Risk-based modelling of faecal indicator organism export from agricultural landscapes

Kenneth Porter

September 2018

A thesis submitted for the degree of
Doctor of Philosophy

Biological and Environmental Sciences, Faculty of Natural Sciences

University of Stirling

Declaration of authorship

I, Kenneth D.H. Porter, declare that this thesis has been composed by me and it embodies the results of my own research. Where appropriate I have acknowledged the nature and extent of work carried out in collaboration with others.

....

Kenneth Porter

27th January 2017

Abstract

Microbial contamination of watercourses can threaten ecosystem services related to clean water; for example, recreational bathing, shellfish harvesting and potable water supplies. This is because pathogens associated with faeces from warm blooded animals can cause gastrointestinal illness in exposed human beings. Microbial water quality impacts from point sources associated with wastewater transfer and treatment have been reduced through engineering solutions. However, as these sources of contamination have been reduced diffuse sources have become more important. Diffuse pollution describes water quality impacts originating from accumulations of many small, spatially distributed, inputs. These sources of pollution are difficult to manage because their loading and connectivity to sensitive receptors varies spatially and temporally. The Sensitive Catchment Integrated Mapping Analysis Platform (SCIMAP) is a risk-based approach that has been developed to map sources of diffuse sediment and conservative nutrient pollution allowing for efficient targeting of mitigation efforts which are often expensive and occupy valuable productive land. SCIMAP has been well received within the regulatory community in the United Kingdom and its development to account for diffuse microbial pollution is therefore timely. The primary goal for this thesis was to explore SCIMAP's application to microbial pollution, highlight areas for improvement and work towards a new SCIMAP framework that accounts for microbial diffuse pollution. An initial application of SCIMAP, as it exists, revealed that the time-integrated approach currently employed may be inappropriate for sources of microbial pollution that are likely to vary temporally due to microbial die off. Furthermore, an enhanced description of land use incorporating spatial distributions of the numbers and types of livestock may improve SCIMAP's

performance. Spatial variations in microbial source loading arising from differences in the persistence of *E. coli* (an indicator of faecal pollution) in the faeces of different livestock was investigated within a controlled environment facility. This controlled experiment provided a novel non-linear description of *E. coli* growth in ovine and 2 types of bovine faeces for a period of 30 days post defecation. Potential variation in rainfall induced *E. coli* release from faecal matrices associated, with beef cattle, dairy cattle and sheep were explored using rainfall simulation. An asymptotic model of *E. coli* release with increasing rainfall depth was developed and no difference was discovered in the profile of release from sheep, beef cattle and dairy cattle. Finally lessons from these investigations were combined to propose a framework for an evolution of SCIMAP allowing for a better description of microbial source and transfer risk. This new version of SCIMAP will provide a decision support tool allowing for more efficient targeting of mitigation efforts reducing microbial impacts to important ecosystem services relying on clean water.

Acknowledgments

This work was funded by the Natural Environment Research Council (NERC) via the IAPETUS doctoral training partnership (NE/L002590/1). We are grateful to the Environment Agency for providing spatial datasets licensed under the Open Government Licence v2.0 for chapter 3.

Thanks goes to my supervisors David Oliver, Richard Quilliam and Sim Reaney for all their support throughout the project; to Mark Allen who as director of the IAPETUS doctoral training partnership provided support and inspiration; to all those who helped in the laboratory and field work (Diana Mclaren, Anisha Keswani, Jan Moser, Zarah Pattison); to Ian Washbourne, Lorna English, James Weir, Sylvia Hodgson and Stuart Bradley for technical support; to Tim Paine and all those who attended his Semin[R] series and contributed to discussions on data analysis; to Chris Hodgson who hosted me for a period of research at Rothamsted Research's North Wyke Farm Platform; and to my partner Diana Mclaren who has provided support throughout the project; despite having her own PhD to work on. Thanks must also go to my parents, Ian and Margot Porter, who have helped financially and grammatically throughout my studies and have also provided inspiration via their own passion for science. Finally, thanks go to the Scottish landscape and mountain bikes, whose partnership has kept me sane through difficult times.

Contents

Figures.

Figure 2.1. Theoretical response of faecal indicator organisms at different scales: (a) the survival of FIOs within individual faecal deposits; (b) the burden of FIOs at the scale of individual fields; (c) FIO population variations in the sediment sink; (d) 3 scenarios with differing FIO – discharge relationships (i) FIO peak prior to the discharge peak (sediment FIO stores relatively more important), (ii) FIO peak subsequent to discharge peak (landscape stores relatively more important), (iii) FIO peak concurrent with discharge peak; (e) variation of FIO concentrations at the catchment outlet which are a result of combining the processes represented in (a) to (d).

Figure 2.2. Variable relative importance of landscape and sediment sources to FIO flux under differing meteorological conditions. Under successive rainfall events the landscape store of FIO becomes depleted and stream-bed sediment stores are replenished. Under dry weather the sediment store of FIOs declines due to die-off and deposits from animals replenish the landscape store.

Figure 3.1. Maps illustrating the two study catchments: The Wyre in Lancashire, North West England (left); and The Yealm in Devon, South West England (right). Numbered points indicate sample locations and associate with the sample locations indicated in figure 2 and table 2. River data is from an Ordnance Survey MasterMap Topography layer.

Figure 3.2. A bar plot illustrating the proportions of the contributing area associated with each sample point occupied by different land cover types. Please refer to online version for colour.

Figure 3.3. SCIMAP fitted results for (a) the Yealm and (b) the Wyre. The top panels show hexagonally binned scatterplots depicting how model performance changes with changing the risk weighting for each land cover. The colour of the hexagonal bin depicts how many simulations fall into that part of the plot. The bottom panels show boxplots which depict the variation in the risk weighting of the 1% best performing simulations.

Figure 3.4. An ordination plot showing the dissimilarity in land cover mosaic across the contributing catchments associated with sample points from the Yealm (grey) and Wyre (black). Increasing distance between points illustrates increasing dissimilarity in land cover make up between catchments.

Figure 3.5. Boxplot illustrating the variability in the concentration of *E. coli* across samples from the Wyre catchment. Letters above the boxes indicate significant differences in FIO concentrations between sites as determined by a Dunns test.

Figure 3.6. Boxplot illustrating the variance in performance associated with the 1% best performing simulations when observed data is subset according to season.

Figure 3.7. A scatterplot of optimum risk weighting identifiability and percent coverage of the associated land cover.

Figure 4.1. Scatter plot of *E. coli* concentrations through time. Where growth occurred lines are predictions of nonlinear (asymptotic exponential) mixed effects modelling. Where no growth was apparent lines illustrate linear mixed effects models. Dashed lines indicate 95% confidence intervals derived from normal non parametric bootstrap. Where the prediction is coloured black there was no

improvement in model performance (change in AIC <2) when the present average/present average +2°C treatment was incorporated.

Figure 4.2. A scatter plot of *E. coli* concentration growth rate against moisture content. Colours indicate present average or present average +2°C temperature treatments.

Figure 4.3. Boxplot showing logit transformed initial moisture content in the faeces of three livestock types from two seasons. Different letters and colours illustrate where results of a Tukey post-hoc test revealed differences between livestock and season combinations. Y axis labels have been back transformed to improve interpretability.

Figure 5.1. Rainfall simulator construction. From the top left moving clockwise: placement of soil boxes underneath a spray nozzle; water pressure and flow gauges; rainfall simulator construction to prevent interference from wind.

Figure 5.2. The left plot shows the spatial variation of simulated rain within the rainfall simulator; boxes indicate placement of soil-runoff boxes and the circles indicate the placement of rainfall collection pots. The distribution of droplet size is shown in the right hand figure.

Figure 5.3. Boxplot of E. coli die off in the mains water feeding the rainfall simulator and in de ionised water.

Figure 5.4. Discharge of surface runoff from soil boxes across stock type treatments and through time.

Figure 5.5. *E. coli* load in surface run off from soil boxes through time.

Figure 5.6. Logit transformed *E. coli* release as a percentage of the total faecal burden applied to each soil-run off box. The y axis has been back transformed to enhance interpretability. The solid line illustrates the results of a asymptotic mixed effect model. Dashed lines are 95%, bias-corrected and accelerated (BCa), bootstrap prediction intervals.

Figure 5.7. Dot plot of model random effects (deviation of individual faecal deposits from the population model) against stock type.

Figure 6.1. Scatter plot illustrating how initial concnetration of *E. coli* and the extent to which *E. coli* can grow in a feacal deposit contributes to an overall source hazard weighting. Stock season combinations occuring in the top right of the plot will have the highest source hazard; stock season combinations at the bottom left of the plot have the lowest source hazard weighting. An overall source hazard weighting is therefore the distance form the origin.

Figure 6.2. From left to right: bar plot of median daily rainfall amount; bar plot of average number of days with rain.

Figure 6.3. Scatter plot of a hazard weighting reflecting the relative likelihood of rain on a day in a given month against a hazard weighting associated with the amount of rain that occurs on a single day in a given month. A hazard weight that reflects both theses aspects of mobilisation hazard is then the distance from the origin.

Figure 6.4. The bottom bar plot illustrates stock type and month differentiated hazard weightings that capture aspects of hazard associated with the concentration of *E. coli* in fresh faeces; the extent to which *E. coli* grows in faeces;

the relative chance of rain on a day of a given month; and the amount of rain that falls in a single day of a given month. The top bar plot shows how source hazard and mobilisation hazard contribute to the overall hazard weight in different months.

Tables.

Table 3.1. A description of SCIMAP land cover classes and how they are derived from CEH LCM land cover classes.

Table 3.2. Catchment characteristics of each of the sub-catchments investigated. Connectivity is defined as the lowest value of topographical wetness index along a flow path as per Lane *et al.* (2009). Number of samples indicates the number of records remaining after sub setting all the available data by the days where flow is >60% of the highest flow recorded.

Table 3.3. Table summarising the influence of land cover risk weighting on SCIMAP performance. Mean risk weightings and associated standard deviation for the 1% best performing models. *p* value indicates the results from a *t*-test and the confidence with which we can reject the null hypothesis that there is no variation in model performance as a result of the risk weighting assigned to a land cover type.

Table 4.1. Controlled environment facility settings.

Table 4.2. Average (n=5) initial, maximum, and day of maximum *E. coli* concentration for faeces from three different livestock types under four temperature regimes. All values are given as $\log_{10}$ CFU/g dry faeces.

Table 4.3. Table of model parameters associated with asymptotic models for each of the livestock types. Numbers in parentheses indicate lower and upper bounds of a 95% confidence interval.

Table 5.1. Summary information for faecal deposits; values in parentheses are standard deviations. Model random effect is the deviation of an individual faecal deposit from the population model. Here the standard deviation of the random

effects associated with each stock type are shown to demonstrate individual variability associated with different stock types.

1. General introduction

Microbial contamination of watercourses can threaten ecosystem services related to clean water for example recreational bathing water, shellfish harvesting, and potable water supply. Improvements in microbial water quality have been achieved through engineering solutions associated with point sources of pollution such as sewer network overflows and outlets from waste water treatment works. However, as inputs from point sources have been reduced diffuse sources have become more important. Diffuse sources of microbial pollution provide inputs that by themselves do not have a great impact on water quality, but the accumulated impact of many inputs has the potential to impact ecosystem services relying on clean water. Diffuse sources of microbial pollution often originate from agricultural activities and include direct deposits of faeces by livestock and spreading of manures and slurry; although wildlife can also provide a source of microbial contamination. Diffuse sources of microbial pollution are difficult to manage because their source loading and connectivity to sensitive receptors varies spatially and temporally. Pollution mitigation measures such as fencing of water courses and vegetative buffer strips are costly and can occupy valuable productive land. Therefore, mitigating infrastructure should be targeted toward areas where they will provide the most improvement in water quality. In the fields of diffuse sediment and nutrient pollution the Sensitive Catchment Integrated Mapping Analysis Platform (SCIMAP) has proven valuable in identifying areas that contribute to diffuse pollution. The overarching goal of this thesis is to explore the application of SCIMAP to the problem of catchment microbial dynamics.

1.1 Aims and Hypotheses

(i) To investigate the performance of the current iteration of SCIMAP when it is applied to diffuse microbial pollution and identify areas for improvement.

Hypotheses: SCIMAP, as it exists, considers conservative nutrient and sediment pollution and is therefore unlikely to capture temporal variations in source risk of microbial pollution due to microbial die off.

(ii) Under controlled conditions, replicate the proliferation of *E. coli* seen within faecal deposits in field trials and study *E. coli* survival in the faeces of three types of livestock in parallel.

Hypotheses: There will be a greater magnitude of *E. coli* growth: (i) under warmer conditions more conducive to *E. coli* survival and (ii) within bovine faeces which have a higher water content than ovine faeces.

(iii) Create livestock type differentiated *E.coli* release profiles under simulated rainfall.

Hypothesis: *E. coli* is released more rapidly from bovine faeces which have a higher moisture content.

(iv)Finally, the culmination of the thesis will be a framework with which information on the persistence of FIOs in faeces and FIO release from faeces under rainfall can be used to adapt the SCIMAP method for diffuse microbial pollution.

1.2 Thesis Structure

The thesis is structured to guide the reader through the research questions outlined above. The state of the knowledge in the field of catchment microbial dynamics is discussed in chapter 2 and concludes with a discussion surrounding the need for predictive modelling frameworks that address different and diverse problems from forecasting pollution to informing the deployment of mitigation measures. The SCIMAP modelling framework is put into context with existing frameworks for modelling the dynamics of microbial transfer through the environment and the need to adapt SCIMAP for diffuse microbial pollution is discussed.

There are 3 data chapters. In order to reflect the demands of contemporary academic writing each chapter is written as if it were appearing in an academic journal. Therefore, each chapter is written to allow it to stand as an individual piece of writing and is structured with its own abstract, introduction, methods, results, and discussion sections.

Chapter 3 applies SCIMAP in its present form to diffuse microbial pollution and provides a benchmark to test future developments. It also highlights weaknesses in the current approach and informs the rest of the thesis. The SCIMAP framework has its foundations in the source-mobilisation-delivery-impact continuum and this provides a narrative for its development to account for diffuse FIO pollution. To address some of the problems associated with SCIMAP's treatment of source loads of diffuse pollution chapter 4 investigates the persistence of a commonly used indicator of faecal pollution, *E. coli*, in the faeces of different livestock under environmentally relevant temperature conditions that vary diurnally and are typical of temperatures experienced in spring and summer. Chapter 5 investigates the profile of *E. coli* release from the faeces of different livestock under simulated rainfall.

New knowledge gained through the experimentation in chapter 4 and 5 is synthesised in chapter 6 to develop a concept for developing SCIMAP for application to diffuse FIO pollution. Challenges and next steps are discussed.

2. Faecal indicator dynamics in the catchment continuum: recent developments and future research challenges for microbial compliance parameters.

