1.66	0.05	-0.61	6.34	-0.97
6	0.94	11.01	0.14	0.00	1.45	-9.75
412	0.93	-0.17	0.01	-0.06	6.09	-0.17
15	0.93	0.56	0.01	-0.10	6.76	-1.01
432	0.93	-4.14	0.11	-0.41	4.66	-3.10
277	0.93	-1.75	0.01	-0.04	7.33	1.91
16	0.93	-6.44	0.02	-0.04	7.30	-4.67
442	0.92	-6.53	0.04	-0.03	8.50	-8.42
342	0.92	6.55	0.06	-0.59	4.89	3.24
376	0.91	-7.28	0.07	-0.22	8.08	-5.06
398	0.91	0.03	0.01	-0.06	7.40	-0.17
416	0.91	0.08	0.01	0.01	6.12	-1.25
266	0.91	-0.04	0.17	-0.48	3.47	-0.77
271	0.90	11.53	0.04	-0.30	5.16	2.65
434	0.90	8.64	0.18	-0.38	2.52	-33.14
340	0.89	4.83	0.11	-0.57	4.84	-1.61
334	0.88	0.62	0.03	-0.47	8.00	-1.40
330	0.88	-21.45	0.04	-0.09	5.85	-0.26
343	0.88	2.02	0.04	-0.56	7.32	-4.62
337	0.87	-1.08	0.07	-0.22	5.25	-2.69
445	0.87	14.61	0.10	-1.18	5.59	-11.67
75	0.87	-2.51	-0.03	-0.05	6.06	14.68
384	0.87	-15.22	0.00	0.01	9.16	6.47
339	0.87	-0.57	0.03	-0.05	6.69	1.06
34	0.87	1.90	0.01	-0.01	7.44	-2.51
365	0.86	-10.27	0.02	-0.03	5.69	0.99
433	0.86	-7.93	0.01	-0.74	5.99	31.82
25	0.86	-0.07	-0.01	0.03	8.07	1.71
297	0.85	1.47	0.01	0.19	5.31	-1.93
222	0.85	3.95	0.06	-0.20	4.81	-1.28
450	0.84	-0.67	0.07	-0.12	6.60	-6.28
270	0.83	-5.85	0.19	-0.44	4.52	11.96
45	0.83	7.90	0.03	0.00	5.89	-3.30
401	0.83	-0.19	0.11	-0.40	6.87	-39.88
437	0.82	-4.39	0.05	-0.12	6.05	-1.72

89	0.81	0.26	0.00	-0.01	7.13	-0.09
447	0.81	-3.92	0.07	-0.17	7.17	-13.75
2	0.80	-4.52	0.01	-0.01	8.39	-3.32
422	0.80	8.85	0.03	-0.09	5.74	-13.73
402	0.80	-0.64	0.06	-0.77	6.08	2.97
346	0.80	-3.87	0.13	-0.58	5.07	-8.81
429	0.80	-18.47	0.02	-0.19	8.73	-1.61
279	0.79	1.37	0.06	-0.19	5.10	-0.09
448	0.79	-0.36	0.01	-0.05	3.41	-1.29
357	0.79	0.63	0.05	-0.23	6.28	5.37
320	0.78	8.42	0.14	-0.60	2.40	-3.10
80	0.78	12.01	0.07	-0.11	3.75	-2.56
208	0.76	-0.81	0.17	-0.07	4.00	1.19
20	0.75	0.51	0.01	-0.02	5.12	1.02
7	0.75	0.00	0.00	0.00	6.75	1.56
212	0.75	2.36	0.01	-0.02	6.58	-1.71
211	0.73	-2.62	0.01	-0.01	6.82	-0.10
203	0.72	12.18	0.00	-0.04	6.37	-0.64
213	0.72	-2.84	0.02	-0.08	6.41	2.33
9	0.72	14.62	0.02	-0.13	4.30	-0.41
449	0.71	2.52	-0.11	-0.01	-0.03	1.10
17	0.71	-16.88	0.00	0.01	7.08	0.42
59	0.70	5.61	0.03	-0.08	5.26	0.14
136	0.70	-5.94	0.00	-0.03	8.35	0.00
32	0.70	5.82	0.06	-0.05	2.87	-7.19
426	0.69	1.52	0.08	-1.38	1.01	15.54
192	0.69	7.97	0.08	-0.30	4.30	-0.80
126	0.69	3.47	0.01	-0.03	7.08	-1.28
344	0.68	0.35	0.19	-0.08	2.90	1.10
305	0.67	1.77	0.00	-0.01	6.39	-0.37
146	0.66	32.98	0.02	-0.33	6.64	1.02
436	0.66	-3.73	0.09	-0.21	5.58	-9.62
201	0.65	50.72	0.02	0.00	6.13	-4.30
311	0.65	27.93	0.10	-0.48	0.63	1.60
410	0.64	7.72	0.00	-0.02	4.89	2.43
349	0.62	2.61	0.01	-0.24	6.98	0.00
193	0.60	10.92	0.03	-0.11	5.28	2.39
415	0.59	15.01	0.03	-0.13	4.96	-2.58
215	0.58	24.28	0.09	-0.11	0.72	2.55
138	0.57	2.86	0.00	0.00	6.09	-0.55
319	0.57	10.26	0.05	-0.16	2.08	-2.61
424	0.57	3.53	0.00	0.00	4.85	-1.23
312	0.56	26.34	0.00	-0.01	8.03	-2.06
368	0.55	-2.54	0.10	-0.43	2.01	4.54
24	0.55	5.21	0.02	-0.09	2.87	-0.23
103	0.54	20.27	0.06	-0.18	2.20	-1.08
440	0.54	2.09	-0.08	-0.05	7.78	-1.94
232	0.53	17.43	0.01	-0.03	5.41	-0.07
310	0.52	32.56	0.25	-0.44	0.27	-0.98
43	0.50	7.25	0.01	-0.02	4.65	0.16
452	0.47	22.21	0.10	-0.47	1.11	-1.58
22	0.45	19.86	0.06	-0.05	1.62	-3.82
350	0.43	7.60	-0.01	0.02	8.24	0.04
227	0.40	19.85	0.01	-0.02	5.06	-1.29
23	0.39	5.00	0.01	-0.02	1.77	0.18
325	0.39	36.43	0.17	-0.05	0.80	-10.90
356	0.31	9.19	0.00	0.07	-0.24	0.07
42	0.31	13.60	0.02	-0.02	2.97	-0.31
414	0.30	8.43	0.00	-0.02	3.06	-0.27
321	0.30	18.75	0.09	-0.84	0.96	-0.53

30	0.30	15.94	0.04	-0.25	1.43	2.83
127	0.26	38.47	0.02	-0.08	2.85	-1.25
341	0.23	11.77	0.11	-0.49	0.98	2.45
116	0.23	11.96	0.01	-0.06	2.17	-0.06
333	0.23	15.81	0.08	-0.13	0.99	0.81
167	0.22	21.42	0.04	-0.09	0.54	-1.05
373	0.21	45.33	0.03	-0.09	3.20	1.25
439	0.17	67.65	-0.01	-0.08	-0.77	0.32
214	0.15	57.73	0.00	0.00	0.40	-0.05
131	0.15	16.45	0.00	0.00	0.99	-0.11
327	0.13	23.05	0.02	-0.24	2.20	4.64
242	0.11	15.39	0.00	0.00	0.98	0.01

C.2 Shape-based clustering outcomes and catchment cluster allocation

Table C.4: Catchment allocation to cluster, for reservoirs and lochs. Clustering based on seasonal profile extracted from time series 2011-2016 with daily values interpolated, logged, and means taken for every day of the year. The k-shape algorithm was used and run five times with k=2. The two resulting clusters were identified as either “sine” (seasonal profile looking as expected) or “v-shape” (seasonal profile with dip in summer and peak in winter) and catchment allocation to cluster noted. Final group was allocated according to which cluster the catchment fell into predominantly.

Water source					Cluster run	Allocated group
	1	2	3	4	5	
Reservoirs						
1	sine	sine	sine	sine	vshape	sine
2	sine	vshape	sine	sine	vshape	sine
3	sine	sine	sine	sine	sine	sine
4	sine	vshape	sine	sine	vshape	sine
5	sine	sine	sine	sine	sine	sine
6	sine	sine	sine	sine	sine	sine
7	sine	vshape	sine	sine	vshape	sine
8	sine	sine	sine	sine	sine	sine
9	vshape	vshape	vshape	vshape	vshape	vshape
10	vshape	vshape	vshape	vshape	vshape	vshape
11	vshape	vshape	vshape	vshape	vshape	vshape
12	sine	vshape	sine	sine	sine	sine
13	sine	sine	sine	sine	sine	sine
14	vshape	vshape	vshape	vshape	vshape	vshape
15	sine	sine	sine	sine	sine	sine
16	sine	sine	sine	sine	sine	sine
17	sine	sine	sine	sine	sine	sine
18	sine	sine	sine	sine	sine	sine
19	vshape	vshape	vshape	vshape	vshape	vshape
20	sine	sine	sine	sine	sine	sine
21	vshape	vshape	vshape	vshape	vshape	vshape
22	sine	vshape	sine	sine	vshape	sine
23	sine	sine	sine	sine	sine	sine
24	sine	sine	sine	sine	sine	sine
25	sine	sine	sine	sine	sine	sine
26	sine	sine	sine	sine	sine	sine
27	vshape	vshape	vshape	vshape	vshape	vshape
28	sine	sine	sine	sine	sine	sine
29	vshape	vshape	vshape	vshape	vshape	vshape
30	vshape	vshape	vshape	vshape	vshape	vshape
31	sine	sine	sine	sine	sine	sine
32	sine	sine	sine	sine	sine	sine
33	sine	sine	sine	sine	sine	sine
34	vshape	vshape	sine	vshape	vshape	vshape
35	sine	vshape	sine	sine	vshape	sine
36	sine	sine	sine	sine	sine	sine
37	sine	sine	sine	sine	sine	sine
38	vshape	vshape	vshape	vshape	vshape	vshape
39	sine	sine	sine	sine	sine	sine
40	sine	vshape	sine	sine	vshape	sine
41	sine	sine	sine	sine	sine	sine
42	vshape	vshape	vshape	vshape	vshape	vshape

43	sine	sine	sine	sine	sine	sine
44	vshape	vshape	sine	vshape	vshape	vshape
45	vshape	vshape	vshape	vshape	vshape	vshape
46	sine	sine	sine	sine	sine	sine
47	sine	sine	sine	sine	sine	sine
48	sine	sine	sine	sine	sine	sine
49	sine	vshape	sine	sine	vshape	sine
50	sine	vshape	sine	sine	vshape	sine
51	vshape	vshape	sine	vshape	vshape	vshape
52	sine	vshape	sine	sine	vshape	sine
53	sine	sine	sine	sine	sine	sine
54	sine	sine	sine	sine	sine	sine
55	vshape	vshape	sine	vshape	vshape	vshape
56	sine	vshape	sine	sine	vshape	sine
57	sine	sine	sine	sine	sine	sine
58	sine	sine	sine	sine	sine	sine
59	sine	sine	sine	sine	sine	sine
60	sine	vshape	sine	sine	vshape	sine
61	sine	sine	sine	sine	sine	sine
62	sine	sine	sine	sine	sine	sine
63	vshape	vshape	sine	vshape	vshape	vshape
64	vshape	vshape	vshape	vshape	vshape	vshape
65	sine	sine	sine	sine	sine	sine
Lochs						
1	sine	sine	sine	sine	sine	sine
2	sine	sine	vshape	vshape	vshape	vshape
3	sine	sine	vshape	vshape	vshape	vshape
4	vshape	vshape	sine	sine	sine	inconclusive
5	sine	sine	sine	sine	sine	sine
6	sine	sine	sine	sine	sine	sine
7	sine	sine	vshape	vshape	vshape	vshape
8	sine	sine	sine	sine	sine	sine
9	sine	sine	sine	sine	sine	sine
10	sine	sine	sine	sine	sine	sine
11	vshape	vshape	vshape	vshape	vshape	vshape
12	sine	vshape	sine	vshape	vshape	inconclusive
13	sine	sine	sine	sine	sine	sine
14	sine	sine	sine	sine	sine	sine
15	sine	sine	sine	sine	sine	sine
16	vshape	vshape	vshape	vshape	vshape	vshape
17	sine	sine	sine	sine	sine	sine
18	sine	sine	sine	sine	sine	sine
19	sine	sine	sine	sine	sine	sine
20	sine	sine	sine	sine	sine	sine
21	sine	sine	vshape	vshape	vshape	vshape
22	sine	sine	sine	sine	sine	sine
23	sine	sine	vshape	vshape	vshape	vshape
24	sine	sine	sine	sine	sine	sine
25	sine	sine	vshape	vshape	vshape	vshape
26	sine	sine	vshape	vshape	vshape	vshape
27	sine	sine	sine	sine	sine	sine
28	sine	sine	vshape	vshape	vshape	inconclusive
29	vshape	vshape	vshape	vshape	vshape	vshape
30	sine	sine	sine	sine	sine	sine

31	sine	sine	sine	sine	sine	sine
32	sine	sine	vshape	vshape	vshape	vshape
33	sine	sine	sine	sine	sine	sine

C.3 TOC climate sensitivity relationships and group allocations

Table C.5: Results of Spearman's rank correlation tests run per catchment on sampled TOC concentrations from 2013 – 2016 and corresponding 3-day antecedent rainfall (short rainfall), 60-day antecedent rainfall (long rainfall), and 60-day antecedent mean temperature (temperature), split by season. Rho values are given with highlights indicating statistical significance ($p<0.05$). Included are also season of maximum and minimum mean TOC concentrations, overall TOC median and the ratio between the concentration median and concentration range per catchment.

	TOC concentrations				Summer			Autumn			Winter			Spring			all			Category
	Max	Min	Med	Rat	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	
1	Sum	Win	4.35	0.85	0.77	0.58	0.03	0.47	-0.30	-0.15	0.15	-0.01	0.43	0.47	0.10	0.24	0.38	-0.09	0.25	Rainfall
2	Sum	Spr	3.85	0.86	0.69	0.36	0.76	0.74	0.78	-0.15	0.02	0.21	-0.07	0.43	-0.11	0.76	0.31	-0.04	0.58	Rainfall+Temp
3	Sum	Win	2.9	1.09	0.75	0.12	0.68	0.76	0.05	-0.05	-0.03	0.13	0.09	0.76	-0.34	0.07	0.55	-0.17	0.39	Rainfall+Temp
4	Aut	Sum	7.4	0.9	0.26	0.43	0.39	-0.03	-0.03	0.16	0.08	-0.33	0.66	0.30	-0.26	0.62	-0.01	-0.39	0.72	Wet up
5	Sum	Win	3.9	0.72	0.73	-0.04	0.36	-0.37	-0.33	0.76	0.38	-0.57	0.44	0.59	-0.15	0.71	0.11	-0.58	0.84	Rainfall+Temp
6	Aut	Sum	8.75	0.24	0.17	0.61	0.15	0.25	0.50	-0.67	-0.23	-0.43	0.53	0.06	0.38	0.24	0.07	0.28	0.13	Wet up
7	Sum	Spr	1.5	2.33	0.72	0.53	0.52	0.57	0.13	0.08	0.49	-0.47	0.27	0.79	0.20	0.43	0.67	0.10	0.41	Rainfall
8	Aut	Spr	6.5	0.6	0.54	-0.11	0.37	0.41	0.40	0.27	0.24	-0.19	0.50	0.76	-0.76	0.49	0.09	-0.55	0.70	Rainfall
9	Sum	Win	1.1	2.17	0.67	-0.03	0.67	0.73	-0.56	-0.19	0.34	-0.35	0.65	0.43	-0.54	-0.11	0.47	-0.48	0.35	Rainfall+Temp
10	Sum	Win	2.4	0.99	0.72	-0.31	-0.25	0.39	-0.08	0.29	-0.12	0.13	0.16	0.55	0.19	0.31	0.39	-0.26	0.35	Rainfall
11	Sum	Win	4.45	0.53	0.63	0.22	0.27	0.73	0.12	0.63	0.25	-0.39	0.51	0.46	0.04	0.52	0.35	-0.25	0.62	Rainfall+Temp
12	Sum	Win	3.6	0.71	0.81	-0.14	0.82	0.51	-0.37	0.16	0.05	-0.16	0.45	0.58	-0.41	0.12	0.41	-0.47	0.59	Rainfall+Temp
13	Sum	Win	3.6	0.86	0.84	-0.12	0.39	0.64	-0.14	0.69	0.27	0.70	0.64	0.23	0.00	0.79	0.20	-0.41	0.78	Rainfall+Temp
14	Win	Spr	1.7	1.3	0.68	-0.04	-0.08	0.88	0.37	-0.14	0.67	-0.32	0.62	0.93	0.20	0.48	0.76	0.09	0.15	Rainfall
15	Sum	Win	12.25	0.4	0.58	0.32	0.70	0.11	0.12	0.61	0.38	-0.35	0.65	-0.11	-0.35	0.61	-0.08	-0.54	0.85	Rainfall+Temp
16	Sum	Spr	4.8	0.54	0.37	-0.17	0.23	0.38	0.69	-0.10	0.00	-0.06	0.82	-0.35	-0.21	-0.35	-0.02	-0.10	0.52	Wet up
17	Sum	Spr	5.8	0.57	0.20	-0.83	0.34	0.12	-0.22	0.00	0.31	-0.44	0.13	0.63	-0.07	0.83	0.06	-0.58	0.58	Rainfall+Temp
18	Sum	Win	3.1	0.83	0.73	-0.42	0.61	0.55	0.00	0.73	-0.07	-0.10	0.32	0.73	0.21	0.49	0.46	-0.32	0.66	Rainfall+Temp
19	Sum	Spr	3.7	0.62	-0.04	0.44	0.08	-0.23	-0.23	0.69	-0.19	-0.41	0.60	0.20	-0.30	0.52	-0.09	-0.50	0.78	Temp
20	Aut	Spr	9.8	0.23	-0.02	0.76	0.14	0.27	0.48	-0.65	-0.04	-0.04	0.70	0.04	0.39	0.32	0.15	0.50	-0.05	Wet up
21	Win	Spr	6.1	0.49	0.01	-0.20	0.46	0.25	0.47	-0.32	0.74	-0.35	-0.18	0.20	0.10	0.28	0.46	0.25	0.26	Rainfall
22	Sum	Sum	3.5	0.51	0.21	0.16	0.26	0.64	0.38	-0.47	-0.20	-0.20	0.35	0.34	-0.37	0.56	0.41	0.01	0.07	Rainfall
23	Aut	Spr	6.5	0.22	-0.07	0.16	0.19	0.12	0.13	-0.35	-0.08	-0.26	0.52	-0.11	0.12	-0.05	0.09	0.15	0.01	None
24	Aut	Win	7.3	0.7	0.27	0.76	0.58	0.52	0.61	0.49	0.87	0.38	0.43	0.82	0.81	0.13	0.52	0.55	0.70	Wet up
25	Win	Aut	7.75	0.44	0.31	0.60	0.59	0.42	0.64	-0.17	0.09	-0.40	0.65	0.40	0.36	0.56	0.16	-0.01	0.60	Rainfall+Temp
26	Sum	Spr	3.5	0.68	0.80	0.69	0.22	0.58	0.61	-0.38	0.72	0.36	0.66	0.61	0.36	0.21	0.76	0.46	0.35	Rainfall
27	Aut	Sum	3.2	0.41	0.13	0.36	0.28	-0.11	0.45	-0.50	-0.02	-0.18	0.07	0.02	0.22	-0.46	0.11	0.31	0.00	Wet up