2.1. Abstract

Agricultural landscapes, through the management of livestock and their manure, provide a source of faecal pollution to the environment, which can impact upon important ecosystem services such as clean and safe bathing, drinking and shellfish harvesting water. The potential impact of diffuse faecal pollution is recognised in legislation throughout the world; for example, via the Bathing Water Directive in the EU and the Clean Water Act in the USA. Regulators monitor microbial water quality using faecal indicator organisms (FIOs) as an indicator of faecal pollution. In order to meet stringent health-based standards set out in legislation there must be the capability to predict water quality impairment, identify impaired water bodies and identify opportunities for water quality improvement. The aim of this critical review was to investigate to what extent previously identified challenges have been addressed and determine opportunities for further research. The review identified that knowledge of FIO fate and transfer in catchments has advanced in recent years, but that research is still required to fully understand the dynamics of FIO regrowth in faeces following deposition and factors determining variable patterns of FIO persistence in soil. In addition, field relevant information regarding persistence of FIOs in sediment is required to supplement recent laboratory investigations. Understanding the episodic flux of FIOs remains a challenge and developing FIO release kinetics from faecal, soil

and sediment matrices may allow for further understanding of the emergence of faecal pollution following rainfall events. There is also a need to develop understanding of how landscape sources of FIO connect to watercourses and to develop tools that allow mitigation efforts to be spatially targeted.

2.2 Introduction

Microbial pollution of watercourses has gained increased attention in recent years, with growing recognition of the breadth of impacts it can have on both the environment and human health (Eze *et al.* 2014). The contamination of receiving waters with pathogens has the potential to threaten the ecosystem services that clean and safe waters provide, such as recreational, drinking and shellfish harvesting services. Environmental monitoring and quantification of faecal indicator organisms (FIO) such as *Escherichia coli* and intestinal enterococci provides an internationally accepted regulatory framework to help understand levels of faecal pollution in the wider environment, though their relationship with specific pathogens is a topic of much debate (Bradshaw *et al*. 2016; Pachepsky *et al.* 2016). Environmental monitoring of pathogens is less common largely due to cost constraints, their lower abundance in the environment and the wide range of microorganisms that could be chosen as a specific target for quantification.

Common sources of microbial contamination of the aquatic environment include agricultural activities and end-of-pipe outflows associated with sewage treatment. In agricultural catchments, there is a tendency for rainfall to promote diffuse inputs of FIOs to receiving waters, for example via hydrological pathways such as overland flow and subsurface drainage (Murphy *et al.* 2015). At small scales, individual diffuse inputs (e.g. livestock faecal deposits) might pose little risk to the aquatic environment, but with increasing scale the cumulative impact of diffuse contributions can seriously threaten water quality. In catchments dominated by urban environments there is a greater risk of microbial water pollution from point sources at waste water treatment works (Kay *et al.* 2010).

The potential risk that FIOs pose to ecosystems service delivery is recognised in legislation throughout the world. For example, the United States Clean Water Act requires a Total Maximum Daily Load (TMDL) assessment of a pollutant to be carried out once a water body has been identified as being impaired. In 2014 microbial pollution caused more U.S. water quality impairments than metals, nutrients, oxygen depletion and sediment (Pandey *et al.* 2014). In the European Union, the Water Framework Directive (2000/60/EC) requires identification of impacted water bodies and implementation of ‘programmes of measures’ (POMs) to improve water quality, including meeting the regulatory standards set out in the Bathing Water Directive (2006/07/EC) and Shellfish Waters Directive (2006/113/EC), which aim to reduce public health risk by minimising exposure to water contaminated with faecal microbes. In response, a holistic catchment-based approach is needed to understand pollution inputs and their environmental interactions and thus enable effective management of water quality for multiple benefits. This approach has led to increased efforts to promote the integrated catchment management (ICM) approach (McGonigle *et al.* 2012; Falkenmark 2007) and has helped to drive forward multi-pollutant research agendas and raise the profile of FIOs as a priority pollutant alongside more widely studied nutrients, such as nitrogen (N) and phosphorus (P). Furthermore, the source-mobilisation-delivery-impact (SMDI) continuum developed for diffuse nutrient pollution (Haygarth *et al.* 2005) is useful for conceptualising how generic diffuse pollutants, including FIOs, interact with the environment to become a threat to water quality at the catchment scale. To ensure effective prediction, management and reduction of FIO concentrations in receiving waters, a robust understanding of FIO behaviour throughout the entire SMDI continuum is required.

The field of catchment microbial dynamics represents an interdisciplinary and multi-scaled research agenda cutting across rural and urban environments, linking science and policy, and necessitating the integration of microbial fate and transfer dynamics from the hillslope through to the coastal environment. The state-of-knowledge and breadth and scope of research opportunities associated with this emerging research agenda were outlined by Kay *et al.* (2007b) who provided a roadmap of challenges and priorities for the research community to tackle in order to advance this field. Therefore, rather than provide a comprehensive and exhaustive evaluation of the literature in the field of catchment microbial dynamics, the aim of this review is to provide a critical and timely update on recent international research progress, assess how those challenges identified by Kay et al. (2007b) are being met and to highlight the pressing issues that still remain or have since emerged. Critically, this review will cover the advances made through research in agricultural systems and focus on FIOs as a regulatory compliance parameter rather than detail research needs associated with specific pathogens.

2.3 An existing research agenda

In summarising the state-of-the-science, Kay *et al.* (2007b) concluded with seven core observations. Briefly these conclusions were:

1. studies show reductions in FIO flux following implementation of on-farm measures such as exclusion of livestock from watercourses;
2. export coefficients, whereby parts of the landscape are assigned a coefficient describing its ability to deliver FIOs to watercourses, provide a means to spatially target mitigation measures however a limited

understanding of the episodic nature of FIO transfer provides challenges in this respect;

3. predictive approaches, such as the use of export coefficients, operate at a scale of >1km^2, however features influencing FIO transfer are often far smaller making the use of these approaches for spatially targeting of BMPs challenging;
4. a limited understanding of the survival of FIOs in different environmental matrices and a lack of field-relevant information on the transfer of FIOs to watercourses hampers predictive modelling of instream FIO concentrations;
5. monitoring exercises often utilise a regular monitoring regime which is limited in its potential to explore the episodic nature of FIO export, a key knowledge gap;
6. the assumption that FIOs associate with sediment is challenged by a limited understanding of how FIOs partition with sediments of different characteristics;
7. cooperation between water and agricultural sectors as facilitated by changes in the structures providing financial support for farmers provides opportunities for the control of FIOs in the EU.

These observations provide a benchmark against which research progress can be evaluated. These observations have been categorised into four themes for the purpose of this review, namely: (i) fundamental data needs, (ii) event dynamics, (iii) landscape drivers of FIO risk and (iv) managing diffuse FIO pollution. For each theme we provide a critical update on the research developments and the advancing state of knowledge in an effort to identify where challenges remain, and new opportunities are emerging in the field of catchment microbial dynamics.

2.4 Fundamental data needs

A number of issues identified above fall under the general theme of 'fundamental data needs'. Key knowledge requirements have been highlighted including a need for further understanding of the variability in FIO concentrations within fresh faeces and FIO population dynamics in the faeces of livestock, in soil and within stream bed sediment. An additional knowledge requirement is also discussed: the potential contribution of FIOs from wildlife faeces to watercourse pollution (Kay *et al.* 2007b).

Diffuse FIO pollution can be described as source limited because the number of FIOs available for transport is finite and varies through time depending on microbial die-off (Evanson and Ambrose 2006) and management practices (Donnison *et al.* 2008; Oliver *et al.* 2014). Therefore, knowledge of the proliferation and persistence of FIOs in the landscape is a key first step in understanding the potential for catchment transfers of FIOs (figure 2.1a and 2.1b). The size of the landscape store of FIOs can be defined as the sum of the daily input of FIOs and the FIO inputs from previous days, which will be declining as a result of die-off (Oliver *et al.* 2010). Determining the initial concentration of FIOs in fresh faeces remains a challenge due to high variability seen within and across studies.

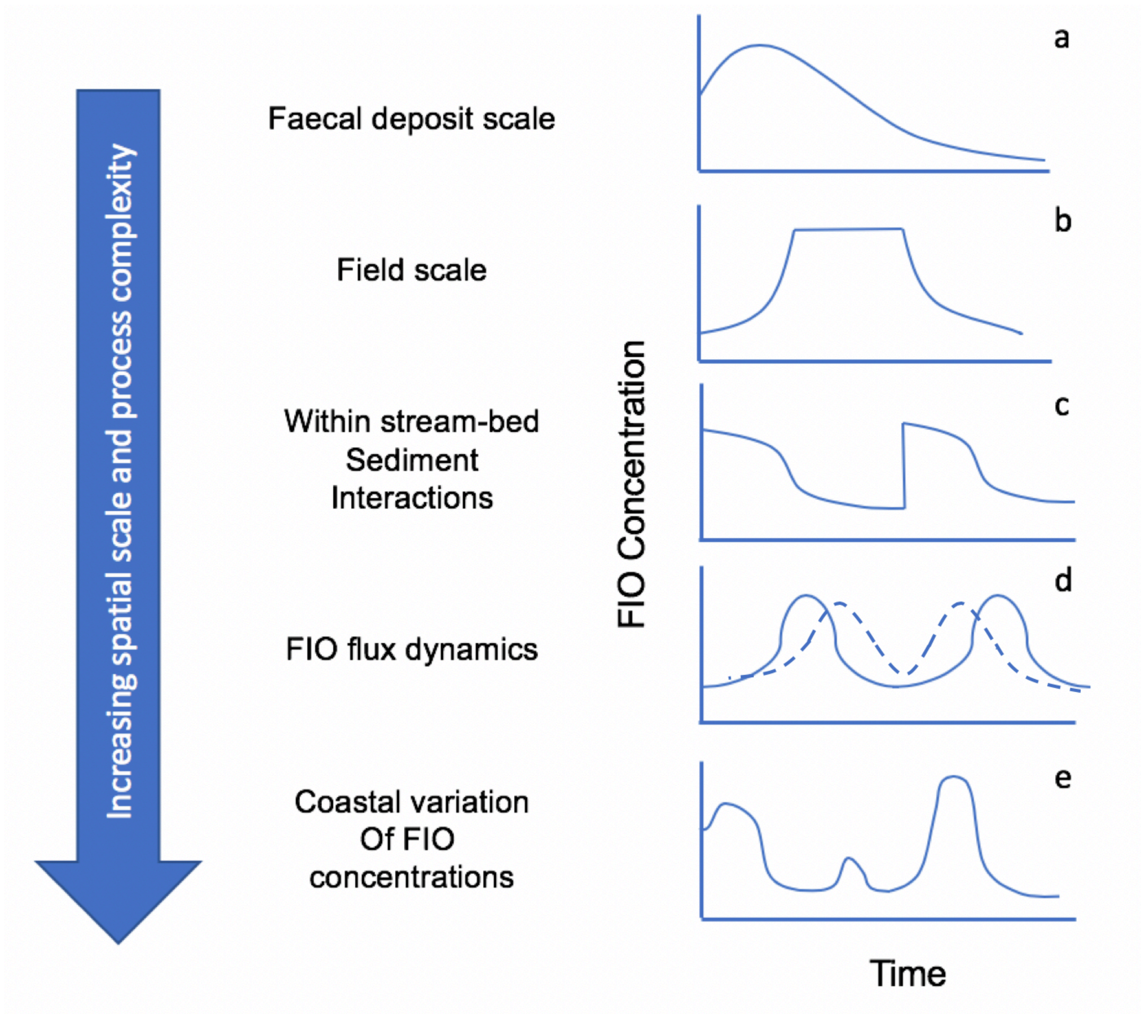

Figure 2.1. Theoretical response of faecal indicator organisms at different scales: (a) the survival of FIOs within individual faecal deposits; (b) the burden of FIOs at the scale of individual fields; (c) FIO population variations in the sediment sink; (d) 3 scenarios with differing FIO – discharge relationships (i) FIO peak prior to the discharge peak (sediment FIO stores relatively more important), (ii) FIO peak subsequent to discharge peak (landscape stores relatively more important), (iii) FIO peak concurrent with discharge peak; (e) variation of FIO concentrations at the catchment outlet which are a result of combining the processes represented in (a) to (d).

For example in dairy cattle faeces Moriarty *et al.* (2008) observed concentrations of *E.coli* ranging between 2.6×10^{3} to 2.0×10^{7} g^{-1} CFU (wet weight); Muirhead (2009) observed a range between 1.7×10^{3} and 4.9×10^{6} and Soupir *et al.* (2008) reported a variation of 1.16×10^{5} to 8.63×10^{7}MPN g^{-1} (wet weight). These examples report wet weight counts and are difficult to compare as variations in moisture content are likely to alter the weight of samples and thus counts per weight. Dry matter content should also be reported to allow conversion between wet and dry weight counts for comparison (Oliver *et al.* 2014).

It is important to determine what drives the variability of FIO concentrations in fresh faeces in order to understand the landscape reservoir of FIOs. Animal age has been shown to influence FIO inputs to land with average *E. coli* concentrations of 1.62×10^{7} in adult sheep and 6.04×10^{8} g^{-1} (wet weight) in lambs having been recorded (Moriarty *et al.* 2011). However, there is little reliable information on the different contributions of FIO from animals across different age bands. Hormones associated with host stress may influence the survival of bacteria residing in the gut of livestock (Freestone *et al.* 2008). Diet has also been implicated as a factor with silage fed during housed periods reducing the burden of FIOs within faeces compared with the diet typical of grazing regimes on pasture (Oliver *et al.* 2014). Silage is more acidic than grazed grass reducing rumen pH which may impact upon micro-organisms in the gut (Donnison *et al.* 2008). Diet is unlikely to be the only factor influencing shedding rates by cattle and removal of cattle from faecally contaminated pasture is likely to have reduced the opportunity for FIO re-ingestion. Development of FIO fate and transfer models requires data and understanding of the FIO burden within the faeces of different animals under different conditions. Understanding of initial concentrations of FIO in faecal deposits is growing; however, limited data means model development often requires the use of parameters reported in international literature, which may not represent local conditions (Oliver *et al.* 2014). Models of FIO fate and transfer may have to be calibrated to individual circumstances through collection of new data which can be time consuming and expensive. Thus, there is a need for a systematic review and meta-analysis to bring disparate datasets together and create a library of global FIO concentrations across livestock dominated regions. Such analysis could begin to address important questions as to whether concentrations remain stable in different regions over time, follow trends or differ dramatically year on year. Subsequent development of models describing how the shedding rates of livestock vary would also be a useful addition to existing process-based models.

The persistence of FIOs in faecal deposits do not necessarily follow normally expected decay rates within faecal deposits (Kay et al., 2007b). Wang *et al.* (2004) determined that the first order kinetic often used to approximate FIO persistence outside the alimentary canal described *E. coli* die-off only between 3 and 20 days post defecation and Stocker *et al.* (2014) showed that a Weibull model, describing initial slow inactivation followed by more rapid inactivation, rather than a linear semi-logarithmic is preferable for the description of FIO persistence in a faecal matrix. A key weakness in FIO fate and transfer models is a failure to recognise complex FIO survival dynamics such as the potential regrowth of FIOs outside of the alimentary canal. Further evidence has emerged over recent years highlighting the importance of including a consideration of FIO regrowth. For example, Soupir *et al.* (2008) studied dairy cow faeces and observed *E. coli* growth for 7 and 4 days, and enterococci growth for 13 and 4 days in spring and summer respectively. In New Zealand, enterococci growth occurred within sheep faeces during all four seasons with peak concentrations reached after 11, 28, 14 and 24 days in spring, summer, autumn and winter respectively while *E. coli* growth in sheep faeces was observed in spring, summer and autumn only with peak concentrations reached after 11, 8, and 14 days respectively (Moriarty *et al.* 2011). Soupir *et al.* (2008) explored the limitations of first order die-off kinetics and noted that because it does not capture regrowth the statistically derived intercept can over-estimate the initial concentration of FIOs. In addition, if an experimentally derived initial concentration is used within the first order die-off kinetic the persistence of FIOs is likely to be under predicted because the equation is shifted downward. While progress has been made in understanding growth of FIOs there remains a need for more complex models of die-off that account for variable rates of growth and death under different environmental conditions (Soupir *et al.* 2008a). These models need to combine durations and rates of growth to able to describe the magnitude of FIO population growth occurring within faecal matrices. It is possible that small differences in the specification of models describing FIO persistence in individual deposits may lead to more than just trivial differences in landscape scale predictions of FIO burden. However, the use of more complex FIO survival models would need to be justified by comparing their performance against predictions from simpler linear decay functions.