	TOC concentrations				Summer			Autumn			Winter			Spring			all			Category
	Max	Min	Med	Rat	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	
28	Sum	Win	3.85	1.07	0.68	0.36	0.40	0.34	-0.38	0.54	0.58	0.10	0.66	0.02	-0.44	0.65	0.17	-0.60	0.79	Rainfall+Temp
29	Sum	Win	8.2	0.48	0.37	0.56	0.20	0.42	-0.01	0.62	-0.08	0.23	0.13	0.52	0.22	0.52	0.18	-0.24	0.74	Wet up
30	Win	Sum	0.9	3.03	0.50	-0.02	0.21	0.61	0.53	0.12	0.45	-0.11	0.26	0.80	0.02	0.54	0.66	0.14	0.23	Rainfall
31	Aut	Spr	7.4	0.19	0.17	0.38	0.45	-0.18	0.15	0.09	-0.03	-0.69	0.51	-0.11	-0.32	-0.11	0.12	-0.24	0.45	None
32	Sum	Spr	4.5	0.37	0.16	0.47	0.52	0.15	0.58	-0.18	0.00	-0.35	0.48	0.12	-0.29	-0.10	0.09	0.06	0.40	Wet up
33	Aut	Spr	7.7	0.23	0.29	0.38	-0.01	0.28	0.74	-0.21	0.01	0.19	-0.03	0.08	-0.28	-0.15	0.19	0.30	-0.07	Wet up
34	Sum	Win	3.7	0.76	0.70	0.55	0.48	0.41	0.28	0.70	0.13	-0.43	0.59	0.73	-0.06	0.58	0.28	-0.36	0.77	Rainfall+Temp
35	Sum	Win	3.3	0.64	0.60	0.19	0.27	0.41	0.18	0.05	0.52	-0.04	0.75	0.69	0.19	0.63	0.36	-0.23	0.63	Rainfall+Temp
36	Win	Sum	5.4	0.4	0.08	0.14	-0.08	0.19	0.50	-0.07	0.05	0.29	0.21	0.08	0.25	-0.31	0.17	0.51	-0.38	Wet up
37	Win	Sum	3.5	0.79	0.18	0.09	-0.45	-0.04	0.51	-0.34	0.05	0.45	0.20	0.05	0.42	0.09	0.09	0.45	-0.39	Wet up
38	Sum	Spr	4.25	0.51	0.01	0.20	0.75	-0.01	0.50	0.10	0.31	0.07	0.27	0.46	0.44	-0.20	0.15	0.19	0.39	Temp
39	Sum	Spr	15.9	0.23	-0.01	0.60	0.68	0.00	0.70	0.01	0.05	0.08	0.45	0.05	0.13	0.04	0.04	0.49	0.26	Wetup
40	Aut	Sum	3.6	0.37	-0.12	-0.46	-0.02	0.03	0.29	-0.50	-0.05	-0.84	-0.07	0.09	-0.53	-0.47	0.30	0.34	-0.16	None
41	Aut	Sum	9.6	0.18	0.11	-0.38	-0.09	-0.50	0.70	-0.68	0.34	-0.53	0.65	0.11	0.38	-0.31	0.31	0.48	-0.17	Wet up
42	Sum	Win	5.25	0.87	0.68	0.50	0.35	0.07	-0.10	0.34	0.14	-0.44	0.73	0.33	-0.18	0.17	0.21	-0.22	0.69	Rainfall
43	Win	Spr	6.8	0.3	-0.25	0.20	-0.42	-0.61	0.20	-0.67	-0.43	-0.24	-0.22	-0.26	-0.37	-0.48	-0.21	0.45	-0.46	None
44	Aut	Win	6.5	0.3	0.33	0.48	0.52	-0.09	0.02	0.01	0.31	-0.73	0.60	0.00	-0.12	0.77	-0.08	-0.35	0.71	None
45	Aut	Spr	2.7	1.67	-0.11	0.08	0.11	0.56	0.34	0.11	-0.31	-0.39	0.18	-0.67	0.26	0.32	0.01	-0.05	0.37	None
46	Aut	Spr	8.65	0.44	0.32	-0.42	0.49	-0.44	-0.08	0.24	0.16	-0.50	0.59	-0.06	-0.53	0.85	-0.11	-0.55	0.82	None
47	Aut	Spr	2.85	0.64	0.21	0.52	0.37	0.21	-0.32	0.05	0.38	-0.08	0.37	0.68	-0.08	0.33	0.25	0.00	0.32	None
48	Sum	Spr	4.25	0.56	0.86	0.87	-0.03	0.62	0.27	-0.53	0.76	0.36	0.26	0.77	-0.36	0.45	0.57	-0.11	0.49	Rainfall
49	Sum	Spr	2.1	1	0.48	-0.22	0.52	0.68	-0.29	0.22	0.94	0.38	0.37	0.27	0.03	0.35	0.63	-0.31	0.64	Rainfall
50	Aut	Spr	7.3	0.39	0.13	0.17	0.18	0.03	0.48	-0.60	-0.23	-0.46	0.41	0.27	-0.05	0.52	0.02	-0.03	0.33	Wet up
51	Sum	Sum	3.2	0.44	0.03	0.32	0.09	0.17	0.59	-0.58	0.01	0.16	0.11	0.13	-0.20	-0.02	0.13	0.40	-0.11	Wet up
52	Aut	Spr	6.05	0.15	0.36	0.79	0.11	0.09	0.72	-0.26	0.30	-0.43	0.26	-0.25	-0.76	-0.30	0.19	0.25	0.33	Wet up
53	Sum	Win	3.95	0.52	-0.16	0.13	0.69	-0.16	0.02	0.58	-0.60	-0.46	0.12	0.11	-0.68	0.16	-0.10	-0.25	0.69	Temp
54	Spr	Sum	8.55	0.58	0.26	0.24	0.34	0.14	0.31	-0.10	0.05	0.35	-0.29	0.08	0.47	-0.51	0.19	0.60	-0.41	Wet up
55	Aut	Spr	4.6	0.39	0.21	0.24	0.48	0.20	0.65	-0.25	0.51	-0.42	0.07	-0.48	-0.79	-0.12	0.08	0.13	0.22	Wet up
56	Aut	Sum	7	0.22	0.51	0.15	0.31	0.05	0.55	-0.11	-0.39	-0.04	-0.12	-0.49	-0.39	-0.06	0.00	0.09	0.08	None
57	Sum	Spr	5	0.09	0.86	0.32	-0.02	0.31	-0.44	-0.19	-0.30	-0.17	-0.33	-0.26	0.31	-0.32	0.27	0.09	0.09	Rainfall
58	Aut	Spr	6.5	0.28	0.25	0.51	0.31	-0.29	0.27	0.21	0.07	0.00	0.55	0.00	-0.68	0.24	0.08	-0.11	0.66	None
59	Spr	Sum	4.2	0.11	0.14	0.79	-0.12	0.22	0.40	-0.03	-0.01	-0.72	-0.01	-0.03	-0.19	0.18	0.15	0.17	0.11	Wet up
60	Sum	Win	4.4	0.34	-0.31	0.77	0.60	-0.57	-0.43	0.64	0.50	-0.50	1.00	-0.63	-0.11	0.95	-0.32	-0.42	0.72	Temp

	TOC concentrations				Summer			Autumn			Winter			Spring			all			Category
	Max	Min	Med	Rat	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	
61	Sum	Aut	5.4	0.7	0.40	-0.26	0.75	-0.44	-0.29	0.76	-0.04	-0.58	0.23	0.28	-0.07	0.76	0.00	-0.46	0.85	Temp
62	Win	Sum	8.4	0.17	-0.39	0.26	0.26	-0.31	0.41	-0.03	-0.27	-0.46	-0.35	-0.59	0.27	-0.74	-0.03	0.58	-0.59	None
63	Sum	Win	2.2	0.56	0.15	-0.13	-0.02	0.25	0.45	0.33	-0.15	-0.16	0.05	-0.12	-0.06	0.20	0.00	-0.17	0.37	None
64	Sum	Spr	3.6	0.57	0.10	-0.08	0.71	0.08	-0.25	-0.01	0.41	0.26	0.50	0.32	-0.23	0.75	0.13	-0.43	0.81	Temp
65	Sum	Win	4.7	0.34	0.24	0.56	0.76	0.17	0.42	0.17	-0.29	-0.60	0.24	-0.49	-0.51	-0.10	-0.01	-0.21	0.64	Temp
66	Win	Sum	3.8	0.47	0.14	-0.10	-0.02	0.33	0.70	0.02	0.16	-0.02	-0.19	-0.44	0.24	-0.69	0.04	0.35	-0.28	Wet up
67	Sum	Spr	4.05	0.74	-0.55	0.13	0.51	-0.67	-0.53	0.60	-0.22	-0.88	0.19	-0.12	-0.23	0.79	-0.27	-0.43	0.84	Temp
68	Win	Sum	7.1	0.43	-0.06	0.32	-0.29	0.18	0.54	-0.24	0.03	0.05	-0.33	-0.15	-0.06	-0.17	0.05	0.43	-0.43	Wet up
69	Aut	Spr	10	0.25	0.09	0.58	0.24	-0.07	0.45	-0.29	0.03	-0.60	0.78	0.15	0.25	0.34	0.04	0.02	0.33	Wet up
70	Aut	Sum	3.5	0.5	0.12	0.40	0.48	-0.22	0.08	-0.38	-0.14	-0.51	0.50	-0.06	-0.47	-0.09	-0.01	-0.09	0.42	Wet up
71	Sum	Win	3.9	0.63	0.52	-0.01	0.93	-0.27	-0.20	0.84	0.45	-0.55	0.47	0.08	-0.47	0.62	0.09	-0.46	0.90	Temp
72	Sum	Spr	10.35	0.25	0.73	-0.20	0.78	0.12	-0.32	0.23	0.73	0.72	0.43	0.84	0.32	0.56	0.67	0.10	0.51	Rainfall+Temp
73	Aut	Sum	8.1	0.15	0.52	0.81	0.38	-0.05	0.68	-0.77	0.03	-0.35	0.53	0.03	0.10	0.47	0.11	0.25	0.34	Wet up
74	Aut	Spr	12.55	0.32	0.07	0.31	0.71	-0.01	0.09	0.01	0.76	0.18	0.50	0.03	-0.37	0.87	0.18	-0.16	0.79	Temp
75	Sum	Spr	4.35	0.47	0.26	0.24	0.79	-0.38	-0.42	0.47	0.42	-0.38	0.65	-0.38	-0.72	0.41	-0.06	-0.40	0.82	Temp
76	Sum	Win	6.1	0.42	-0.09	0.73	0.77	-0.24	-0.25	0.82	-0.05	-0.07	0.46	0.20	-0.44	0.10	0.07	-0.19	0.72	Wet up
77	Sum	Aut	4.5	0.47	0.24	0.27	0.18	0.31	0.56	-0.14	0.19	-0.31	-0.21	-0.03	-0.17	-0.17	0.21	0.13	0.08	Wet up
78	Win	Sum	5.9	0.45	0.39	0.50	-0.06	0.32	0.27	-0.45	-0.19	-0.30	0.16	0.20	0.17	-0.61	0.14	0.29	-0.35	None
79	Aut	Win	8.95	0.45	0.55	0.51	0.73	0.16	-0.42	0.89	0.15	-0.40	0.71	-0.34	-0.35	0.74	0.05	-0.26	0.88	Temp
80	Win	Spr	6.9	0.38	0.12	0.44	0.33	0.08	0.37	-0.34	-0.33	-0.74	0.34	-0.03	-0.34	-0.41	0.09	0.03	0.24	Wet up
81	Sum	Win	6.85	0.42	0.13	-0.10	0.91	-0.01	-0.61	0.51	0.34	-0.53	0.67	0.32	-0.29	0.58	0.02	-0.47	0.87	Temp
82	Aut	Aut	9.95	0.23	0.61	0.42	0.39	0.31	0.42	0.04	0.44	-0.14	0.31	0.25	0.27	0.07	0.44	0.44	0.06	None
83	Sum	Spr	5.5	0.62	0.23	0.48	0.39	0.16	0.30	-0.11	-0.18	-0.51	0.48	-0.14	-0.35	0.46	0.04	0.04	0.63	Wet up
84	Aut	Spr	3.3	0.33	-0.31	0.59	-0.08	0.00	0.40	-0.24	0.01	-0.40	0.26	0.14	-0.24	0.04	0.14	0.21	0.01	Wet up
85	Win	Spr	5.9	0.14	-0.17	0.28	-0.05	0.37	0.55	-0.52	0.31	-0.45	0.38	-0.47	-0.25	-0.44	0.01	0.14	0.11	None
86	Aut	Spr	9.1	0.16	0.08	0.82	0.08	0.11	0.47	-0.06	-0.31	0.14	-0.16	0.27	-0.08	-0.27	0.03	0.43	0.15	Wet up
87	Aut	Spr	9.5	0.31	0.18	0.75	0.49	0.07	0.48	0.24	0.40	-0.05	0.45	-0.16	0.33	0.31	0.29	0.20	0.49	Wet up
88	Aut	Spr	3.55	0.43	0.31	0.39	0.75	0.11	-0.06	0.57	-0.42	-0.69	0.45	-0.38	-0.34	0.16	-0.06	-0.14	0.76	Temp
89	Sum	Spr	4.2	0.45	0.01	0.40	0.32	-0.13	0.33	-0.08	-0.17	-0.49	0.32	-0.11	-0.47	-0.25	-0.10	-0.05	0.31	Wet up
90	Aut	Sum	3.05	0.35	0.29	0.38	0.48	0.09	0.63	-0.22	-0.52	-0.52	0.48	0.48	-0.22	-0.36	0.26	0.16	0.35	Wet up
91	Sum	Win	5.45	0.6	-0.16	0.74	0.66	-0.02	-0.29	0.88	-0.11	-0.20	0.54	-0.19	-0.47	0.15	-0.22	-0.33	0.78	Wet up
92	Sum	Spr	7	0.41	0.16	0.64	0.72	0.13	0.64	-0.01	0.35	-0.21	0.69	0.05	-0.24	0.20	0.21	0.15	0.59	Wet up
93	Win	Sum	4.2	0.78	0.02	0.27	0.47	-0.05	0.19	-0.11	-0.03	-0.45	0.36	0.04	-0.29	-0.45	0.00	-0.02	0.29	None