Over recent years progress has been made in addressing the need for more complex models of FIO persistence. For example, Oliver *et al.* (2010) observed *E. coli* growth in cattle faeces for up to ten days post defecation and incorporated average growth into a field scale FIO burden model. However, this first approximation of FIO regrowth utilised a static rate of growth. The rate of growth is likely to vary spatially and temporally due to variations in abiotic factors such as temperature, humidity, rainfall and solar radiation (Soupir *et al.* 2008a), and this should be captured within new models of FIO persistence in faeces. Martinez *et al.* (2013) applied the Q_{10} approach to *E. coli* survival in order to understand how rates of growth and death vary according to temperature. The Q_{10} temperature coefficient is a measure of the rate of change of a system as a result of increasing the temperature by 10°C. This coefficient can be used to correct die-off models to account for the effect of temperature. The method has been applied to both the growth and die-off phase of FIOs. The approach shows promise, but further exploration is limited by available datasets. For example, without further information regarding the duration of the growth phase under different conditions and variability in animal shedding rates it is difficult to apply these coefficients in models describing the variation in landscape burden due to the growth and decay of FIO populations.

There are limited data regarding the potential for wildlife to contribute FIO pollution to watercourses. For example, Daszak (2000) highlighted the potential for wildlife to contribute to the landscape store of FIOs but important information on the survival of FIOs within wildlife faeces is scarce (Guber *et al.* 2015). More recently there has been increasing momentum with respect to addressing this knowledge gap. For example, *E. coli* and enterococci have been shown to survive and grow once inoculated into Canada Goose faeces (Moriarty *et al.,* 2012) and Guber *et al.* (2015) studied FIO persistence in the faeces of White-Tailed deer and reported differences in FIO survival dynamics from those observed in the faeces of domestic animals that have been more comprehensively studied. This highlights a need to understand variabilities in the persistence of FIOs occurring naturally in the faeces of wildlife. Depending on the niche space wild animals fill they will be more or less likely to contribute FIOs to watercourses and further information on the survival of FIOs within the faeces of a variety of wild animals would be useful if we are to understand the contribution wildlife make to faecal pollution of watercourses. The impact of wildlife on microbial water quality has the potential to be as significant as some farm management practices, such as spreading of collected manure or grazing of livestock on pasture (Muirhead *et al.* 2011) especially where catchments have a higher proportion of non-agricultural land cover (Kiefer *et al.* 2012). Furthermore, mitigation efforts such as vegetated buffer strips have the potential to attract wildlife leading to a concentration of wildlife faeces on stream margins. While attention toward this knowledge gap is growing, research is needed before the contribution wildlife make to the landscape reservoir of FIOs is fully understood. This can perhaps complicate the communication of risk management to farmers and landowners if these catchment stakeholders then feel that wildlife sources are receiving less regulation than farmed animals despite perceptions that wildlife contributions could impact water quality. Evidence is therefore needed to help underpin communications about proportional contributions of different animals and livestock to the different catchment communities responsible for managing land and water.

FIOs are able to persist in the soil contributing to the landscape reservoir of FIO (Texier *et al.* 2008; Muirhead *et al.* 2009) and the ability of FIOs to persist in soils is important as it provides the opportunity for their transfer to watercourses over successive rainfall events. Muirhead *et al.* (2009) compared the faecal and soil reservoirs and found the soil reservoir to be only two orders of magnitude less than the faecal reservoir; 3.4×10^4 to 6.3×10^6 MPN m^{-2} (wet weight) in the soil compared with 4.2×10^7 to 1.4×10^{11} MPN m^{-2}. It is important to understand the variation in the prevalence and persistence of FIOs within this store if the landscape reservoir of FIOs is to be determined. Texier *et al.* (2008) studied *E. coli* persistence within cow faeces and in surrounding soil and *E. coli* survived for two months while a faecal matrix existed, it declined when the faecal matrix broke down. Meanwhile *E. coli* populations in the top soil layer (0-5 cm) remained at a constant concentration of 10^3 to 10^4 cells g^{-1} dry soil and *E. coli* prevalence in the root zone (5-25 cm) appeared to vary with soil type and was reduced in well drained soils (< 10^2 cells g^{-1} dry soil compared with 10^3 to 10^4 cells g^{-1} dry soil in poorly drained soils). Evidence suggests that the soil reservoir of FIOs depends on soil characteristics (Texier *et al.* 2008), such as organic matter content (Oliver *et al.* 2005a), and some progress has been made in understanding the relationship between soil characteristics and FIO persistence in soil stores. VanderZaag *et al.* (2010) discovered that *E. coli* was able to survive for two times longer in soils of a field amended with dairy cattle manure and inorganic fertiliser compared with soils from an adjacent riparian zone. Bucci *et al.* (2015) investigated the impact of freeze/thaw cycles on the survival of *E. coli* in soil and observed a decrease of up to four orders of magnitude in faecal coliforms (FC) and enterococci following winter freezing and Oliver *et al.* (2012) report a drop of *E. coli* concentrations at a headwater catchment during freezing conditions. However, Adhikari *et al.* (2007) report *E. coli* survival of six months in frozen soil. There is, however, little data regarding the influence of freeze/thaw and wetting/drying cycles on the fate and transfer of FIOs in the environment. This represents a critical limitation in current understanding because these cycles are likely to determine the extent to which FIOs persist and are transported over multiple rainfall events or snowmelt periods.

Further knowledge regarding the influence of soil temperature, type and nutrient status on the persistence of FIOs in the soil store would be useful for understanding the spatial heterogeneity of the landscape burden of FIOs and the extent to which FIOs can move toward watercourses over successive rainfall events.

While survival of manure derived FIOs in soil has been studied, evidence is emerging that suggests *E. coli* are able to develop into naturalized populations adapted to conditions in the environment. Differences in the genome of *E. coli* derived from faecal deposits and soil have been observed and *E. coli* have been shown to compete for niche space in the soil, exploiting a wide range of nutrient sources at temperatures as low as 15°C (Texier *et al.* 2008; Brennan *et al.* 2013). The survival dynamics of these naturalised populations of *E. coli* are influenced by the indigenous microbial community and environmental conditions such as soil temperature, moisture and nutrient status (Ishii *et al.* 2010). The presence of a naturalised population of FIOs in the soil may impact on the performance of FIOs as indicators of faecal pollution in water courses. It is possible that transfer of these organisms to watercourses may be perceived as faecal contamination. It would be useful to understand how these naturalised populations of *E. coli* interact with their environment in comparison with faecally derived *E. coli* in order to realise how the presence of these organisms influences the perceived level of faecal contamination in watercourses.

FIOs, once delivered into streams, are able to settle and survive in stream bed sediments creating an important reservoir of FIOs (Muirhead *et al.,* 2004; Pachepsky and Shelton 2011; Pandey *et al.* 2013) which can be remobilised under high flow conditions (Droppo *et al.* 2011; Piorkowski *et al.* 2014) or resuspended by livestock trampling. Current process-based modelling tends not to account for FIO persistence in a sediment reservoir or FIO/sediment association and dissociation. Understanding the persistence of FIOs in the sediment reservoir will improve our ability to assess and predict microbial water quality by providing information on this legacy store of FIOs that have the potential to contribute to pollution. It has been suggested that FIOs are able to survive in stream-bed sediment because it provides nutrients and a surface on which to grow (Shelton *et al.* 2014). Some progress has been made in determining what drives variable FIO population dynamics in sediment utilising laboratory experiments. For example, Shelton *et al.* (2014) investigated the impact of nutrient concentrations on FC population change and found a nitrogen peak stimulated FIO growth with subsequent population decline at the rate prior to the nutrient spike. An influence of bed shear stress has been observed with thicker substrata enhancing FIO population survival (Walters *et al.* 2014). It is also likely that predatory bacteria indigenous to stream-bed sediment will impact FIO persistence by grazing and depleting their numbers (Walters *et al.* 2014). Temperature has also been shown to have an impact with slow population decline at 4°C. In addition, fine particle and organic carbon content reduces inactivation rates and sensitivity of inactivation to temperature (Garzio-Hadzick *et al.* 2010). The experiments by Shelton *et al.* (2014), Walters *et al.* (2014), and Garzio-Hadzick *et al.* (2010) provide valuable information on FIO survival in sediments. However, field relevant studies are required to assess the appropriateness of extrapolating findings from these laboratory studies to field conditions.

Kay *et al.* (2007b) recognised that more complex models of FIO persistence in the landscape are needed and further knowledge of FIO die-off has been achieved. However, understanding a period of FIO population growth immediately following deposition in the environment remains a challenge. The lack of a large dataset representative of *E. coli* prevalence and persistence in livestock faeces at a national to international scale provides a significant barrier to the development of robust FIO risk prediction tools. In addition, the potential for wildlife to contribute to FIO water pollution has been raised and this is now emerging as a new research agenda. The variability of FIO persistence in soil and river sediments has been a topic of investigation and the ability of soils and sediments to provide an accommodating habitat for FIOs as they are transferred to sensitive receptors over multiple rainfall events is a research priority. While preliminary laboratory results have suggested factors that influence FIO persistence in stream bed sediments field relevant data are now needed to assess the appropriateness of extrapolating these findings into the real world.

2.5 Event dynamics

Landscape accumulations of FIOs can be mobilised by rainfall. Once mobilised, rainfall-runoff can transport FIOs across and through the soil, and if delivered to an aquatic receptor this can lead to wider contamination of watercourses. It is important to understand the episodic flux of FIOs following rainfall events (figure 2.1c) because it is widely recognised that storm events account for >90% of the total load of FIOs transported to watercourses (Kay *et al.* 2007b; McKergow and Davies-Colley 2010; Kay *et al.* 2010). More recent evidence has suggested that even small increases in discharge (0.25 - 0.8 L s^{-1}), that still represent a hydrograph ‘surge’ but at much smaller magnitude under base flow conditions, can also lead to an order of magnitude increase of FIO concentrations in headwater streams (Oliver *et al.* 2015). Despite some focus on FIO flux dynamics during storm events in recent years this remains an area of much uncertainty and an area of significant opportunity in terms of substantiating an evidence base concerning how scale of study, and storm and catchment characteristics can influence FIO flux responses. The impact of different storm typologies and antecedent conditions on the transfer of FIOs to watercourses and the role hysteresis can play in predicting FIO flux is unclear. Perhaps a barrier that prevents a catalogue of empirical storm datasets being produced is the lack of perceived investigative novelty associated with such monitoring campaigns relative to research exploring the mechanisms of FIO fate and transfer. Alternatively, it may be due to the nature of the resources required to enable the capture of good quality storm pollutant data through hydrographs, being labour intensive and technically challenging if sampling in remote locations.

Hysteresis analysis is a useful way to describe how the peak in discharge relates to the peak of a pollutant such as FIOs (Williams 1989) and intra event and intra catchment variations in the FIO – discharge hysteresis patterns observed (i.e. peak FIO load occurs either before, after and at the same time as the discharge peak) highlights the complexity of the processes involved in the flux of FIOs throughout catchment systems. For example, studies associated with large catchments have observed *E. coli* concentration peaks prior to discharge peaks (McKergow and Davies-Colley 2010) whereas small headwater catchments have shown no clear trends of hysteresis (Oliver *et al.* 2015). The hysteresis responses observed are likely to be influenced by scale making comparison between studies difficult. A useful resource for the research community would be the development of a database of hysteresis rules associated with different catchment types, storm typologies and management regimes. Further understanding of what drives variation in FIO – discharge hysteresis patterns will be critical to help develop robust and transferable rules of cell emergence during different types of storm event.

The general agreement that an order of magnitude increase in FIO load will occur during high flow conditions suggests that rainfall is activating a rapid hydrological pathway such as overland flow to facilitate the efficient transport of FIOs (Collins *et al.* 2005). Using a flow path separation technique (loadograph recession analysis) Murphy *et al.* (2015) associated high loads of FIO with faster flow paths including overland flow, preferential flow, conduit flow and large fissure flow highlighting these as key pathways for FIO transport to watercourses immediately following rainfall. Overland flow dominated transport of FIOs during high flow may lead to diffuse landscape sources becoming more important than urban point sources which are likely to be relatively more important during base flow conditions (Kay *et al.* 2010).

Drainage of agricultural land is another conduit for faecal pollution providing rapid transport at times of heavy rainfall. Oliver *et al.* (2005) investigated the relative importance of drained and undrained pasture and while drainage reduced overland flow there was little difference in FIO export from the two treatments as in the drained treatment overland flow was routed to mole and tile drains which routed cells to the watercourse. Additionally, Falbo *et al.* (2013) and Buchanan *et al.* (2013) found that roadside ditches adjacent to agricultural land provide conduits for *E. coli* transfer to watercourses that bypasses landscape features which may attenuate FIOs. Therefore, management of rapid pathways that bypass the general landscape may provide an opportunity for water quality improvement.

Complex patterns of FIO – discharge hysteresis highlights the complexity of the processes driving FIO flux and understanding what drives the variable hysteresis patterns observed will be useful. It has been suggested that the relative importance of landscape and sediment sources of FIO between different catchments may drive some of this complexity (Hudson 2003). Wilkinson *et al.* (2011) investigated stormflow increases of *E. coli* concentration and developed a model that made better predictions of storm flow dynamics within large catchments compared to small catchments. The authors highlight the spatial heterogeneity in sources of FIOs as a significant challenge for modelling FIO flux in smaller catchments as the relative contribution of landscape sources compared with stream bed sediment sources is greater in smaller catchments. In larger catchments the stream bed sediment sink of FIOs is likely most important with successive settling and remobilisation of FIOs over multiple storm events influencing FIO flux (Wilkinson *et al.* 2011). It is important to note that the relative contribution of landscape and sediment stores will vary not just between catchment typologies but also through time. It is likely that over successive rainfall events the landscape burden of FIOs is depleted (Evanson and Ambrose 2006) and sediment stores are replenished shifting importance of landscape stores to sediment stores as illustrated in figure 2.2. Modelling this shifting of importance from one store to the other may prove to be useful in predicting watercourse FIO flux and will require combining knowledge of FIO persistence, mobilisation, depletion, and replenishment in different environmental matrices. Further understanding of the mobilisation of FIOs from faecal deposits into soil and sediment sinks and remobilisation to watercourses will therefore be important in understanding FIO flux and the relative importance of different stores under different flow regimes (Droppo *et al.* 2011; Muirhead *et al.* 2009).