	TOC concentrations				Summer			Autumn			Winter			Spring			all			Category
	Max	Min	Med	Rat	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	
94	Win	Spr	3.4	0.15	-0.21	0.13	0.05	-0.10	0.36	-0.04	-0.44	-0.45	0.52	-0.11	-0.20	0.16	-0.17	0.06	0.13	None
95	Aut	Sum	8.9	0.36	-0.21	0.67	0.49	0.74	0.30	-0.56	0.31	0.48	0.68	0.18	-0.04	0.26	0.46	0.55	0.14	Rainfall
96	Aut	Win	10.8	0.43	0.22	0.82	0.17	0.34	0.24	-0.30	0.25	-0.28	0.36	0.20	0.09	0.01	0.28	0.44	0.31	Wet up
97	Aut	Win	5.95	0.26	0.40	-0.27	0.49	0.22	0.49	0.14	0.55	-0.49	0.00	0.39	0.32	-0.14	0.29	0.07	0.17	None
98	Sum	Spr	6.5	0.2	0.35	0.60	0.70	0.09	-0.05	0.20	-0.28	-0.37	0.19	-0.31	-0.76	-0.31	-0.12	0.17	0.43	Wet up
99	Win	Spr	3.8	0.09	0.66	0.31	0.28	-0.16	-0.26	-0.45	-0.38	0.05	0.09	-0.22	-0.29	-0.15	0.03	0.15	-0.07	None
100	Aut	Sum	3.6	0.74	0.19	0.44	0.17	0.55	0.56	-0.59	-0.06	-0.35	0.38	0.24	-0.26	-0.12	0.17	0.14	0.19	Wet up
101	Aut	Sum	5.3	0.16	-0.07	0.81	-0.07	0.10	0.23	-0.07	-0.41	-0.05	0.38	-0.21	-0.40	-0.43	-0.05	0.26	-0.03	Wet up
102	Sum	Spr	5.75	0.85	0.68	0.30	0.46	0.66	0.08	0.75	0.42	-0.29	0.45	0.59	0.34	0.35	0.50	-0.21	0.77	Rainfall+Temp
103	Sum	Spr	6.4	0.14	-0.35	0.32	0.18	-0.49	0.77	-0.02	-0.11	-0.11	0.42	-0.22	0.00	-0.17	-0.14	0.09	0.54	Wet up
104	Sum	Win	2.6	0.92	0.45	0.09	0.42	0.61	-0.18	0.28	0.28	0.52	0.46	0.14	0.32	0.01	0.23	-0.30	0.63	Rainfall
105	Sum	Win	4.45	0.88	0.70	0.35	0.79	0.59	0.51	0.03	0.39	-0.13	0.56	0.29	0.02	0.71	0.29	-0.20	0.60	Rainfall+Temp
106	Aut	Spr	3.15	0.62	-0.18	0.67	0.37	0.56	0.86	-0.31	0.81	0.27	-0.05	0.70	0.54	-0.48	0.41	0.58	-0.03	Wet up
107	Aut	Spr	8.9	0.12	-0.09	0.66	0.28	0.23	0.52	-0.69	0.34	-0.37	0.30	-0.13	-0.18	-0.44	0.15	0.13	0.14	Wet up
108	Aut	Spr	18.8	0.2	0.01	0.63	0.50	0.03	-0.30	-0.22	-0.10	0.00	0.68	0.22	-0.23	0.32	0.08	-0.11	0.60	Wet up
109	Sum	Win	3.2	0.87	0.74	-0.02	0.36	0.38	0.28	0.13	0.05	-0.32	0.25	0.72	0.22	0.44	0.46	-0.06	0.53	Rainfall
110	Sum	Spr	2.1	0.83	0.15	0.48	0.41	-0.24	0.85	0.07	-0.18	-0.53	-0.13	0.21	-0.26	0.74	0.00	-0.10	0.65	Wet up
111	Aut	Spr	2.7	1.55	0.35	0.46	0.71	-0.04	0.35	0.59	0.37	-0.39	0.25	0.27	0.29	-0.30	0.16	0.15	0.32	Temp
112	Sum	Spr	3.15	0.71	0.40	0.72	0.35	0.67	0.36	-0.37	0.55	0.69	-0.50	0.85	0.24	-0.14	0.66	0.38	-0.12	Wet up
113	Sum	Spr	2.3	1.35	0.79	-0.18	0.44	0.54	0.06	0.28	0.72	0.41	0.45	0.61	0.00	0.67	0.49	-0.27	0.62	Rainfall+Temp
114	Sum	Spr	2.4	0.77	0.18	-0.25	0.44	0.38	-0.11	-0.04	0.20	-0.29	0.16	-0.16	-0.23	-0.03	0.13	-0.27	0.43	Temp
115	Aut	Sum	3.7	0.46	-0.17	0.70	0.43	-0.15	0.14	-0.55	0.13	-0.67	0.29	0.60	0.01	-0.86	0.32	0.49	0.24	Wet up
116	Aut	Win	5.5	0.44	0.54	0.69	0.11	0.40	0.23	0.56	0.13	-0.88	0.33	0.62	-0.05	-0.09	0.26	-0.07	0.72	Wet up
117	Sum	Spr	6.05	0.81	0.62	0.31	0.45	0.60	0.15	0.68	0.62	-0.31	0.51	0.59	0.34	0.35	0.54	-0.18	0.75	Rainfall+Temp
118	Sum	Win	6.75	0.59	-0.03	0.02	0.69	-0.48	-0.71	0.63	0.25	-0.18	0.25	0.47	-0.07	0.84	-0.20	-0.58	0.84	Temp
119	Sum	Win	3.2	0.97	0.63	0.05	0.52	0.80	0.54	-0.18	-0.78	-0.30	-0.25	0.64	-0.26	0.78	0.20	-0.35	0.57	Rainfall
120	Sum	Spr	2.8	0.77	-0.04	0.08	0.19	0.14	0.34	-0.23	-0.25	-0.65	0.26	0.09	-0.26	-0.24	-0.03	-0.18	0.29	None
121	Win	Spr	7.05	0.52	0.44	0.72	0.25	0.70	0.48	-0.69	0.60	0.19	0.20	0.36	0.57	-0.16	0.51	0.59	-0.07	Rainfall
122	Sum	Win	3.6	0.61	0.61	-0.06	0.56	0.70	0.13	0.14	0.12	0.07	0.65	0.24	-0.34	0.51	0.23	-0.40	0.73	Rainfall+Temp
123	Sum	Win	4.45	0.72	0.80	-0.05	0.59	0.05	-0.73	0.51	0.22	0.20	0.17	0.64	-0.08	0.57	-0.11	-0.48	0.80	Rainfall+Temp
124	Sum	Win	9.7	0.57	0.34	0.35	0.38	0.12	-0.38	0.51	0.20	-0.15	0.45	0.51	-0.43	0.73	0.20	-0.58	0.76	Temp
125	Aut	Sum	8	0.37	-0.10	0.10	-0.24	0.43	0.42	-0.57	0.21	0.10	0.39	0.23	0.26	0.07	0.22	0.46	-0.20	Wet up
126	Aut	Spr	5.4	0.28	-0.03	0.36	-0.04	0.37	0.60	-0.62	-0.06	0.34	0.38	-0.08	0.52	0.02	0.23	0.63	-0.36	Wet up

	TOC concentrations				Summer			Autumn			Winter			Spring			all			Category
	Max	Min	Med	Rat	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	Short rainfall	Long rainfall	Tempera-ture	
127	Aut	Win	7.35	0.33	-0.22	0.38	0.62	0.67	-0.17	0.23	0.40	-0.30	0.44	0.05	0.26	0.85	0.05	-0.31	0.85	Rainfall+Temp

C.4 Model estimates for TOC median values for projected climate data

Table C.6: Outputs from the chosen multiple linear regression model for TOC median projections. Model estimate = estimate value derived when using projected climate change data (for AAT and SER); lower confidence and upper confidence interval = 5th and 9th percentiles; projected TOC median value = median TOC the model estimates (inverse log-transformed); projected change = difference between the projected median and the current (2013 – 2016) median.

ID	Model estimate	lower confidence interval	upper confidence interval	Projected TOC median value	Projected change in TOC median
2	2.408791	2.121973	2.695609	11.12051	2.470507
6	2.686025	2.230565	3.141484	14.67323	8.773232
7	2.036096	1.686219	2.385973	7.660645	3.960645
9	2.417386	1.988614	2.846158	11.2165	4.716498
15	1.720312	1.343481	2.097143	5.58627	1.73627
16	1.504174	1.107466	1.900881	4.500434	-3.69957
17	3.055744	2.432945	3.678543	21.23698	14.23698
20	1.994402	1.632942	2.355863	7.347811	5.047811
21	1.905598	1.635015	2.17618	6.723424	3.523424
22	3.286708	2.467634	4.105781	26.75464	19.05464
23	3.31982	2.436432	4.203207	27.65537	22.65537
24	3.23054	2.40033	4.060751	25.29332	18.89332
25	1.719525	1.338736	2.100314	5.581875	-1.16813
32	1.887549	1.541125	2.233974	6.603166	3.103166
33	1.370374	1.001625	1.739123	3.936822	0.236822
34	2.886481	2.4011	3.371862	17.9301	5.380104
42	2.061883	1.72303	2.400736	7.860756	3.660756
43	1.850856	1.538447	2.163265	6.365267	2.165267
45	2.078871	1.812226	2.345516	7.995436	0.595436
59	2.926347	2.390816	3.461878	18.65935	9.759349
75	1.077273	0.724324	1.430223	2.936661	2.036661
80	2.78807	2.349006	3.227134	16.24963	8.949634
89	2.409434	2.130175	2.688692	11.12766	4.62766
103	2.77434	2.356195	3.192486	16.02805	8.62805
116	2.039823	1.757373	2.322273	7.689245	4.389245
126	2.427549	2.11355	2.741548	11.33107	4.831073
127	2.4256	2.094059	2.757141	11.30901	1.309011
138	2.258692	1.658387	2.858997	9.570563	4.170563
146	2.382277	1.825009	2.939546	10.82954	1.029539
167	2.159931	1.833859	2.486004	8.670542	5.070542
192	2.641912	2.307338	2.976485	14.04002	10.84002
193	2.420667	2.03486	2.806475	11.25337	6.753365
201	3.238406	2.774613	3.702198	25.49305	6.693045
203	3.076988	2.553609	3.600368	21.69297	13.94297
208	2.23483	1.765855	2.703806	9.344898	6.644898
211	2.429529	1.937682	2.921377	11.35354	4.003537
212	1.687558	1.541074	1.834042	5.406262	1.906262
213	3.045093	2.512833	3.577353	21.01198	15.71198
214	2.973506	2.304969	3.642043	19.56038	9.210377
215	2.997342	2.50657	3.488115	20.03223	13.13223
222	3.346562	2.747061	3.946062	28.4049	19.6549
227	2.963003	2.497857	3.42815	19.35602	13.95602
232	2.711386	2.259218	3.163554	15.05012	7.050118

242	2.34154	1.510295	3.172785	10.39724	6.14724
266	1.311027	1.014721	1.607332	3.709981	0.109981
268	1.566868	1.260208	1.873529	4.791618	0.941618
270	1.292904	0.827197	1.758611	3.643351	-1.75665
271	2.640219	2.326294	2.954144	14.01627	1.76627
273	1.011649	0.694175	1.329122	2.750131	-0.14987
276	1.015869	0.585798	1.445939	2.761762	1.261762
277	1.263364	0.855894	1.670834	3.537301	-0.3627
278	2.042902	1.658058	2.427747	7.712962	4.612962
279	1.436008	1.020982	1.851034	4.20388	1.00388
280	2.489043	2.193335	2.78475	12.04973	8.899733
296	2.377466	1.940957	2.813975	10.77756	7.277557
297	2.77365	2.281711	3.26559	16.01699	8.71699
299	1.920128	1.315546	2.52471	6.82183	3.67183
305	2.771435	2.220072	3.322797	15.98154	8.931545
310	2.050067	1.85347	2.246663	7.76842	1.86842
311	2.81649	2.377393	3.255587	16.71806	7.61806
312	3.083578	2.584217	3.582939	21.83639	5.936394
319	2.286274	1.91494	2.657608	9.838208	6.038208
320	2.607453	2.229311	2.985596	13.56446	9.764463
321	2.370217	1.649533	3.090901	10.69971	7.299711
327	2.377912	1.871838	2.883985	10.78236	6.282362
330	3.058554	2.505389	3.611719	21.29674	12.89674
333	2.392587	2.080473	2.704701	10.94176	6.341761
334	1.614292	1.35149	1.877093	5.024329	-1.47567
335	0.925138	0.655216	1.195059	2.522215	0.122215
337	1.669841	1.435505	1.904177	5.311323	1.711323
338	0.890363	0.699452	1.081274	2.436014	1.336014
339	1.492317	1.227145	1.757488	4.447387	-1.35261
340	2.12435	1.82246	2.426241	8.367459	2.317459
341	1.95954	1.540725	2.378356	7.096064	4.396064
342	1.611898	1.226566	1.99723	5.012318	1.112318
343	1.350902	1.104834	1.59697	3.860906	0.260906
344	2.409464	2.040162	2.778765	11.12799	6.92799
346	2.350026	2.030224	2.669828	10.48584	3.985838
349	2.885168	2.369673	3.400663	17.90657	9.006567
350	3.067798	2.493784	3.641811	21.49452	10.69452
356	2.490077	1.960419	3.019735	12.0622	5.962203
357	2.888821	2.411973	3.365668	17.9721	11.1721
365	1.937646	1.705409	2.169883	6.942387	1.442387
368	2.070179	1.575148	2.56521	7.926242	4.326242
373	2.782664	2.367373	3.197956	16.16202	6.562025
376	2.221763	2.021604	2.421922	9.223581	5.523581
384	1.867687	1.598538	2.136836	6.473306	-3.22669
388	1.223817	0.960295	1.487338	3.40014	-0.94986
393	1.305417	1.090506	1.520327	3.689225	-1.56077
398	1.975791	1.672414	2.279168	7.212322	2.412322
401	1.700627	1.414062	1.987192	5.477381	1.027381
402	1.298501	1.064823	1.532178	3.663799	0.613799
407	1.77894	1.477292	2.080587	5.923572	0.173572
410	2.535153	2.062123	3.008184	12.61837	9.318368
411	1.989182	1.581597	2.396767	7.309552	5.209552
412	1.659249	1.436531	1.881966	5.255362	2.655362
415	2.373953	2.140038	2.607869	10.73977	5.239768

416	0.703496	0.444874	0.962117	2.020804	-1.5792
418	2.234456	1.559072	2.90984	9.3414	6.4914
422	2.549207	1.885644	3.21277	12.79696	8.546956
424	1.981892	1.352745	2.611038	7.256456	4.856456
425	1.51892	1.126948	1.910892	4.567291	2.867291
426	0.831121	0.563482	1.09876	2.29589	0.19589
427	1.78478	1.260185	2.309374	5.958268	1.508268
429	2.640083	2.250343	3.029822	14.01436	4.514362
432	1.868055	1.469029	2.267081	6.475689	1.025689
433	1.64387	1.336442	1.951298	5.175158	0.825158
434	2.051122	1.626867	2.475377	7.776623	1.676623
436	2.909544	2.384815	3.434272	18.34842	11.34842
437	2.338275	1.975023	2.701527	10.36334	5.663344
440	2.347251	2.019476	2.675025	10.45678	4.50678
442	2.267259	1.890472	2.644047	9.65291	0.70291
445	2.65458	2.327807	2.981353	14.21901	7.369011
446	1.755287	1.510737	1.999838	5.78511	1.38511
447	2.537995	2.179601	2.89639	12.65428	4.554277
448	2.853396	2.389938	3.316854	17.34659	13.29659
449	1.791747	1.448973	2.134522	5.999927	3.799927
450	1.713241	1.416756	2.009726	5.546911	1.996911
451	1.834045	1.649952	2.018139	6.259156	1.809156
452	2.768509	2.321841	3.215177	15.93486	8.834858
456	1.814263	1.511314	2.117211	6.136549	0.086549
1311	2.383036	2.132348	2.633725	10.83776	8.03776
1361	2.577167	1.937398	3.216936	13.1598	9.659802
3001	1.685605	1.425163	1.946046	5.395712	1.445712
3251	3.140442	2.549912	3.730972	23.11409	13.16409
4141	2.217424	1.426924	3.007925	9.183644	5.983644
4391	2.368658	2.055694	2.681622	10.68305	2.133046

C.5 Model estimates for *E. coli* median values for projected land use data

Table C.7: Outputs from the chosen multiple linear regression model for *E. coli* median projections. Model estimate = estimate value derived when using projected land capability data (for LCA and LCL); lower confidence and upper confidence interval = 5th and 9th percentiles; projected *E. coli* median value = median *E. coli* the model estimates (inverse log-transformed); projected change = difference between the projected median and the current (2011 – 2016) median.