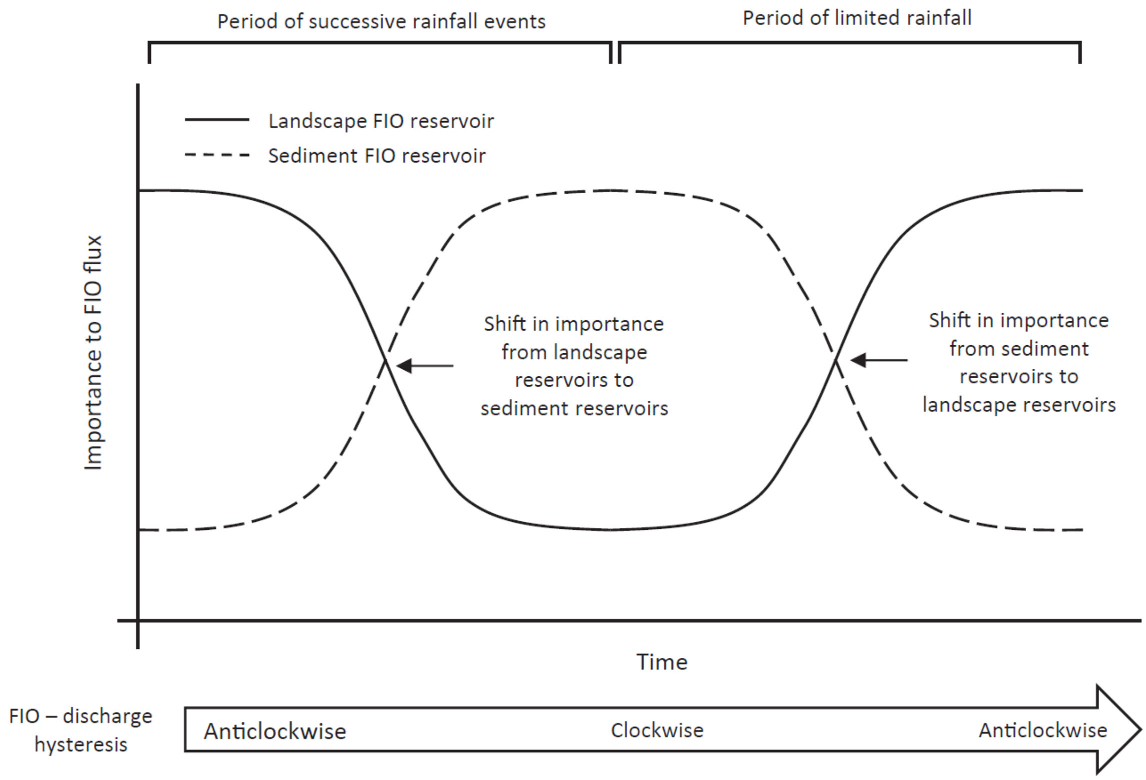

Figure 2.2. Variable relative importance of landscape and sediment sources to FIO flux under differing meteorological conditions. Under successive rainfall events the landscape store of FIO becomes depleted and stream-bed sediment stores are replenished. Under dry weather the sediment store of FIOs declines due to die-off and deposits from animals replenish the landscape store.

The rate at which FIOs are mobilised from different matrices is likely to influence the emergence of FIOs in watercourses following rainfall. Hodgson *et al.* (2009)

investigated release kinetics of FIOs (*E. coli* and enterococci) from different faecal matrices under simulated rainfall in a laboratory environment and discovered FIO release became less likely as faecal material aged and percentage dry matter increased. The study also recorded differences in FIO release related to the physical structure of different types of faeces with sheep faeces remaining intact after simulated rainfall and dispersal of cattle faeces occurring much more readily. While laboratory studies are a useful first approximation, they must be combined with field relevant studies to ensure their applicability in the field. The relative risk of different kinds of livestock faeces will be a function of initial concentration of FIOs, FIO persistence and likelihood of mobilisation from the faecal matrix. New insight into the dynamics of source risk in space and time could emerge through a combination of new and existing knowledge relating to these parameters. This will require development of more complex die-off kinetics and information on release kinetics of FIOs from faecal matrices, which will be important for understanding the risk of disparate agricultural practices such as the grazing of different types of livestock on pasture or spreading of slurry and manure. Blaustein *et al.* (2015) investigated the impact of the rate of rainfall on the release of FIOs from solid manure and observed a two-stage response with initial rapid FIO release followed by slower release suggesting easily removed FIO are mobilised quickly with further bacteria becoming harder to remove from the faecal matrix. However, whether this is due to differing morphology between individual organisms, an FIO's association with colloids or a bacteria's position in a faecal matrix is unknown. This may mean that certain proportions of a faecal deposit's burden of FIOs are more or less prone to mobilisation and understanding the size of the easily mobilised proportion and how this proportion differs between the faeces of different animals is likely to aid our understanding of FIO risk to watercourses.

The rate at which FIO mobilisation occurs is likely to depend on the faecal matrix investigated. For example, Guber *et al.* (2013) investigated FIO release from slurry and observed slower release compared to direct faecal deposits (Blaustein *et al.* 2015a). It is suggested that this difference is a result of the more complex physical structure of solid manure that contains not just animal faeces but also plant material (Blaustein *et al.* 2015a). FIO release kinetics are also likely to vary through time with potential influences from the formation of faecal crusts, wetting/drying and freeze/thaw cycles. However, there is a lack of data in this regard and it cannot be accounted for in assessments of FIO mobilisation and transfer. Field relevant release kinetics for different types of faecal matrix commonly associated with livestock manure management should be developed and their incorporation into process based models may improve prediction of in-stream FIO flux. Despite knowledge of the survival of FIOs in a soil reservoir there is little knowledge regarding the remobilisation of FIOs from this store once dissipated from the faecal source within which the FIO load was delivered to pasture. Work exists regarding the vertical leaching of FIOs through soil (Aislabie *et al.* 2011) but there is little knowledge regarding the mobilisation of FIOs from soil stores into overland flow and it is this hydrological pathway that is likely to affect FIO flux in watercourses following rainfall (Collins *et al.* 2005; Murphy *et al.* 2015).

Artificial high flow events have shown the stream bed sediment as a significant reservoir of FIOs (Cho *et al.* 2010; Muirhead *et al.* 2004) and Wilkinson *et al.*

(2011) describes a ‘shunting’ mechanism whereby FIOs are transferred further downstream over successive rainfall events moving from one sediment store to the next. This concept not only applies to rivers but also to the wider landscape with FIOs moving from faecal matrices into a soil sink then through successive soil and stream-bed sediment sinks depending on the size of the catchment. Understanding this process will require knowledge of FIO release kinetics from soil and sediment stores. A number of techniques exist for calculating re-suspension of stream-bed sediment FIO stores. For example Droppo *et al.* (2011) use average associated bacteria and critical bed shear stress as parameters in a sediment resuspension model to predict how bacteria are resuspended over high flow events; reach specific variation in sediment resuspension has been reported (Piorkowski *et al.* 2014) and Cho *et al.* (2010) attempt to capture reach specific sediment properties (Piorkowski *et al.* 2014) affecting shear stress and entrainment coefficients; and Pandey *et al.* (2012; 2013) developed and tested an empirical equation that represents FIO resuspension from stream bed sediments. Pandey *et al.* (2013) compared model predictions of watercourse *E. coli* concentrations between simulations including and excluding effects of a sediment reservoir and observed an increase from 10^7 to 10^{14} CFU/s highlighting the importance of including impacts from a sediment store in models predicting watercourse FIO pollution.

Evidence suggests that many of the FIOs transferred in run-off exist as highly mobile, free-living organisms that are less likely to settle into a soil or stream-bed sediment store (Muirhead *et al.* 2005). The extent to which a catchment ‘shunting’ mechanism occurs is likely to depend on the proportion of FIOs that exist as flocs or adsorbed to particles of soil or organic matter. Modelling approaches often assume FIOs associate with sediment particles driving the settling of FIOs into streambed sediment. However, this assumption may not be justified (Kay *et al.* 2007b). Deriving proportions of attached and free-living organisms has been attempted via a number of techniques: centrifugation, settling and filtration. Muirhead *et al.* (2005) investigated the transport state of *E. coli* from cowpats using centrifugation and found that the majority of *E. coli* was transported as free organisms with only 2-26% of the *E. coli* attached to soil particles. Other centrifugation experiments have focussed on additional FIOs including faecal coliform, *E. coli*, and enterococci with attachment rates of 15 to 30% (Cizek *et al.* 2008). This suggests that once mobilised *E. coli* are highly mobile and when mobilised can be efficiently transported to watercourses and through a catchment systems. In contrast, other experiments record higher levels of FIO attachment to soil particles. For example, using field derived soils packed into soil boxes Soupir *et al.* (2010) recorded 28 to 49% attachment to particles present in run-off and Characklis *et al.* (2005) report 30 to 55% attachment to particles in urban river samples which suggests settling of attached FIOs and retention in the landscape is more likely. A study utilising the settling method whereby particles are allowed to settle by gravity for a period of time as determined by Stokes Law which describes the rate a particle descends depending on its density reports a rate of FIO attachment to soil particles of 15.5 to 41% (Liu *et al.* 2011). Filtration utilises a method based on the ability of membrane pores relative to the size of particles to separate attached FIOs from free living cells. Using the filtration method Krometis *et al.* (2009) reported 78% *E. coli* attachment after one hour of contact time. Soupir *et al.* (2008b) investigated FIO attachment to different particle sizes using a series of different sized filters and records *E. coli* attachment of 9.5% with most attachment associated with particles 3μm and 8μm in size. Literature on the attachment of FIOs to soil particles and manure colloids demonstrates the variability of this critical process in the transfer of FIOs through landscapes. Despite this research their remains inadequate information about the factors that drive the variability in FIO attachment to soil particles and manure colloids to make predictions, at the landscape scale, of the proportion of the reservoir of FIOs that may be associated with particles and thus potentially more settleable. This information is important as it will provide insight into the relative importance of landscape and sediment stores of FIOs under different conditions.

The lack of a standard method for separating attached and unattached FIOs as well as difficulties associated with available techniques may contribute to uncertainty in FIO attachment to soil particles and manure colloids (Liang *et al.* 2014). Liang *et al.* (2014) proposed a novel method utilising flow cytometry which showed promise as a simple way of measuring the proportion of attached FIOs in samples of river and run off water and sediment. However, most flow cytometers can only accept particles which are <100μm wide, which limits investigation to clay soils. Large particle flow cytometry is a growing field of study and flow cytometers that can accept particles up to 2000μm do exist. The availability of species-specific fluorescent tags is another limiting factor. The technology has been applied only to laboratory inoculated samples due to the high microbial diversity of environmental samples. More specific staining techniques such as fluorescent *in-situ* hybridization (FISH) and fluorescent labelled anti-bodies (FITC) may improve the method for field derived samples (Liang *et al.* 2014).

Another method useful for particle separation is field flow fractionation (FFF), which exploits the tendency for particles in laminar flow to move toward the edge of a channel. A transverse field, for example a thermal gradient, that can be modulated forces particles of different sizes depending on its strength to remain at a stable position at the edge of a channel (Williams and Caldwell 2014). Baalousha *et al.* (2011) review the use of FFF in the characterisation of natural colloids and manufactured nanoparticles present in environmental systems. However, to this author's knowledge the technique has not been applied to the field of catchment microbial dynamics. FFF may prove useful in the separation of free-living bacteria and sediments in environmental samples.

Determining what drives FIO flux (Figure 2.1d) is a key priority for policy makers (Kay *et al.* 2007b). It has been highlighted that catchment characteristics such as area and patterns of precipitation influence FIO flux, suggesting that further understanding of the interaction between catchment characteristics and FIO emergence in rivers may allow development of tools for predicting the FIO contamination of watercourses. Mobilisation of FIOs from faeces has been investigated but the development of reliable and transferable FIO release kinetics for faecal, manure, slurry, soil and sediment matrices remains a challenge. Attachment of FIOs to particles of soil and organic matter is likely to be a significant factor influencing FIO transport. Realising the proportion of attached and flocced FIOs at the catchment scale is an obvious challenge and this may be in part due to difficulties associated with laboratory procedures that delineate attached, flocced and free-living bacteria from environmental samples.

2.6 Landscape drivers of FIO risk

Determining sources of FIO pollution is an important step in the management of diffuse FIO pollution (Kay *et al.* 2007b). Managing diffuse sources of FIO pollution is especially challenging because sources vary spatially and temporally in both loading and connectivity and hence not every part of the landscape contributes to pollution to the same extent (Heathwaite *et al.* 2005).

To describe the spatial heterogeneity of FIO sources a number of studies have employed a regression approach, relating watercourse concentrations of FIOs with land use information and land uses associated with management of livestock and their faeces have been associated with higher watercourse concentrations of FIOs (Kay *et al.* 2010, Tetzlaff *et al.* 2012, McGrane *et al.* 2014). A limitation associated with regression approaches is that they do not consider the physical processes driving FIO variability in watercourses and do not account for the spatial heterogeneity of hydrological connectivity among catchments (Tetzlaff *et al.* 2012). Therefore, regression approaches can only highlight sources of FIO at a scale coarser than the scale at which mitigation measures operate, creating difficulties for implementing this technique for the spatial targeting of mitigation.

An alternative approach to deriving the origin of FIO pollution is microbial source tracking (MST). MST is an emerging technology that uses quantitative polymerase chain reaction (qPCR) to investigate the genome of FIOs enumerated from the environment to determine the species from which the FIO originated. MST, in determining whether faecal pollution is derived from human, livestock or wildlife sources, would be useful in source apportionment and targeting mitigation. Harwood *et al.* (2013) provides a recent review of the technology and highlights a number of challenges including: the techniques sensitivity to dilute samples; the effect of PCR inhibiting substances in environmental samples; and the efficiency of DNA recovery. Within the microbial source tracking domain, much attention has been given to human faecal sources. However, the development of MST markers for animals is less established, especially in regard to wild animals, providing challenges for the use of this technology in apportioning diffuse sources of FIOs (Harwood *et al.* 2013).

The SMDI concept shows how a source of pollution is only a risk if it can be transferred and delivered to a watercourse (Haygarth *et al.* 2005). The extent to which mobilisation and delivery occurs will vary both spatially and temporally. Heathwaite *et al.* (2005) described the critical source area (CSA) concept whereby a part of the landscape is only a risk to in-stream microbial water quality when a source of pollution exists and that part of the landscape connects to a watercourse. The importance of connectivity to FIO risk in streams is highlighted by Murphy *et al.* (2015) who suggest FIO emergence in streams is limited by hydrological response rather than the magnitude of sources in the contributing catchment. The representation of connectivity in predictive models is crucial but Kay *et al.* (2007b) suggested that current approaches that lump connectivity into hydrological response units are too simplistic. An improved characterisation of connectivity in the landscape would be useful for predictive modelling of FIO contamination of watercourses.

Attempts to apportion FIO pollution to characteristics of the landscape are important in determining where management should be focussed and recent experiments employing regression approaches highlight land uses that may be responsible for a large proportion of in-stream contamination. In addition, experiments exploring MST highlight organisms that are responsible for pollution. However, these approaches do not consider landscape connectivity hindering their use for spatial targeting of diffuse FIO pollution management. Understanding how sources of FIO pollution in the landscape connect to watercourses and sensitive receptors remains a research priority.

2.7 Managing diffuse FIO pollution

The modelling of FIO transport through catchments represents an effort to bring understanding of how FIOs interact with different environmental matrices together within a single framework. These models can then be used to predict the loading of FIO to a sensitive receptor, such as a bathing beach or shellfish harvesting area, hence allowing policy makers to make informed decisions about the management of water quality in a catchment of interest. The development of a process based generic model of FIO transport through the landscape is challenging as these approaches require a significant amount of data for model development, calibration, predictive uncertainty assessment and validation (Beven 2014). Rather than creating one generic model, this challenge may be addressed through the development of a modelling toolbox consisting of a number of models focussed on specific policy/management questions and within certain environments reflecting the hydrological pathways and typical land management.