Catchment ID	**Model estimate**	**lower confidence interval**	**upper confidence interval**	**Projected *E. coli* median value**	**Projected change in *E. coli* median**
1	1.531978	3.627321	1.268304	1.795652	2.627321
2	1.494472	3.456981	1.237458	1.751485	0.456981
6	1.492642	3.448834	1.127837	1.857447	2.448834
7	0.94072	1.561825	0.464197	1.417242	0.561825
8	1.269433	2.558834	0.981419	1.557447	0.558834
9	1.566884	3.791695	1.321077	1.812691	0.791695
15	1.692481	4.432943	1.320445	2.064517	3.432943
16	2.163395	7.700628	1.855356	2.471434	4.700628
17	1.502806	3.494284	1.244598	1.761015	2.494284
18	1.896683	5.663752	1.60551	2.187855	2.663752
20	2.150893	7.592528	1.846942	2.454844	5.592528
21	2.528769	11.53807	2.152132	2.905407	10.53807
22	1.54448	3.685535	1.277874	1.811087	0.685535
23	7.357858	1567.473	4.885822	9.829893	1566.473
24	1.453013	3.275977	1.08752	1.818506	2.275977
25	1.09857	1.999874	0.725811	1.47133	-1.00013
31	1.183484	2.265731	0.839292	1.527675	1.265731
32	1.542855	3.677926	1.277995	1.807715	2.677926
33	1.855009	5.391755	1.552678	2.157339	2.391755
34	1.548648	3.705103	1.280988	1.816307	0.705103
39	1.163147	2.199989	0.816563	1.509732	1.199989
42	1.440296	3.221944	1.186886	1.693705	2.221944
43	1.450765	3.266376	1.091505	1.810025	2.266376
45	1.419459	3.134882	1.165478	1.67344	0.134882
59	1.639663	4.153433	1.390948	1.888379	1.153433
63	2.75159	14.66752	2.274421	3.228759	11.66752
75	1.442438	3.230999	0.916546	1.96833	1.230999
76	2.106579	7.22007	1.834352	2.378805	5.22007
80	2.501317	11.19855	1.653553	3.349082	8.198554
89	2.400651	10.03036	1.642168	3.159135	7.03036
103	4.29219	72.12641	3.155318	5.429061	69.12641
116	3.614995	36.15115	2.795106	4.434884	35.15115
126	1.997288	6.369046	1.547611	2.446965	3.369046
127	2.201476	8.038346	1.605657	2.797295	4.038346
138	5.304862	200.3132	3.774873	6.83485	199.3132
146	2.736272	14.42935	2.286866	3.185677	11.42935
167	2.094856	7.124271	1.669492	2.52022	6.124271
192	2.269675	8.676258	1.647763	2.891588	7.676258
193	1.809463	5.10717	1.275188	2.343739	4.10717
201	1.632422	4.116252	1.368981	1.895863	0.116252
203	2.233883	8.336052	1.640947	2.82682	4.336052
208	2.444353	10.52309	1.746006	3.1427	9.523091
211	2.085907	7.05189	1.418233	2.75358	3.05189

212	1.99651	6.363317	1.718675	2.274346	4.363317
213	2.15927	7.664806	1.575481	2.743058	6.664806
214	2.484545	10.99566	1.638324	3.330765	6.995657
215	2.556553	11.89131	1.812215	3.300892	10.89131
222	1.959817	6.098028	1.625112	2.294522	2.098028
227	4.51544	90.41774	3.505257	5.525622	89.41774
232	3.163302	22.64856	2.602398	3.724206	18.64856
242	8.578804	5316.74	5.527713	11.62989	5312.74
266	2.146726	7.556795	1.844077	2.449375	6.556795
268	2.196734	7.995588	1.876523	2.516945	6.995588
270	1.004787	1.731325	0.635833	1.37374	-0.26867
271	2.659336	13.2868	2.253598	3.065073	10.2868
272	2.765738	14.89076	2.257711	3.273764	12.89076
273	1.984198	6.27321	1.704753	2.263643	5.27321
276	2.000867	6.395467	1.72172	2.280015	4.395467
277	0.858929	1.36063	0.435625	1.282232	-0.63937
278	2.280082	8.777481	1.922507	2.637657	7.777481
279	2.105052	7.207528	1.813643	2.396461	5.207528
280	5.065788	157.5053	4.198331	5.933245	156.5053
294	7.917869	2744.914	5.756055	10.07968	2739.914
295	5.418753	224.5975	4.367383	6.470122	219.5975
296	8.752623	6326.26	6.204897	11.30035	6321.26
297	3.188542	23.25304	2.318297	4.058787	19.25304
299	2.750967	14.65777	1.782037	3.719897	12.65777
302	4.959238	141.4852	4.073254	5.845223	140.4852
303	7.843767	2548.793	5.762983	9.924552	2543.793
305	9.323303	11194.9	6.523656	12.12295	11189.9
310	1.525443	3.59718	1.233131	1.817755	2.59718
311	1.69383	4.440279	1.304218	2.083442	1.440279
312	2.345013	9.433407	1.864229	2.825797	5.433407
316	2.256021	8.545034	1.442531	3.069511	5.545034
319	1.305732	2.690391	0.924211	1.687253	1.690391
320	1.823938	5.196211	1.272763	2.375113	4.196211
321	2.071537	6.937014	1.670977	2.472097	4.937014
327	1.727235	4.625079	1.458992	1.995478	2.625079
330	1.450849	3.266737	1.091915	1.809784	1.266737
332	1.934189	5.918432	1.650083	2.218296	4.918432
333	1.263165	2.536595	0.907491	1.618838	1.536595
334	1.671644	4.320909	1.288401	2.054888	2.320909
335	1.788331	4.979463	1.461819	2.114843	3.979463
336	1.556982	3.744482	1.287105	1.826859	1.744482
337	1.204821	2.336162	0.856642	1.553	0.336162
338	1.896683	5.663752	1.60551	2.187855	4.663752
339	1.871678	5.499196	1.574208	2.169149	2.499196
340	1.344446	2.83606	1.079609	1.609282	1.83606
341	2.426245	10.31631	1.986428	2.866063	8.316313
342	2.213404	8.146797	1.886462	2.540345	6.146797
343	2.096717	7.139405	1.807149	2.386285	5.139405
344	1.523643	3.588914	1.261731	1.785556	2.588914
346	1.531978	3.627321	1.268304	1.795652	1.627321
349	3.079409	20.74554	2.416221	3.742596	16.74554
350	2.337711	9.357501	2.007004	2.668418	5.357501
352	1.031706	1.805848	0.584962	1.478449	-2.19415
356	3.52409	32.9229	2.75511	4.293071	28.9229

357	1.695882	4.451452	1.439842	1.951922	3.451452
365	1.907581	5.736775	1.372938	2.442225	4.736775
367	1.125717	2.082427	0.373314	1.87812	1.082427
368	1.342345	2.828009	0.967363	1.717326	0.828009
369	1.6977	4.461371	1.43705	1.95835	1.461371
373	1.363182	2.90861	0.981536	1.744828	-0.09139
376	1.208989	2.350095	0.860476	1.557501	0.350095
384	2.284249	8.818312	1.924585	2.643914	5.818312
387	2.013369	6.488507	1.734029	2.29271	3.488507
388	1.830005	5.233916	1.519455	2.140554	4.233916
393	2.050876	6.774708	1.768826	2.332926	5.774708
394	1.042293	1.835713	0.682511	1.402075	-0.16429
398	2.409271	10.12584	1.979579	2.838963	9.125844
400	1.033959	1.812176	0.672333	1.395584	0.812176
401	2.000867	6.395467	1.72172	2.280015	5.395467
402	1.075632	1.931846	0.722066	1.429199	0.931846
405	1.204821	2.336162	0.856642	1.553	-0.66384
407	2.088382	7.071847	1.800513	2.376252	6.071847
410	4.019463	54.67119	3.096351	4.942574	53.67119
411	2.247633	8.465301	1.710106	2.785159	7.465301
412	1.796666	5.029509	1.473557	2.119774	4.029509
415	2.11231	7.26732	1.803297	2.421324	5.26732
416	1.529953	3.617961	1.061837	1.998069	2.617961
418	4.550504	93.68008	3.753946	5.347062	92.68008
422	4.018727	54.63024	3.303773	4.733681	53.63024
424	5.437626	228.8958	4.388498	6.486753	226.8958
425	1.684869	4.391747	1.025581	2.344157	3.391747
426	2.067545	6.905395	1.783284	2.351807	5.905395
427	1.91752	5.804061	1.630637	2.204402	3.804061
428	1.431961	3.186901	1.178454	1.685468	0.186901
429	1.549366	3.708482	1.290278	1.808453	0.708482
432	0.763079	1.14487	0.291626	1.234531	-0.85513
433	1.06313	1.89542	0.707454	1.418807	-0.10458
434	1.167315	2.213352	0.820714	1.513915	0.213352
436	1.279834	2.596043	0.920935	1.638733	0.596043
437	1.258997	2.521888	0.904057	1.613937	0.521888
440	1.652832	4.221748	1.348016	1.957648	2.221748
442	1.043301	1.838571	0.588142	1.498459	-1.16143
443	1.490304	3.438446	1.233825	1.746783	-0.56155
444	0.90477	1.471363	0.501375	1.308164	-0.52864
445	1.636163	4.135425	1.338542	1.933784	2.135425
446	0.854761	1.350813	0.429537	1.279985	-0.64919
447	1.594896	3.927814	1.313428	1.876363	0.927814
448	1.346512	2.843995	0.970247	1.722777	0.843995
449	1.346512	2.843995	0.970247	1.722777	1.843995
450	1.048562	1.853544	0.645588	1.451536	-0.14646
451	2.150893	7.592528	1.846942	2.454844	5.592528
452	3.284284	25.68986	2.667135	3.901432	24.68986
456	2.109219	7.241804	1.816838	2.401601	6.241804
1311	2.368454	9.680869	2.021086	2.715822	8.680869
1361	3.005885	19.2041	2.404314	3.607456	18.2041
3001	1.519476	3.56983	1.258385	1.780567	2.56983
3251	1.491119	3.442064	1.236418	1.745821	0.442064
3281	1.486137	3.419987	1.23015	1.742124	1.419987

3851	1.213156	2.364085	0.864278	1.562034	1.364085
4141	6.512253	672.3419	4.491521	8.532986	671.3419
4391	1.665334	4.28744	1.354857	1.975811	0.28744

C.6 *E. coli* climate sensitivity relationships and group allocations

Table C.8: Results of Spearman's rank correlation tests run per catchment on sampled *E. coli* concentrations from 2011 – 2016 and corresponding daily total rainfall, and total rainfall of the preceding 3, 10, and 30 days for the summer season (March to October) and the winter season (November to April). Rho values are given with highlights indicating statistical significance ($p<0.05$). NA = no data.

	Summer				Winter				Category
	Rain_rho	Rain03_rho	Rain10_rho	Rain30_rho	Rain_rho	Rain03_rho	Rain10_rho	Rain30_rho	
1	0.08	0.10	0.21	0.02	0.61	0.29	0.31	-0.23	None
2	0.28	0.08	0.23	0.13	-0.14	-0.09	-0.18	-0.78	None
6	0.41	0.46	0.30	0.25	-0.26	0.43	0.31	0.03	None
7	0.46	0.41	0.29	0.13	0.37	-0.10	-0.12	-0.55	None
8	0.33	0.24	0.29	0.02	0.44	-0.28	0.20	-0.71	None
9	-0.18	0.43	0.33	0.00	-1.00	-0.50	0.50	0.50	None
15	0.26	0.70	0.33	0.10	0.62	0.46	-0.03	0.19	Medium
16	0.11	0.31	0.12	0.04	-0.20	-0.29	-0.05	0.16	None
17	0.09	0.30	0.71	0.09	-0.30	-0.30	-0.70	-0.70	Medium
18	0.00	-0.11	0.10	0.11	0.11	0.00	-0.04	-0.70	None
20	-0.18	-0.29	-0.31	-0.37	-0.32	0.21	0.63	0.95	None
21	0.43	0.52	0.30	0.01	0.29	0.50	0.20	0.20	None
22	0.06	0.13	0.04	-0.03	0.63	0.67	-0.36	-0.05	None
23	0.30	0.66	0.72	0.80	-0.59	0.18	0.35	-0.19	Medium
24	0.23	0.06	0.10	-0.14	0.88	0.84	-0.10	-0.54	None
25	0.16	0.28	0.00	-0.20	-0.27	0.35	-0.07	-0.23	None
30	-0.18	0.18	0.08	-0.19	-0.80	0.40	0.80	0.40	None
31	-0.08	0.52	0.81	0.66	0.40	0.14	0.45	0.35	Long
32	0.46	0.24	0.27	0.15	0.13	0.25	0.33	-0.21	None
33	0.13	0.24	-0.34	-0.09	0.47	0.62	0.58	0.38	None
34	-0.14	-0.01	0.34	0.30	0.09	0.20	0.66	-0.37	None
39	0.30	0.20	-0.07	-0.26	0.70	0.39	0.94	-0.39	None
42	0.16	0.18	0.24	-0.10	0.02	0.16	0.68	0.19	None

	Summer				Winter				Category
	Rain_rho	Rain03_rho	Rain10_rho	Rain30_rho	Rain_rho	Rain03_rho	Rain10_rho	Rain30_rho	
43	0.31	0.52	0.35	0.05	0.01	0.20	0.56	0.58	Medium
45	-0.03	0.19	0.22	0.01	-0.30	-0.07	-0.18	-0.08	None
59	0.18	0.73	0.65	0.44	0.43	0.41	0.41	0.19	Medium
63	0.16	0.61	0.71	0.31	0.13	0.15	0.24	0.05	Medium
75	-0.05	0.28	0.00	0.06	0.78	0.07	-0.51	-0.44	None
76	0.28	0.14	0.07	-0.17	0.67	-0.45	-0.11	0.11	None
80	0.14	0.28	0.32	0.12	0.27	0.40	0.55	0.30	None
89	0.14	0.49	0.50	0.30	0.70	0.90	0.60	0.00	Medium
103	-0.13	0.17	0.39	0.08	-0.06	-0.64	-0.26	-0.20	None
116	0.10	0.15	0.42	0.21	0.23	0.53	0.62	0.42	Medium
126	0.10	0.50	0.45	0.19	0.09	0.44	0.39	0.13	Medium
127	0.26	0.37	0.49	0.31	0.57	0.78	0.38	0.23	Medium
131	0.28	0.51	0.52	0.21	-0.26	0.56	0.34	0.40	Medium
136	0.44	0.29	0.33	0.22	0.35	0.34	0.42	0.51	None
138	0.33	0.34	0.51	0.50	0.02	0.11	0.16	0.30	Long
146	0.05	0.24	0.44	0.60	0.44	0.15	0.06	-0.09	Long
167	0.04	0.42	0.51	0.02	0.16	0.44	0.21	0.19	Medium
192	0.18	0.26	0.49	0.38	-0.45	-0.31	0.38	0.18	None
193	-0.11	0.54	0.57	0.52	0.51	0.43	0.55	0.58	Medium
201	0.14	0.36	0.42	0.38	0.01	0.17	0.27	0.04	None
203	0.10	0.49	0.62	0.34	0.02	0.57	0.39	0.22	Medium
208	-0.31	0.34	0.69	0.56	0.31	0.70	0.37	0.29	Long
211	0.34	0.57	0.30	0.14	0.67	0.81	0.26	-0.57	Medium
212	0.17	0.37	0.39	0.02	0.14	0.86	0.48	0.52	None
213	0.10	0.46	0.45	0.34	-0.32	-0.68	0.00	-0.02	None
214	0.04	0.30	0.65	0.60	0.46	0.75	0.64	0.58	Long
215	0.42	0.72	0.67	0.06	0.02	0.23	0.33	0.24	Medium
222	0.08	0.29	0.70	0.39	0.43	0.53	0.55	0.54	Medium
227	0.04	0.13	0.45	0.41	-0.17	0.42	0.45	0.22	Long

	Summer				Winter				Category
	Rain_rho	Rain03_rho	Rain10_rho	Rain30_rho	Rain_rho	Rain03_rho	Rain10_rho	Rain30_rho	
232	0.05	0.26	0.42	0.41	-0.14	0.19	0.37	0.38	Long
242	0.21	0.48	0.38	0.02	0.14	0.63	0.61	0.27	Medium
266	0.02	0.22	0.15	-0.26	-0.41	-0.56	-0.62	0.72	None
268	0.69	0.46	0.38	0.06	0.93	0.64	-0.18	-0.29	Short
270	0.31	0.22	0.19	-0.27	0.00	0.61	0.61	0.00	None
271	-0.24	-0.01	-0.14	-0.26	0.14	0.23	-0.20	0.00	None
272	0.09	0.42	0.34	-0.22	0.79	0.21	0.36	-0.04	None
273	-0.29	-0.11	-0.28	-0.10	0.81	0.31	0.05	0.04	None
276	-0.31	0.21	0.18	0.05	0.66	0.60	0.60	0.60	None
277	0.32	0.50	0.60	0.29	0.09	0.74	0.31	0.38	Medium
278	-0.01	0.09	-0.04	-0.34	-0.15	-0.40	-0.10	-1.00	None
279	0.13	-0.02	0.12	-0.19	0.59	0.78	0.63	0.30	None
280	-0.02	0.15	0.21	0.31	-0.37	-0.03	-0.71	0.03	None
294	0.28	0.54	0.49	0.38	0.20	0.28	0.11	0.23	Medium
295	0.22	0.41	0.16	0.13	0.24	0.21	0.07	0.18	None
296	0.29	0.69	0.62	0.69	0.43	0.82	0.90	0.40	Medium
297	-0.17	0.35	0.10	0.09	0.49	0.19	-0.44	-0.61	None
299	0.10	0.07	0.05	0.47	-0.32	-0.23	-0.28	-0.18	None
302	-0.16	0.04	-0.07	-0.08	-0.21	-0.06	0.27	-0.15	None
303	0.35	0.72	0.57	0.49	0.06	0.14	-0.13	0.07	Medium
305	0.00	0.53	0.45	0.40	-0.17	0.65	0.09	0.04	Medium
310	-0.03	0.26	0.23	0.06	NA	NA	NA	NA	None
311	0.08	-0.15	0.26	0.13	0.71	0.71	0.71	0.35	None
312	0.17	0.15	0.32	0.15	0.12	0.56	0.51	-0.59	None
316	0.39	-0.10	0.11	0.25	0.58	0.58	0.80	0.53	None
319	0.33	0.14	0.18	-0.10	-0.42	-0.42	0.30	0.24	None
320	-0.02	0.37	0.41	-0.15	NA	NA	NA	NA	None
321	0.38	0.58	0.20	0.31	0.74	0.35	0.41	0.29	None
325	0.21	0.36	0.30	0.28	-0.10	0.10	0.52	0.10	None

	Summer				Winter				Category
	Rain_rho	Rain03_rho	Rain10_rho	Rain30_rho	Rain_rho	Rain03_rho	Rain10_rho	Rain30_rho	
327	0.03	0.17	0.32	0.16	-0.22	-0.15	-0.11	-0.19	None
328	0.01	0.26	-0.02	-0.50	0.46	0.39	0.36	0.10	None
330	0.34	0.14	0.10	-0.43	NA	NA	NA	NA	None
332	0.06	0.61	0.34	0.53	0.36	0.25	-0.61	-0.52	Medium
333	-0.12	0.09	0.38	0.43	0.34	0.68	0.68	0.54	None
334	0.32	0.25	-0.17	-0.35	0.35	0.16	0.02	-0.63	None
335	0.42	0.50	0.48	0.08	-0.11	-0.63	-0.30	-0.22	None
336	0.30	0.09	0.31	0.05	-0.46	-0.12	0.23	0.70	None
337	0.03	0.55	0.43	0.11	-0.68	0.17	-0.17	-0.34	None
338	0.00	0.54	0.56	0.64	-0.39	-0.02	0.06	-0.76	Long
339	0.18	0.42	0.36	0.02	0.39	0.33	-0.39	-0.58	None
340	0.14	0.09	0.18	0.46	-0.65	-0.37	0.16	0.57	None
341	0.03	0.29	0.26	-0.02	-0.34	-0.25	0.31	0.93	None
342	0.12	0.13	-0.03	-0.29	-0.74	0.63	-0.21	0.32	None
343	0.69	0.44	-0.20	-0.47	0.30	0.62	0.26	-0.18	Short
344	0.20	0.02	0.56	0.45	0.41	0.00	0.21	0.00	None
346	-0.12	0.33	0.14	-0.10	0.76	0.41	0.23	-0.11	None
349	0.44	0.54	0.59	0.55	0.64	0.62	0.62	0.26	Medium
350	0.22	0.44	0.49	0.34	0.51	0.49	0.45	-0.36	Medium
352	0.17	0.44	0.33	0.01	-0.41	-0.30	-0.69	-0.43	None
356	0.40	0.48	0.49	0.19	0.34	0.43	0.49	0.10	None
357	0.13	0.31	0.01	0.19	0.45	-0.45	-0.45	-0.45	None
365	-0.15	0.41	0.23	-0.14	0.05	0.23	0.05	-0.14	None
367	0.12	0.04	0.36	0.35	-0.89	0.00	0.00	0.89	None
368	0.04	-0.29	0.02	0.47	0.36	0.63	0.70	0.38	None
369	0.32	0.21	0.41	0.30	0.13	-0.26	-0.19	-0.16	None
373	0.48	0.56	0.97	0.35	0.55	0.51	0.52	0.44	Medium
376	-0.19	0.02	0.72	0.48	-0.16	0.08	-0.20	-0.60	Medium
384	0.36	0.40	0.38	0.42	-0.60	-0.10	-0.60	-0.70	None