A number of key catchment management questions exist that the catchment management toolbox would need to address, for example:

1. what is the concentration of FIOs that will be delivered to recreational waters, shellfish harvesting areas or a drinking water reservoir in response to forecasted rainfall?
2. what and where is the source of FIOs within the landscape?
3. and how can water quality be improved to meet required standards?

The first question is perhaps best addressed through process-based mechanistic models that aim to predict the emergence of FIOs as a result of predicted or observed rainfall. Existing process based models considering FIO transport include COLI (Walker *et al.*, 1990), Hydrologic Simulation Program FORTRAN (HSPF) (Bicknell *et al.* 2001), Spatially Explicit Delivery Model (SEDMOD) (Fraser, 1999), Water Assessment Model with ArcView interface (WAMView) (Bottcher *et al.* 2002), Loading Simulation Program in C++ (LSPC) (Shen *et al.*, 2005), Soil and Water Assessment Tool (SWAT) (Neitsch *et al.* 2005; Bougeard *et al.* 2011), WATFLOOD (Dorner *et al.* 2006), KINEROS/STWIR (Guber *et al.* 2011), an approach based on the MIKE modelling suite (Bedri *et al.* 2014) and others developed by Tian *et al.* (2002), Moore *et al.* (1989), and Fraser *et al.* (1998). While a detailed review of these models is beyond the scope of the current work, de Brauwere *et al.* (2014) provides an in-depth description of process-based models. The processes governing the occurrence survival and transport of FIOs will be numerous and complex, and a lack of understanding and data relating to FIO persistence and transfer (as discussed under theme 1) provides a challenge for process-based modelling. It is possible that not all of the complexity involved in the processes of FIO behaviour in the environment can be captured in these models because understanding or data does not exist for that part of the process or the process cannot be perceived or measured with existing technology (Beven *et al.* 2006). This complexity also requires many parameters, meaning upon model calibration many optimum parameter sets are possible and multiple parameter sets may lead to identical, equifinal, outputs (Beven, 2014). The problem of data availability is compounded in catchment microbial dynamics due to the relative scarcity of data compared with similar diffuse pollution pressures such as nutrient pollution (Kay 2008; Muirhead 2015). There are significant financial costs associated with setting up FIO monitoring campaigns that gather data at the temporal and spatial scales needed to calibrate complex process-based models. Goss and Richards (2008) argue that risk index approaches may be useful interim tools while data and knowledge gaps challenge process-based models.

Questions 2 and 3 are interlinked with water quality improvement focussing on addressing important sources. These questions may best be addressed through use of a phased approach with the first phase utilising a regression-based screening tool (e.g. Kay *et al.* 2010, Tetzlaff *et al.*2012, McGrane *et al.* 2014) predicting whether catchment inputs or point sources are most important in the impairment of a waterbody of interest. The second phase would then involve a more distributed modelling approach that works at the catchment scale to determine where in the landscape pollution may be coming from and target mitigation efforts.

Mitigation of diffuse pollution can be split into two broad categories, changes in management and physical infrastructure such as fencing (Kay *et al.* 2018) and riparian buffer strips. Changes in management can be achieved through financial incentives and awareness raising through stakeholder engagement activities. For example, in the EU the Common Agriculture Policy provides financial support to the farming community based on the concept of cross compliance whereby payments depend on compliance with environmental, conservation and animal welfare legislation (Kay *et al.* 2007b). Physical infrastructure proposed for mitigation of diffuse FIO pollution includes: wetlands (Rea *et al.* 2015; Morató, 2014), retention ponds (Jenkins *et al.* 2015, 2012), streamside fencing (Kay *et al.* 2007a) and buffer strips (Tate *et al.* 2006). These physical infrastructure options are expensive and occupy valuable productive land. Therefore, it is important that their performance under different conditions is understood and spatial targeting exercises are carried out in order to determine where in the catchment these options will make the biggest difference.

A number of opportunities exist for the spatial targeting of diffuse pollution management. For example, regression-based approaches include those described by Kay *et al.* (2010), Tetzlaff *et al.* (2012) and McGrane *et al.* (2014) and can be developed into screening tools which use statistical relationships to differentiate the importance of different sources of FIO, whether they originate from point sources or diffuse catchment inputs. However, Wilkinson (2011) highlights the non-linear and non-stationary response of FIO emergence during storm events suggesting statistical relationships cannot represent the complex process of in-stream FIO flux and in a recent review, de Brauwere *et al.* (2014) suggests that statistical relationships cannot be used in place of more complex mechanistic models. However, it should be noted that the appropriateness of the modelling approach utilised depends on the scientific or policy question asked (Oliver *et al.* 2016b) and while a regression approach's ability to predict FIO load due to a rainfall event is limited it may be a powerful tool for prioritising areas for mitigation.

A challenge associated with this approach is that generalisation of statistical relationships developed at one site is likely to be inappropriate and a large amount of data will be required to derive relationships for specific sites. Therefore, understanding the minimum amount of data required to capture the variation in FIO concentrations is a key question in determining how feasible these approaches are at a national scale.

Where catchment inputs of FIOs are highlighted as a potential significant contributor to water quality impairment then prioritisation of the most significant sources of diffuse pollution at the catchment scale will be necessary. Given the importance of hydrological pathways in facilitating the transfer of FIOs to watercourses this prioritisation will require a good representation of hydrological connectivity from source to receptor and Kay *et al.* (2007b) highlighted the flow connectivity simulation approach carried out by Heathwaite *et al.* (2005) for diffuse phosphorus pollution as a potentially powerful approach for the prediction of diffuse FIO pollution. FIORIT (Oliver *et al.* 2010) is a risk indexing tool whose goal was to create an operational risk assessment framework given the importance of connectivity and lack of understanding for some of the processes involving FIO transport to watercourses. Rules for the allocation of risk between fields were developed by panels of experts with flexibility for location specific influences a priority. It is a useful tool for the prioritisation of mitigation effort as it highlights fields that are relatively more risky to river water quality. A benefit of FIORIT is that new knowledge can be integrated into the framework with ease (Oliver *et al.* 2010).

Another tool utilising the CSA concept is the SCIMAP risk mapping framework. SCIMAP uses spatial land cover data to determine the potential locations of sources and detailed topographic information to calculate how likely it is for a source to become connected to the river network (Reaney *et al.* 2011). The process is risk-based and does not attempt to determine the absolute concentration of pollution in a stream but outputs a risk quotient between 0 and 1, which represents the risk of pollution transfer at that point relative to all the other points in a river network. Therefore, the approach aims to identify where is most and least at risk within the catchment. Where routine monitoring has highlighted a catchment of concern, SCIMAP can be applied to predict where in the catchment mitigation efforts should be concentrated. A benefit of the SCIMAP framework is that it is built around a minimum information requirement philosophy, requires a minimum of three inputs (elevation, land use and spatial rainfall information) and can be run using relatively inexpensive computer hardware. The software is also available for free as a web app (https://my.scimap.org.uk).

The approach has been optimized for conservative diffuse nutrient and sediment pollution (Reaney *et al.* 2011; Milledge *et al.* 2012) and has not been developed for FIOs. Given SCIMAP's use by the regulatory community in England it is timely to develop this framework to account for diffuse FIO pollution. At present SCIMAP is time-integrated, i.e. it shows a long term average of risk in the landscape. FIOs are a non-conservative contaminant; their concentrations vary over time due to population growth and die-off and capturing the temporal variations in FIO risk is a key challenge for the application of SCIMAP to diffuse FIO pollution. Given these limitations, SCIMAP use for diffuse FIO pollution may be problematic; however, there are opportunities for the development of SCIMAP. Abiotic conditions that vary from month to month have a significant influence on the persistence of FIOs in the landscape. Therefore, a time-integrated approach is inappropriate as it is important to understand the month-to-month variation of risk across the landscape. Understanding the effects of various factors that influence growth and die-off of FIOs in the landscape will allow for a weighting of risk according to month given a set of conditions expected of that month. Furthermore, survival in overland flow and transport times needs to be derived in order to discover what proportion of a population of FIOs die or is withheld in the landscape before reaching a watercourse.

The SCIMAP output would represent the risk of FIO contamination occurring rather than the risk of illness due to FIO prevalence in stream water. Quantitative Microbial Risk Assessment (QMRA) aims to investigate public health risk associated with contamination of water (McBride *et al.* 2013). Muirhead *et al.* (2015) use QMRA to develop a risk index creating an output suitable for interpretation by land managers, for example farmers, and water quality managers, such as regulatory authorities and charities. The approach relates farm management practices with resulting instream concentrations of FIOs identifying opportunities for mitigating infrastructure and changes in management that may reduce instream concentrations of FIO.

Many approaches targeted toward modelling FIO transport focus on catchment outlets (Kay 2009). A benefit of both the SCIMAP and risk-index approaches outlined above is that they address the potential for FIO contamination within the stream network. This is important for identifying important catchment inputs and understanding potential harm of within-catchment network receptors such as water supplies and those partaking in water contact recreation such as sport fishing and kayaking.

A tiered approach which moves from coarse to fine scales is likely to provide a resource efficient approach that the policy community can easily implement. Therefore, modelling should be seen as a toolbox for catchment managers with a suite of tools able to provide outputs for a variety of questions.

Increasingly integrated catchment management is moving away from a traditional top-down approach where environmental issues are managed at a relatively high level of governance to a more participatory approach with more involvement from the community and stakeholders within catchments (McGonigle, 2013). It is therefore more important than ever that model outputs are easily interpreted by a variety of end users and model developers should pay close attention to human-computer interaction during model development (Yearley 1999; Whatmore *et al.* 2011). Communicating uncertainty is a significant challenge in this respect and likely to influence stakeholder engagement with academic outputs (Spiegelhalter *et al.* 2011; Retzbach *et al.* 2015).

When considering mitigation options it must be noted that integrated catchment management requires catchment managers to consider the risk of multiple pollutants to multiple receptors. Phosphorus (Schoumans *et al.* 2014), nitrogen (Heppell *et al.* 2014), pesticides (Bloodworth *et al.* 2015), suspended sediment (Rickson, 2014) and FIOs are diffuse pollutants that have the potential to cause harm to the environment and ecosystem services. A priority for research should be the development of modelling tools that determine what mitigation measures should be and where these measures should go in order to represent an efficient use of resources, not just within types of diffuse pollution but across all types of pollution. Such an integrated assessment of mitigation options should also investigate the potential for pollution swapping whereby mitigation of one pollutant leads to an increase in another pollutant (Stevens and Quinton, 2009). A conceptual multi-pollutant framework to identify opportunities for multiple benefits and potential pollutant swapping has recently been proposed (Bloodworth et al., 2015). Concepts from this framework can be built into multi pollutant frameworks, such as SCIMAP, to develop prioritisation tools where there are pressures from multiple contaminants.

Limitations in the understanding of FIO fate and transfer make it difficult to predict what mitigation options will be most effective. Determining the efficacy of different mitigating infrastructures is a research priority. In addition, tools predicting where mitigation efforts should be targeted are required in order to justify occupation of productive land and the cost of mitigation measures. The management of diffuse FIO pollution must be taken in context with other potential pollutants such as sediment and nutrients and opportunities to target multiple pollutants with the same infrastructures must be explored.

2.8 Conclusion

Catchment microbial dynamics is a rapidly evolving field and in recent years there have been a number of advances in the understanding of FIO behaviour in a range of environmental matrices. For example, progress has been made towards more complex models of FIO persistence in the landscape but research understanding the growth of FIOs immediately following deposition under different conditions is required. An emerging research priority is understanding the potential contribution wildlife faeces makes to watercourse FIO pollution and how this may vary spatially and temporally. Kay *et al.* (2007b) highlighted a need for further research into FIO population dynamics in the sediments associated with freshwater systems. FIO persistence in sediments has been investigated in the laboratory and now field relevant studies are needed to assess applicability of these results in the field. Nutrient status and soil type have been implicated in the variable persistence of FIOs in soil and determining the characteristics of soil that drives this influence is a priority. Understanding the episodic nature of FIO emergence in watercourses had been highlighted as a topic of priority (Kay *et al.* 2007b) and challenges in this respect remain. Mobilisation of FIOs from different reservoirs is likely to influence FIO emergence in streams following rainfall and further study is required in order to develop a database of FIO release kinetics from faecal, manure, slurry, soil and sediment matrices. Understanding the proportion of FIOs that are free living, flocced or attached to particles of soil or organic matter at the catchment scale remains a significant challenge for understanding the episodic flux of FIOs in watercourses; flow cytometry and FFF may provide solutions to this problem. Kay *et al.* (2007b) highlighted understanding the connectivity of landscape sources of FIO to watercourses as a research priority. While pathways for FIO transport such as overland flow and field drainage have been investigated understanding FIO connectivity at the catchment scale remains a challenge. Modelling frameworks such as SCIMAP, FIORIT and QMRA approaches provide opportunities in this respect. Addressing the challenges outlined in this review may allow for better management of diffuse FIO pollution of watercourses and the impact upon ecosystem services such as clean and safe bathing and shellfish harvesting water can be reduced.

3. Predicting diffuse microbial pollution risk across catchments: the performance of SCIMAP and recommendations for future development

This chapter was published in Science of the Total Environment:

Porter, K.D., Reaney, S.M., Quilliam, R.S., Burgess, C. and 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**, pp.456-465.

3.1 Abstract

Microbial pollution of surface waters in agricultural catchments can be a consequence of poor farm management practices, such as excessive stocking of livestock on vulnerable land or inappropriate handling of manures and slurries. Catchment interventions such as fencing of watercourses, streamside buffer strips and constructed wetlands have the potential to reduce faecal pollution of watercourses. However, these interventions are expensive and occupy valuable productive land. There is, therefore, a requirement for tools to assist in the spatial targeting of such interventions to areas where they will have the biggest impact on water quality improvements whist occupying the minimal amount of productive land. SCIMAP is a risk-based model that has been developed for this purpose but with a focus on diffuse sediment and nutrient pollution. In this study we investigated the performance of SCIMAP in predicting microbial pollution of watercourses and assessed modelled outputs of *E. coli*, a common faecal indicator organism (FIO), against observed water quality information. SCIMAP was applied to two river catchments in the UK. SCIMAP uses land cover risk weightings, which are routed through the landscape based on hydrological

connectivity to generate catchment scale maps of relative in-stream pollution risk. Assessment of the model's performance and derivation of optimum land cover risk weightings was achieved using a Monte-Carlo sampling approach. Performance of the SCIMAP framework for informing on FIO risk was variable with better performance in the Yealm catchment (r_s = 0.88; p <0.01) than the Wyre (r_s = -0.36; p > 0.05). Across both catchments much uncertainty was associated with the application of optimum risk weightings attributed to different land use classes. Overall, SCIMAP showed potential as a useful tool in the spatial targeting of FIO diffuse pollution management strategies; however, improvements are required to transition the existing SCIMAP framework to a robust FIO risk-mapping tool.

3.2. Introduction

Faecal pollution has the potential to negatively impact upon ecosystem services associated with clean and safe recreational bathing and shellfish harvesting water (Clements *et al.* 2015; Wu *et al.* 2016). Microbial contamination of such aquatic environments can expose humans to harmful pathogens that may cause gastrointestinal illness (Wade *et al.* 2006). Direct measurement of pathogens in environmental water samples is uncommon due to challenges associated with their enumeration in the laboratory, e.g. cost, detection limits etc., and so faecal indicator organisms (FIOs) such as *Escherichia coli* and intestinal enterococci provide an internationally accepted framework for the assessment of faecal pollution of water bodies. In the European Union, the health risks of faecal pollution of aquatic environments are recognised via the Bathing Water (EU, 2006a) and Shellfish Water (EU, 2006b) Directives. Regulators must compare measured FIOs against stringent standards of microbial water quality in order to comply with these directives. Risk assessment tools that can identify 'hotspots' of FIO pollution in catchment systems are therefore welcomed by regulatory agencies as a mechanism to help understand origins of pollution and to spatially target catchment management and interventions for improvements in microbiological water quality (Dymond et al., 2016).

Diffuse sources of FIO pollution, such as organic fertilisers applied to land and excretion of faeces by grazing livestock to pasture, provide challenges to water quality managers. This is because the loading of diffuse sources, and their propensity to connect to watercourses, varies spatially and temporally (Heathwaite *et al.* 2005). The impact of diffuse sources of microbial pollution on watercourses can be reduced through the use of mitigation measures such as streamside fencing (Kay *et al.* 2007a), vegetated buffer strips (Tate *et al.* 2006), wetlands (Morató, 2014) and retention ponds (Jenkins *et al.* 2015). These measures can be expensive and occupy valuable productive land. Therefore, methods to spatially identify and target locations in catchments where interventions will provide the best improvement in water quality are warranted. Past research has used regression approaches to attribute sources of FIOs to different land cover types and/or discrete point sources (Kay *et al.* 2010; Tetzlaff *et al.* 2012; McGrane *et al.* 2014). However, these approaches do not account for the spatial heterogeneity of landscape to watercourse connectivity (Tetzlaff *et al.* 2012).

Alternative approaches include the development of fully process-based models that attempt to account for the mechanisms that govern FIO fate and transfer in more detail. There are, however, limitations in our understanding of FIO fate and transfer that can amplify uncertainties in fully quantitative, process-based risk assessment approaches. For example, there are knowledge gaps regarding the complex behaviour of FIO persistence in different matrices such as faecal deposits (Soupir *et al.* 2008a; Martinez *et al.* 2013; Oliver & Page, 2016), soil (Muirhead *et al.* 2009; Park *et al.* 2016) and stream bed sediment (Pachepsky and Shelton 2011; Shelton et al., 2014; Pandey *et al.* 2013; Pandey *et al.*, 2016). In addition, current understanding of the mechanisms with which FIOs aggregate and attach to particles is limited (Muirhead *et al.* 2005; Liu *et al.* 2011). This understanding is especially important as it is likely to determine the extent to which FIOs settle within the streambed environment or the extent to which cells might be retained in the landscape through processes of filtration by the soil architecture (Engstrom et al, 2015). Such limits in understanding make it difficult for all processes to be considered in complex process-based models (Beven 2006, Cho et al., 2016). These complex models also require a significant amount of data for model parameterisation and validation. This is especially problematic in the field of catchment microbial dynamics due to the relative scarcity of data on FIO concentrations and loads compared to nutrient and sediment flux (Muirhead 2015; Oliver et al., 2016). Semi-quantitative risk assessment frameworks, which provide a basis for decision support, are therefore useful tools to inform on relative risk of FIO transfers in space and time. This is because, despite gaps or limitations in the current evidence-base concerning FIO behaviour in complex catchment systems, they are able to provide a '1st approximation' of risk (Goss and Richards 2008).

The Sensitive Catchment Integrated Mapping Analysis Platform (SCIMAP) has demonstrated significant potential as a framework to inform on catchment-scale risks for diffuse nutrient and sediment pollution (Reaney *et al.* 2011). The approach provides an estimate of in-stream risk relative to the catchment being considered and provides information at multiple spatial scales but within a time integrated framework. SCIMAP is underpinned by the source-mobilisation-delivery-impact (SMDI) continuum (Haygarth *et al.* 2005) and critical source area (CSA) concepts, which describe how a source of pollution can only convert to a pollution risk if there are no interruptions to the SMDI continuum (Heathwaite *et al.* 2005). At present, the SCIMAP approach is optimised for diffuse fine sediment (Reaney *et al.* 2011) and nutrient pollution (Milledge *et al.* 2012) but offers scope for addressing a number of additional diffuse pollutants, including FIOs.

The aim of this study was to assess the effectiveness of the current SCIMAP framework for informing on risk of FIO pollution in contrasting catchment systems by comparing FIO pollution risk predicted by SCIMAP with observed FIO risk, e.g. FIO concentrations. To deliver on this aim the objectives were to: (i) quantify variation in model performance as a result of risk weightings being assigned to a particular land cover type; and (ii) determine whether there was an association between SCIMAP predicted FIO risk and observed FIO risk in our study catchments. The intention was to develop initial risk weightings for land cover types and benchmark model performance on the assumption that FIOs behave similarly to sediment, albeit in a ‘living’ form.

3.3. Materials and Method

Most modelling frameworks predict in-stream pollution by defining a function, e.g. a relationship derived from regression analysis, and these can be described as forward models. Our study adopted an inverse approach (Reaney et al., 2011; Milledge et al. 2012), because it defined a function (in the case of SCIMAP, land cover risk weightings) based on observed FIO concentrations, i.e. the approach queries how a model needs to be parameterised in order to simulate observed pollution and is therefore ‘fitted’ to observed data. This ‘fitted’ approach is described in detail in Milledge *et al.* (2012). Briefly, the fitted approach involves pseudo randomly generating simulations from forward models whose output is compared to observed data. In this case the forward model used is SCIMAP and the user definable parameters are risk weightings for different land cover types. Model outputs were compared against a spatial FIO water quality dataset provided by the Environment Agency. This dataset spans 6 years (2007-2012) and was collected as part of the Catchment Sensitive Farming (CSF) initiative (Environment Agency, 2016). The FIO dataset reported here concerns *E. coli* concentrations, measured using the standard method of membrane filtration, reported across two catchments in England: The River Wyre, Lancashire and The River Yealm, Devon (Figure 3.1).

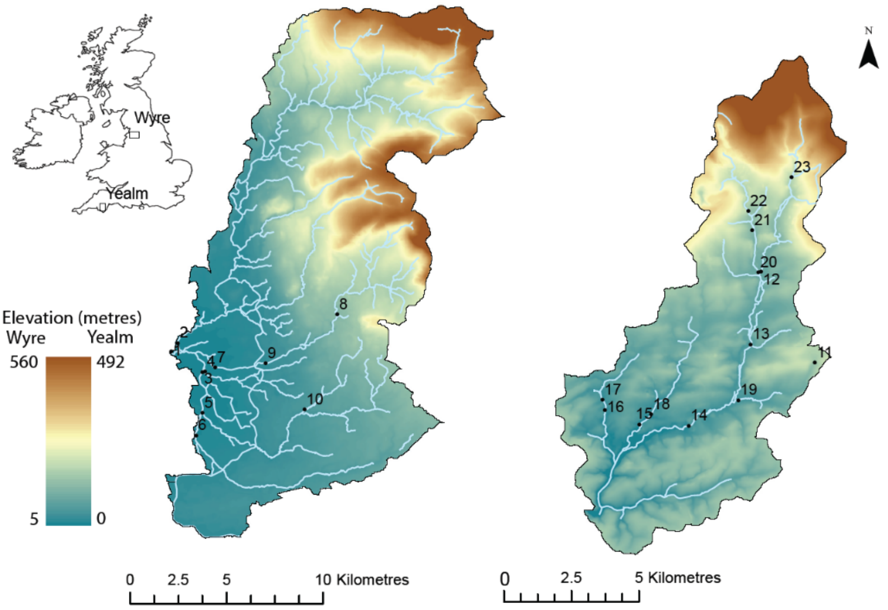

Figure 3.1. Maps illustrating the two study catchments: The Wyre in Lancashire, North West England (left); and The Yealm in Devon, South West England (right). Numbered points indicate sample locations and associate with the sample locations indicated in figure 3.2 and table 3.2. River data are from an Ordnance Survey MasterMap Topography layer.

To evaluate the SCIMAP approach when applied to a FIO dataset we used the same SCIMAP framework that was developed for prediction of diffuse fine sediment risk. This approach was implemented within the SAGA geographical information system (Conrad *et al.* 2015). The SCIMAP risk mapping approach is described in detail in Lane *et al.* (2009) and Reaney *et al.* (2011). Briefly, the approach involves determining the risk of a sediment (or other pollutant)

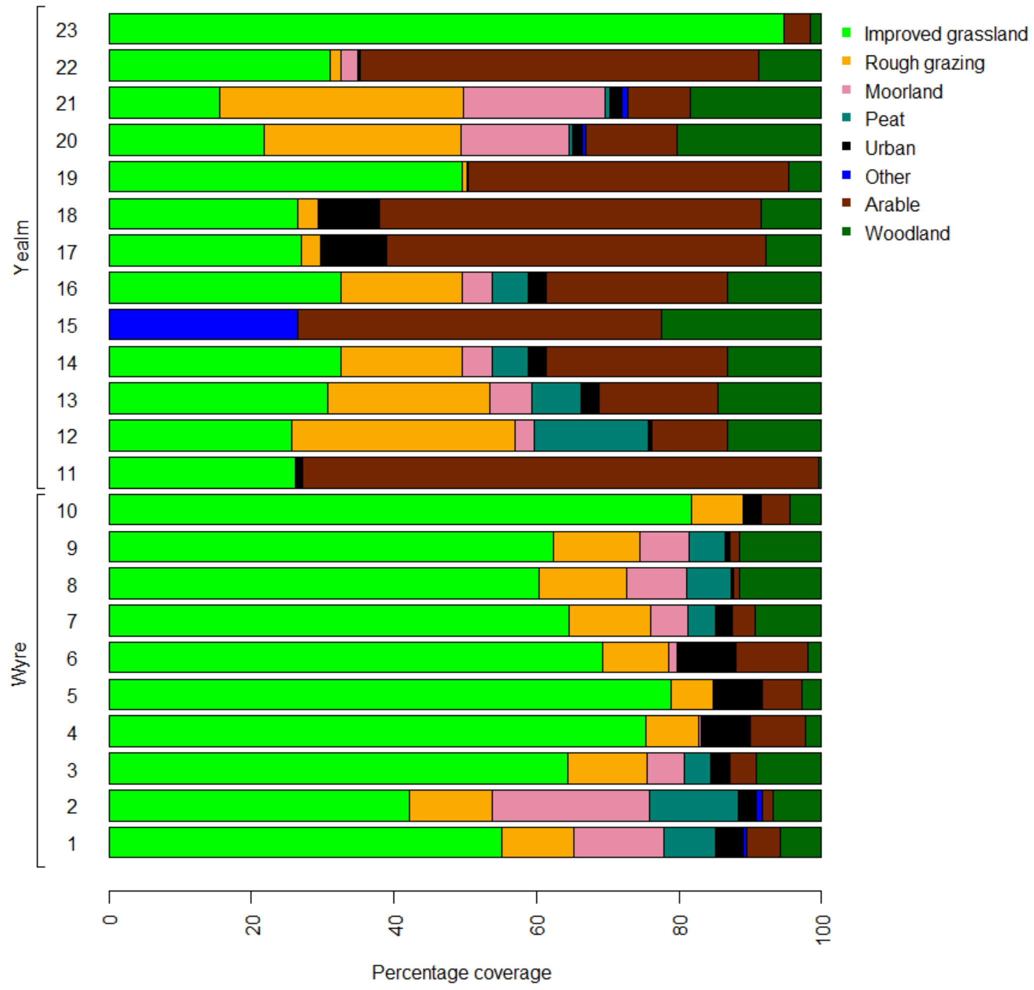

Figure 3.2. A bar plot illustrating the proportions of the contributing area associated with each sample point occupied by different land cover types. Please refer to online version for colour.

source being generated and the risk of the sediment (or other pollutant) source becoming connected to a watercourse, capturing the CSA concept described earlier. For sediment pollution, the risk of a source being generated is defined as a function of topography, land cover and rainfall. These datasets are used to calculate local erodibility based on the land cover, and the erosive potential of overland flow, which is driven by the local slope gradient and the upslope contributing area. Therefore, due to the combination of these factors, each land use is associated with its own risk weighting. The risk of the source connecting to the stream network is determined using the network index of hydrological connectivity (Lane *et al.* 2009), which can be derived from the topographic wetness index. The topographic wetness index calculates the propensity for part of the landscape to generate saturation excess overland flow from topographic information (Beven & Kirkby, 1979). The propensity for a point in the landscape to connect to a watercourse is then defined as the lowest value of topographic wetness index along the flow path to the watercourse. If overland flow is not generated at any point along a flow path, it is not possible for that cell to transmit water further downslope and hence the source of risk is disconnected from the stream network (Lane *et al.* 2009). Once a pollution source has been delivered to a watercourse, pollution risk is concentrated as it is routed downstream and diluted based on the rainfall weighted upslope contributing area, with higher risk inputs concentrating risk and lower risk inputs diluting risk.

SCIMAP adopts a minimum information requirement approach, and the standard version requires three inputs: the generation of a source of risk requires a land cover map and spatially distributed rainfall information; the derivation of a topographic wetness index requires a detailed digital elevation model (DEM); and the concentration and dilution of risk utilises the same rainfall information described previously. In this study; the land cover map utilised was the Centre for Ecology and Hydrology (CEH) Land Cover Map 2007 (Morton *et al.* 2011); rainfall information was Met Office UKCP09: 5 km gridded data - annual averages (Met

Office 2014); and the NextMap digital elevation model (DEM) at a grid resolution of 5m x 5m, developed by Intermap, was used. It is important to balance the information content of the observed dataset with the complexity of the modelling approach.

Table 3.1. A description of SCIMAP land cover classes and how they are derived

CEH LCM broad habitat class	SCIMAP class	Description
Broadleaved mixed and Yew woodland Coniferous woodland	Woodland	Deciduous, mixed, conifer, larch, evergreen and felled forest.
Arable and horticulture	Arable	Freshly ploughed land and annual and perennial crops.
Improved grassland	Improved grassland	Intensively managed grassland for hay, silage. and/or grazing of livestock
Rough grassland Neutral grassland Calcareous grassland Acid grassland	Rough grazing	Semi-natural grassland and managed low productivity grassland.
Fen marsh and swamp Bog	Bog	Herbaceous and mossy swards with a peat depth of > 0.5 metres. Fen, fen meadows, rush pasture, swamp, flushes and springs.
Dwarf shrub heath Montane habitats Inland rock	Moorland	Heather grassland and exposed rock as well as habitats occurring at higher altitudes.
Salt water Fresh water Supra-littoral rock Supra-littoral sediment Littoral rock Littoral sediment	Other	Coastal water, rivers, canals and standing water. Coastal rock and sediment.
Built -up areas and Gardens	Urban	Built up areas including towns, cities, dock sides, industrial estates and car parks. Suburban areas with a mix of built up areas and vegetation.

Therefore, for the purposes of this experiment the 23 land cover classes described in the CEH land cover map were condensed into eight classes: improved grassland, rough grazing, moorland, bog, arable, urban, woodland and other. Table 3.1 shows which of the CEH land cover classes were included in each of these new classes. The rationale for the reduction and merging of classes was that a number of separate classes within the larger CEH Land Cover map were listed where we would not expect a significant difference in the risk weights associated with the cover. For example, the deciduous and coniferous woodland classes were merged since they will have similarly low levels of livestock and represent similar availabilities of FIOs. Here, the fitted approach was used to establish how these land covers needed to be weighted in order to best represent in-stream measured *E. coli* risk. The SCIMAP fitted approach uses a Monte Carlo sampling framework based on the Generalised Likelihood Uncertainty Estimation (GLUE) methodology (Beven and Binley 1992). Here 25000 model realisations with varying land cover risk weightings were generated.