	Summer				Winter				Category
	Rain_rho	Rain03_rho	Rain10_rho	Rain30_rho	Rain_rho	Rain03_rho	Rain10_rho	Rain30_rho	
385	0.09	0.26	0.50	0.50	0.35	0.57	-0.04	-0.47	None
387	-0.02	0.34	-0.03	-0.26	-0.05	0.92	0.72	-0.67	None
388	0.47	0.50	0.26	-0.14	-0.04	0.46	0.79	-0.04	Medium
393	-0.24	-0.21	-0.06	-0.01	0.51	0.68	-0.17	-0.51	None
394	-0.16	0.13	-0.06	0.16	-0.21	0.51	0.17	0.21	None
398	-0.04	-0.08	-0.26	-0.37	0.25	-0.03	-0.14	-0.22	None
400	0.10	0.61	0.30	-0.04	0.22	0.67	0.76	0.04	Medium
401	0.07	0.15	0.06	-0.18	0.14	0.49	0.14	0.37	None
402	0.20	0.21	0.34	0.65	0.32	-0.22	0.71	0.63	Long
405	0.31	0.70	0.73	0.56	0.15	0.25	0.17	0.18	Medium
407	0.11	0.42	-0.02	-0.19	-0.17	-0.85	-0.85	0.07	None
410	0.28	0.54	0.51	0.03	0.40	0.40	0.40	-0.40	Medium
411	-0.14	0.50	0.49	-0.14	0.61	0.82	0.51	0.39	Medium
412	-0.15	-0.08	-0.53	-0.53	0.16	-0.09	-0.19	-0.60	None
414	0.04	0.30	0.37	0.30	0.02	0.11	0.07	-0.03	None
415	-0.05	0.39	0.27	0.22	0.12	-0.39	-0.61	-0.36	None
416	0.09	0.24	0.10	0.15	-0.31	-0.22	0.31	-0.06	None
418	0.12	0.20	-0.02	0.12	-0.10	0.15	0.67	-0.05	None
422	-0.09	0.45	0.53	0.61	0.58	0.59	0.54	0.22	Long
424	-0.13	-0.05	-0.24	-0.34	0.20	0.16	0.22	-0.31	None
425	-0.37	0.32	-0.17	-0.20	0.65	0.88	0.58	0.21	None
426	0.00	0.20	0.16	-0.19	-0.85	-0.78	0.17	-0.54	None
427	0.24	0.25	0.37	0.03	0.26	0.33	0.07	-0.67	None
428	0.15	0.42	0.19	-0.02	0.25	0.42	0.12	-0.11	Medium
429	0.06	0.26	0.21	0.15	0.48	0.32	0.32	0.12	None
432	0.33	0.40	0.10	0.02	0.60	0.49	0.09	0.52	None
433	0.33	0.30	0.38	0.20	0.35	0.40	0.10	-0.18	None
434	0.23	0.13	0.48	0.25	-0.16	0.03	0.27	0.38	None
436	0.17	0.37	0.28	0.22	0.13	0.24	0.17	-0.18	None

	Summer				Winter				Category
	Rain_rho	Rain03_rho	Rain10_rho	Rain30_rho	Rain_rho	Rain03_rho	Rain10_rho	Rain30_rho	
437	-0.18	0.12	0.21	0.00	0.08	0.05	-0.17	-0.26	None
439	0.00	0.42	0.26	0.15	0.24	0.23	0.22	-0.04	None
440	0.18	0.14	0.24	0.02	-0.02	0.21	0.11	-0.09	None
442	0.12	0.39	0.46	0.16	0.22	0.28	0.24	-0.21	None
443	0.05	0.10	0.11	0.05	-0.05	0.22	0.30	0.07	None
444	0.27	0.27	0.31	0.24	-0.17	-0.32	-0.40	-0.29	None
445	0.03	0.17	0.25	0.26	-0.15	0.23	0.23	0.12	None
446	0.27	0.26	0.40	0.27	-0.23	0.24	-0.30	-0.34	None
447	0.08	0.22	0.36	0.32	0.36	0.31	0.21	-0.17	None
448	0.78	0.29	0.28	0.18	0.39	0.62	0.17	0.28	Short
449	-0.23	-0.07	0.24	-0.06	0.37	0.78	0.15	-0.15	None
450	-0.19	0.07	-0.16	0.06	-0.21	0.67	0.67	0.39	None
451	-0.23	0.43	0.03	0.10	-0.13	0.27	-0.18	-0.31	None
452	0.02	0.20	0.01	-0.13	0.05	0.48	0.48	0.31	None
456	-0.04	0.15	-0.18	-0.27	-0.15	-0.05	-0.42	-0.27	None

D. Suggestions for priority catchments

TOC

Catchment	Catchment description	Focus of investigation
34	Reservoir, located in the West, cluster 3, categorised as “Temperature”, small catchment with little slope, very poorly drained soils, mainly semi-natural land cover, large increase in AAT projected, high TOC, aluminium, iron and manganese medians, high increase in TOC projected due to AAT increase (in combination with reduction in SER)	Increase process understanding – test hypotheses of direct climate change impacts on catchments sensitive to temperature, and for multiple water quality parameter Review control options and ES assessment
214	Reservoir, located in the South, cluster 4, categorised as “Rainfall+Temperature”, small catchment with little slope, exclusively sandstone bedrock, very organic soils, mainly semi-natural land cover, high number of livestock, high medians for metals, TOC, and turbidity, high projected increase.	Increase process understanding – test hypotheses of direct climate change impacts and of land use impacts Test hypothesis of increase in variability and peak concentrations of TOC Review control options and ES assessment
349	Loch, located on the Orkney Islands, cluster 4, categorised as “Rainfall”, “inconclusive” shape, little relief, mostly igneous bedrock with a little limestone, high percentage coverage of improved grassland (~44%), medium to high median values for metals and colour/TOC, small increase projected due to small decrease in SER.	Increase understanding of potential lack of buffering
418	River, located in Northeast Scotland, cluster 1, categorised as “None”, medium catchment size, some organic soils, mainly semi-natural land cover, large numbers of deer, high median pH, medium increase projected.	Increase process understanding – what are main controls of DOC production, release, transfer, and loss?
227	Reservoir, cluster 1, categorised as “Wetup”, “sine” shape, medium catchment in the South, mainly sandstone bedrock, very organic soils, mainly semi-natural land cover, slightly higher median values for metals, turbidity and TOC, large increase for TOC projected due to decrease in SER and increase in AAT.	Increase process understanding – test hypotheses of direct climate change impacts on catchments sensitive rainfall over a longer period, and for multiple water quality parameter Review control options and ES assessment
436	Loch, cluster 2, categorised as “Wetup”, “v-shape”, medium sized catchment on the Outer Hebrides, mainly peat soils, mainly semi-natural land cover, low pH median, large increase in TOC projected.	Increase process understanding, especially climate change effect in combination with acidification reversal.

<u>*E. coli*</u>

Catchment	Catchment description	Focus of investigation
349	Loch, located on the Orkney Islands, cluster 4, little relief, mostly igneous bedrock with a little limestone, high percentage coverage of improved grassland (~44%), projected to increase in LC 3.2, high turbidity median values, slightly above average *E. coli* median, slight increase predicted	Critical source areas – Are the areas projected to increase in capability increasing risk for water quality deterioration?
295	River, large catchment in the Northeast, cluster 5, medium well drained soils, mostly semi-natural land cover with some arable and improved grassland cover, large increase in prime land projected, high coliform and *E. coli* median values, large increase predicted	Critical source areas – Why are concentrations high? Scenario analysis – How could land uses change and what would the impact be? ES assessment and stakeholder engagement – Incentivise water positive behaviour
294	River, large catchment in the Northeast, downstream of 295, cluster 5, mainly well drained soils, high percentage of arable (~11%) and improved grassland (~30%) cover, large increase in prime land (~70%) projected, high coliform and *E. coli* median values, large increase predicted	Scenario analysis – How could land uses change and what would the impact be? ES assessment and stakeholder engagement – Incentivise water positive behaviour
415	Reservoir, cluster 2, medium catchment in the Northeast, adjacent to 295, mainly semi-natural land cover, small increase in prime land and land capable for sustaining livestock projected, overall good raw water quality, low *E. coli* concentrations, small increase predicted	Scenario analysis – How could land uses change and what would the impact be?
414	Reservoir, cluster 1, medium catchment in the Northeast, some sedimentary bedrock and mainly well drained soils, mostly semi-natural land cover and some coniferous forest (~17%), large increase in prime land (~37%) projected, low *E. coli* concentration, large increase predicted	Scenario analysis – How could land uses change and what would the impact be? ES assessment and stakeholder engagement – Incentivise water positive behaviour
227	Reservoir, cluster 1, medium catchment in the South, mainly sandstone bedrock, very organic soils, no agricultural land use despite large proportion of land capable for sustaining livestock, small increase projected, low *E. coli* concentrations, medium increase predicted	Scenario analysis – How could land uses change? Critical source areas – Are areas projected to increase in capability increasing risk for water quality deterioration?

TOC catchments were chosen to represent each climate sensitivity category and get a spectrum of colour annual time series shapes and clusters from the initial cluster analysis.

E. coli catchments were chosen for their close proximity but different current land uses and projections in land capability, giving a good coverage. More detailed information from these catchments could help understand patterns and inconsistencies observed between catchment characteristics and water quality outcomes in terms of FIO contamination, stakeholder preferences and possibilities for working together to achieve positive water quality and ES outcomes.

Two catchments (227 & 349) were chosen for both parameter and could serve to get a more holistic picture.