Modelled risk values for 10 locations in the River Wyre catchment and 13 locations in the River Yealm catchment were compared with associated observed measurements of *E. coli* concentration. Locations and an overview of catchment characteristics for each location are shown in figures 3.1 and 3.2 and table 3.2. Previous studies have found that >90% of FIO loading to water occurs during high flow conditions following rainfall (Kay et al. 2007b; McKergow and Davies-Colley 2010; Kay et al. 2010) and Kay *et al.* (2007b) noted that many studies employ a regular sampling regime which biases toward low flow and while this was observed in the EA dataset, sufficient data representative of high flow conditions were deemed to be present. In order to avoid a bias toward the many base flow samples present within the EA dataset for both catchments, the data that were associated with flow that was >=60% of the highest flow were subset. This operationally defined exceedance threshold retained high flow events while excluding data associated with base flow conditions. Flow data were not available for all of the locations used in this experiment so flow information from a local gauging station was used; location 9 for the Wyre and location 14 for the Yealm (figure 3.1). This approach assumed that if it was high flow at one point in the catchment it was also high flow at the other points in the catchment. While this approach represented an approximation, we argue that it remains valid given that it is being used within a risk-based framework, i.e. it is the relative magnitude of *E. coli* concentrations that is important rather than the absolute concentration. The number of records remaining at each site after this sub-setting procedure was used is shown in table 3.2. A high number of samples were associated with locations 1 and 9 in the Yealm catchment and location 14 in the Wyre catchment. These locations were equipped with autosamplers programmed to sample after a flow threshold considered to be high flow was met. Samples from other sites were acquired using manual grab sampling.

Table 3.2. Catchment characteristics of each of the sub-catchments investigated. Connectivity is defined as the lowest value of topographical wetness index along a flow path as per Lane *et al.* (2009). Number of samples indicates the number of records remaining after sub setting all the available data by the days where flow is >60% of the highest flow recorded.

	Location	Mean connectivity/ standard deviation	Mean slope(°)/ standard deviation	Mean elevation (metres)/standard deviation	Number of samples
Wyre	1	0.71/0.04	0.24/0.66	8.6/0.4	706
	2	0.63/0.13	5.41/5.63	141.33/129.28	20
	3	0.71/0.14	0.89/1.00	13.72/3.71	21
	4	0.75/0.13	0.62/0.79	12.06/3.06	21
	5	0.69/0.12	1.78/1.93	56.32/33.92	20
	6	0.68/0.13	1.75/2.28	45.35/25.38	20
	7	0.67/0.12	1.97/1.65	42.62/20.2	19
	8	0.62/0.12	7.49/7.26	223.17/95.61	21
	9	0.62/0.12	6.88/7.05	199/99.89	888
	10	0.64/0.11	3.39/3.36	104.54/42.18	23
Yealm	11	0.88/0.10	3.42/1.71	128.91/8.76	40
	12	0.83/0.12	7.34/4.66	274.93/123.84	40
	13	0.59/0.05	7.16/1.79	44.08/2.39	42
	14	0.81/0.14	6.51/4.74	125.58/85.16	193
	15	0.77/0.04	6.05/3.32	33.49/13.81	47
	16	0.45/0.07	14.1/5.06	20.98/4.57	41
	17	0.78/0.13	5.96/4.61	59.72/19.29	30
	18	0.78/0.14	5.64/4.13	55.08/18.18	50
	19	0.83/0.12	5.82/3.62	85.25/32.87	41
	20	0.79/0.11	2.45/0.82	57.8/0.56	42
	21	0.63/0.13	7.23/2.84	87.19/3.32	27
	22	0.49/0.05	20.4/7.14	122.77/10.83	28
	23	0.88/0.07	5.00/2.47	262.31/44.18	42

Statistical analysis

All statistical analysis was carried out using the R statistics package (R Core Team 2015) and third-party packages (Augie 2015; Carr et al. 2014; Sarkar & Andrews 2013; Neuwirth 2014; Wickham 2007, 2014, 2015; Wickham & Francois 2015; Deepayan 2008). All *E. coli* counts underwent $\log_{10}$ transformation prior to statistical analysis. The observed *E. coli* measurements used in our study were derived from a median of the subset data for each location and were converted into risk values by determining their rank order to allow comparison with the relative risk nature of the SCIMAP output. The Spearman's rank correlation coefficient (r_s) comparing observed risk with the simulation output was used as the objective function. This statistical comparison measures the extent to which the relative order of the locations in the observed and simulated datasets match and avoids assuming the observed dataset includes the most and least risky locations in the catchment. For each catchment this assessment provided 25000 Spearman's correlation coefficients; one associated with the comparison of each of the randomly generated combinations of land cover risk values and the observed in-stream *E. coli* risk. One sample *t*-tests were used to assess whether the land cover risk values associated with the best modelled outputs (i.e. top 1% of r_s) were significantly different from 0.5. Because *E. coli* concentrations in the Wyre catchment were not normally distributed, a Kruskal Wallis test was used to investigate differences in *E. coli* concentration among the observed data across all sites of the Wyre, with a Dunn's test used to determine which sites were different from one another. Differences at the $p < 0.05$ level (95% confidence interval) were considered statistically significant.

3.4. Results

The SCIMAP fitted approach provided three outputs that elicit information on the influence of different land covers on the risk of FIO pollution in streams and rivers. Two-dimensional density (2-Dd) plots and boxplots (figure 3.3) depict the relationship between land cover risk weighting and model performance. The 2-Dd plot is a scatter plot of risk value against the Spearman's correlation coefficient, derived from comparing the SCIMAP output associated with that risk value and observed FIO risk. The scatterplot is divided into hexagonal sections whose saturation determines the number of models that fall into that part of the plot. Results from *t*-tests (table 3.3) determine the confidence with which we can reject the null hypothesis that the mean risk weighting of the 1% best performing models is significantly different from 0.5 and therefore either contributes to diffuse pollution (risk weightings >0.5) or dilutes pollution (risk weightings <0.5). Together these results provide insight into the performance of SCIMAP's prediction of diffuse FIO pollution risk by providing the maximum correlation achieved and the potential uncertainty associated with model outputs, which is driven by the 'identifiability' of optimum risk weightings for land cover types. Identifiability, or the ease at which an optimum risk weighting can be derived, is represented by the standard deviation of risk value in the 1% best performing models. Larger standard deviations suggest that it is harder to identify an optimum risk weighting for land cover types.

3.4.1 The River Yealm catchment

The results for the Yealm catchment suggest that improved grassland and woodland should be assigned low risk values with respect to their contribution to FIO pollution of water. The 2-Dd plots (figure 3.3) show improvement in model performance as the risk weighting for these land covers decreases. In addition, the boxplots (figure 3.3) show low mean risk weightings associated with best 1% performing models. By contrast, the results suggest that rough grazing should be assigned high risk weightings with the 2-Dd plots showing improving performance of SCIMAP as the risk weighting increases. The boxplot shows a high mean risk weighting for the best 1% performing models affirming this result. The 2-Dd plot infers that arable land cover should be associated with a medium amount of risk, with model performance peaking at risk values approaching 0.54. This was supported further by both the box plots (figure 3.3) and mean risk weighting of the best 1% performing models (table 3.3). Risk weightings associated with the remaining land cover types (moorland, bog and urban) were not influential on the performance of the model predicting in-stream FIO risk. The mean risk weighting for these land covers was approaching 0.5 with a large standard deviation (table 3.3); therefore, the mean risk weighting was not significantly different from 0.5 (table 3.3). This was also apparent in the 2-Dd plots, as represented by a ‘flat top’ in the output (figure 3.3).

Of the land covers shown to have an impact on FIO diffuse pollution risk, only the risk weightings associated with improved grassland and woodland were highly identifiable. The optimum risk weighting for rough grazing was harder to identify (table 3.3). Overall the performance of SCIMAP in the prediction of FIO risk in the Yealm catchment was good with a maximum r_s of 0.88 (p<0.01).

Table 3.3. Table summarising the influence of land cover risk weighting on SCIMAP performance. Mean risk weightings and associated standard deviation for the 1% best performing models. *p* value indicates the results from a *t*-test and the confidence with which we can reject the null hypothesis that there is no variation in model performance as a result of the risk weighting assigned to a land cover type.

Land cover type	Yealm		
	Optimum mean/ standard deviation	*p* value	Summary of influence on FIO risk
Improved Grassland	0.08/0.05	<0.001	Low risk
Rough Grazing	0.78/0.16	<0.001	High risk
Moorland	0.5/0.29	>0.05	Not influential
Bog	0.5/0.29	>0.05	Not influential
Urban	0.5/0.29	>0.05	Not influential
Arable	0.54/0.23	<0.01	Medium risk
Woodland	0.19/0.05	<0.001	Low risk

Land cover type	Wyre		
	Optimum mean/ standard deviation	*p* value	Summary of influence on FIO risk
Improved Grassland	0.63/0.32	<0.001	Medium risk
Rough Grazing	0.58/0.26	<0.001	Medium risk
Moorland	0.52/0.29	>0.05	Not influential
Bog	0.49/0.30	>0.05	Not influential
Urban	0.52/0.30	>0.05	Not influential
Arable	0.18/0.22	<0.001	Low risk
Woodland	0.04/0.04	<0.001	Low risk

3.4.2 The River Wyre catchment

The performance of SCIMAP in predicting FIO risk in the Wyre catchment was poor with no correlation between predicted risk and observed risk (r_s = -0.357, p > 0.05). The 2-Dd plots in figure 3.3 provide little insight into the influence of all land cover risk weightings on model performance.

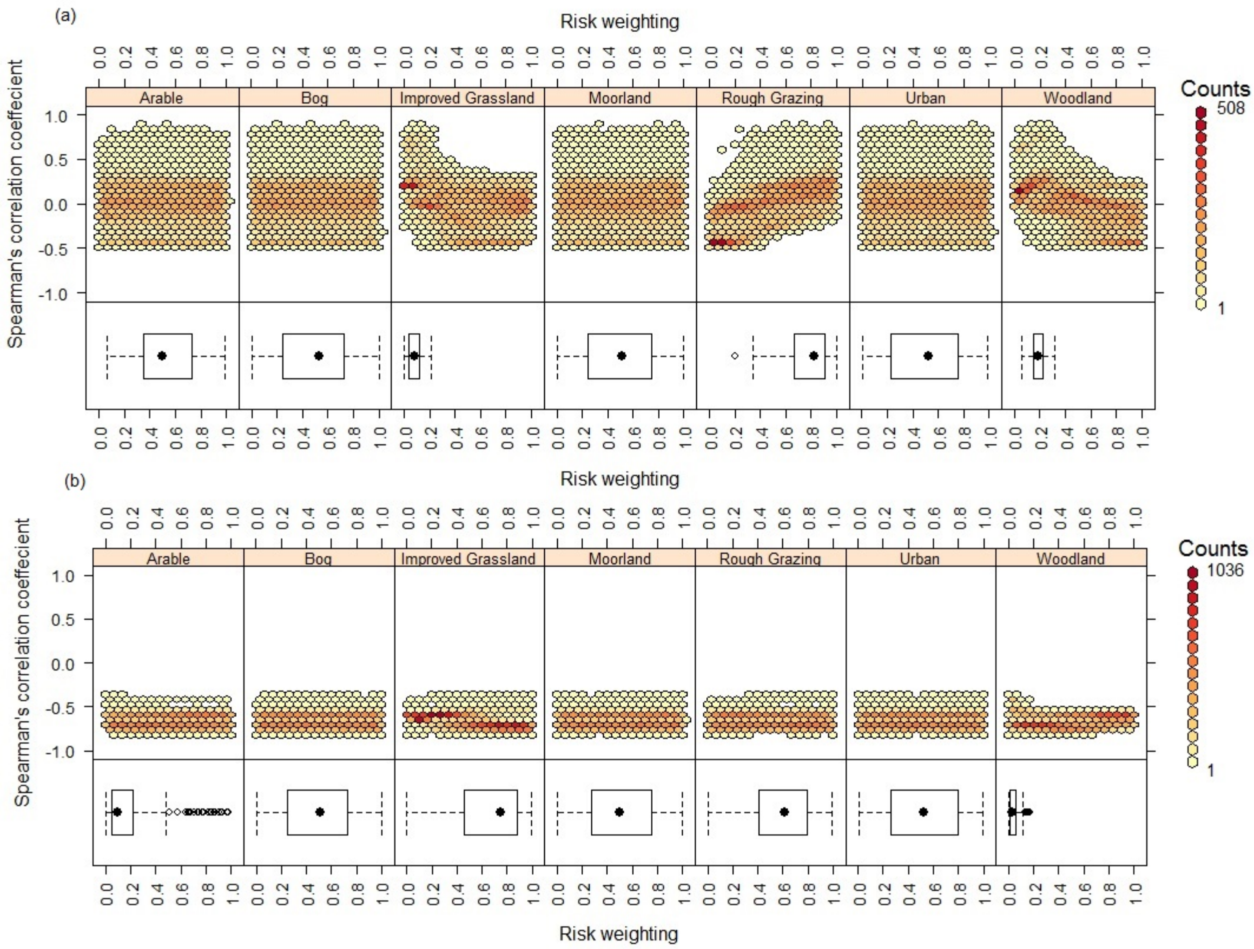

Figure 3.3. SCIMAP fitted results for (a) the Yealm and (b) the Wyre. The top panels show hexagonally binned scatterplots depicting how model performance changes with changing the risk weighting for each land cover. The colour of the hexagonal bin depicts how many simulations fall into that part of the plot. The bottom panels show boxplots which depict the variation in the risk weighting of the 1% best performing simulations.

However, when the risk weightings from the best 1% models are depicted as boxplots (figure 3.3) relationships can be seen. Arable and woodland show better model performance with lower risk weightings while model performance appears to improve when improved grassland and rough grazing is assigned a medium risk weighting. Land covers associated with moorland, bog and urban areas do not appear to influence model performance. These results are supported by the results of a *t*-test (table 3.3). The risk weighting associated with woodland is more identifiable while risk weightings associated with the remaining land covers of influence are less identifiable (table 3.3).

Ordination plots can be used to illustrate the variability in land cover between the contributing catchments associated with the sample points in the Wyre. Here non-metric multidimensional scaling (Kruskal 1964a) was used, utilising a Bray Curtis dissimilarity index (Bray and Curtis, 1957) (figure 3.4). The approach plots the coverage of each land use against the coverage of all other land uses creating a space with, in this case, 7 dimensions and then reduces the number of dimensions to 2 to allow visualisation. The extent to which the new 2-dimensional space represents the original 7 dimensional space is described with a value of stress. In this case a solution with a stress of 0.09 was achieved which Kruskal (1964b) describes as a fair representation of the original multi-dimensional space. Each point on the plot represents one sub catchment and increasing dissimilarity in land cover make up is associated with increasing distance between points. It was clear from this plot that the sub catchments associated with the Yealm river basin were more dissimilar than those associated with the Wyre river basin. There was less variability in the composition of land cover in the Wyre and sub catchments appear to gather into two clusters. This similarity between sub catchments was also apparent in the FIO concentrations observed in the Wyre. Figure 3.5 shows a boxplot illustrating the variability in FIO concentration at each of the sampling points in the Wyre catchment and a significant difference in the distributions of FIO concentrations at each of the sites was observed (P < 0.05). The Dunns test did reveal that there was a degree of clustering of sites. For example, five of the ten sites were associated with group a and/or b and five were associated with group c (Fig 5).

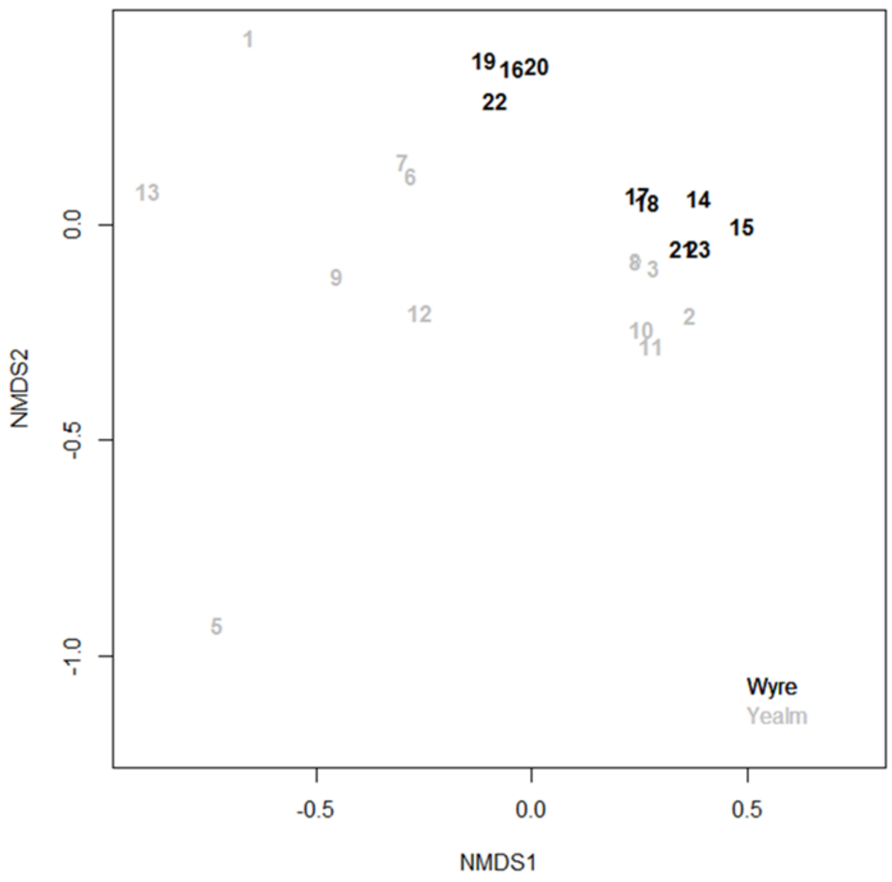

Figure 3.4. An ordination plot showing the dissimilarity in land cover mosaic across the contributing catchments associated with sample points from the Yealm (grey) and Wyre (black). Increasing distance between points illustrates increasing dissimilarity in land cover make up between catchments.

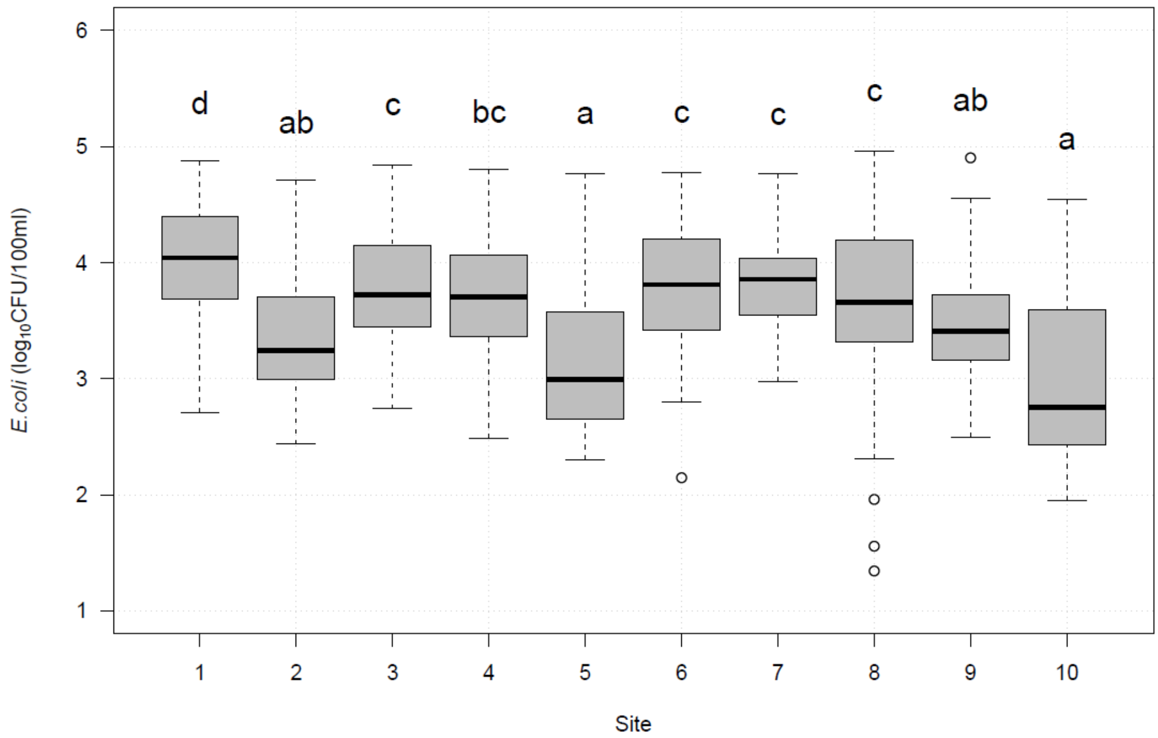

Figure 3.5. Boxplot illustrating the variability in the concentration of *E. coli* across samples from the Wyre catchment. Letters above the boxes indicate significant differences in FIO concentrations between sites as determined by a Dunns test.

The potential for seasonal differences in SCIMAP's performance can be investigated by comparing model output with observed data split into winter or summer months. This revealed that there was some variance in model performance depending on the season of interest (figure 3.6). For the Yealm, SCIMAP performance appeared to reduce during winter months while an opposite more pronounced effect was observed for the Wyre. The possibility of a relationship between the identifiability of a land cover's optimum risk weighting and it's representation in a catchment is explored in figure 3.7. As the percent coverage of a land cover class increased the identifiability of an optimum risk weighting appeared to decrease. The pattern was more pronounced in the Yealm catchment.

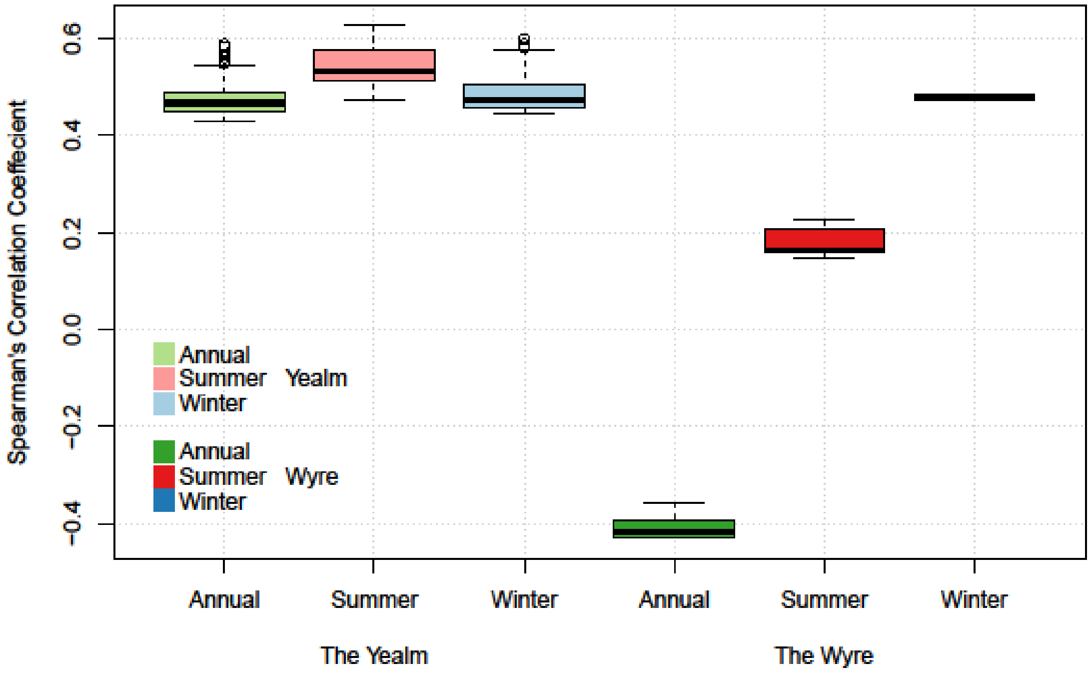

Figure 3.6. Boxplot illustrating the variance in performance associated with the 1% best performing simulations when observed data is subset according to season.

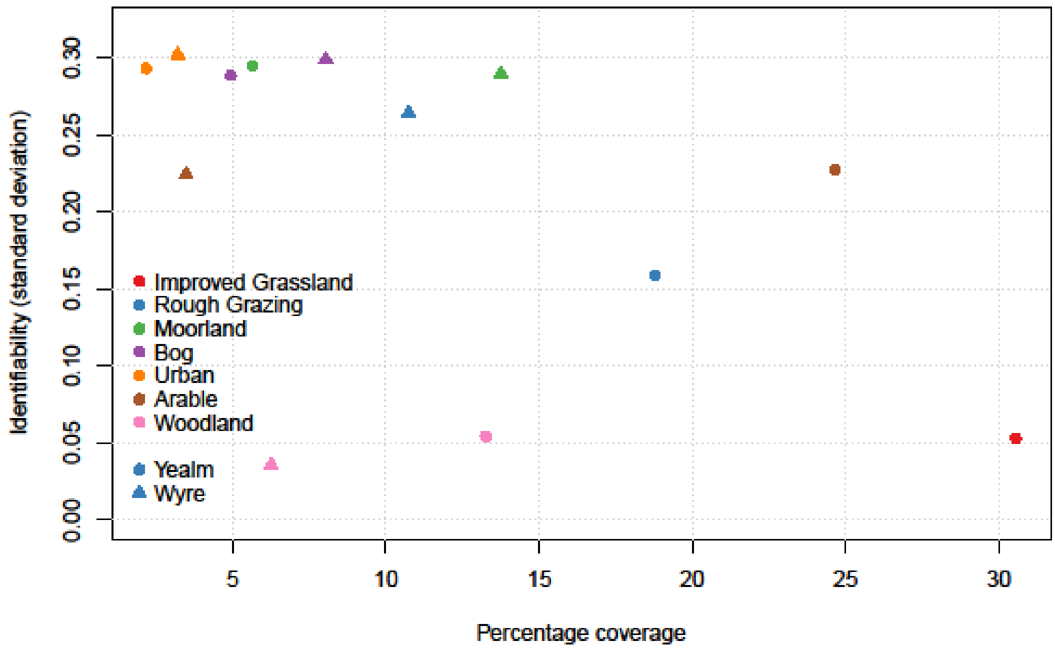

Figure 3.7. A scatterplot of optimum risk weighting identifiability and percent coverage of the associated land cover.

3.5. Discussion

This study provides a novel application of the SCIMAP model fitted against historical *E. coli* data collected across two UK catchments. The performance of SCIMAP in the prediction of diffuse FIO risk in the catchments studied was variable, with a higher degree of agreement between predicted and observed FIO risk in the Yealm than in the Wyre catchment. Even where SCIMAP performed well there was variability in the certainty with which risk weightings could be applied to land cover types. There are several reasons why model performance might be poor or why assignment of land cover risk weightings was uncertain. First, high risk weightings may offset low risk weightings resulting in a wide range of optimum risk values; this problem is associated with covariance between one or more land covers where the land cover mosaic is similar between catchments. Second, it is possible that a land use either does not exist in a catchment or represents only a small proportion of the catchment meaning that the signal from this land use is weak. Third, a land cover class may be too broad combining too many different availabilities of FIOs (Reaney et al. 2011). Finally processes that influence FIO fate (e.g. die-off, persistence, affinity to particles, etc.) may be important to consider alongside processes that govern transfer and SCIMAP, in its current form, does not adequately account for the former.

In the Wyre catchment, areas associated with improved grassland and rough grazing were assigned a medium risk which was unexpected as these areas are associated with agricultural practices such as increased spreading of farmyard manure and slurry and livestock grazing providing a high availability of FIOs (Kay et al., 2010). When the relative coverages of land covers are similar between catchments the fitted approach cannot determine which land cover is responsible for a change in in-stream risk, resulting in high risk weightings being offset by low risk weightings or *vice versa*. It is possible that there was too much similarity in land cover between the sub-catchments investigated in the Wyre catchment. The overall composition of land covers in the sub catchments shows how sub catchments of the Wyre largely fall into two groups of similar land cover mosaic. Statistical analysis revealed that there was a difference in the *E. coli* concentration between sites of the Wyre, however it was not apparent whether this difference was large enough for the SCIMAP fitted approach to delineate risk values for the different land covers. While too few catchments were investigated in this study to determine the level of difference in land cover between catchments required for the SCIMAP fitted approach to be applied optimally, Milledge *et al.* (2012) investigated 11 catchments across England and that dataset was used to determine how identifiability of land cover risk values associated with increasing diversity of land cover types in catchments. However, Milledge et al (2012) used a diffuse nutrient pollution dataset, whereas for FIOs we were limited to two catchments, largely because in the UK there are a limited number of spatial datasets of FIOs across catchment systems (Oliver *et al.*, 2016).

Even when an influence on model performance was apparent, a large standard deviation in optimum risk weightings was seen for many land cover classes in both the Wyre and the Yealm, indicating that the identifiability of an optimum risk weighting was low and could only be applied with a relatively low degree of certainty. As representation of a land cover decreases, its signal is likely to decrease making its optimum risk weighting harder to identify. Previously, a positive relationship between percentage coverage of land cover in a catchment and the identifiability of its optimum weighting has been observed when considering nutrient pollution (Milledge *et al.* 2012). This also appeared to be the case for FIO pollution, with decreasing standard deviation in optimum risk weightings as percent coverage of a land cover class increased. All of the catchments studied in the Wyre were dominated by improved grassland leaving little space for other land covers. This may, in part, explain the large standard deviations observed for this catchment for all of the land covers. In contrast, the increased variation in land covers in the Yealm may have resulted in smaller standard deviations recorded for improved grassland, rough grazing and woodland land covers. It has been shown that the use of the SCIMAP fitted approach can be improved when close consideration is given to the location of sampling points (Reaney et al., 2011). Thus, ensuring that contributing catchments of monitoring points vary in their land cover make up as much as possible is a clear priority in order to maximise the identifiability of land cover risk weightings. This is likely to be challenging in catchments where there is a dominance of one land cover, for example those associated with agriculture. Thus, carefully considered sampling campaigns to ensure appropriate spatial coverage of an observed FIO dataset across catchments are essential in order to verify SCIMAP performance in predicting FIO risk following future model refinements. This reinforces the need for good quality monitoring distributed across stream networks, not just end-point receptors.

Optimum risk weightings may be hard to identify if a land cover class is too broad encompassing many different availabilities of FIOs. It is possible that the availability of FIOs in the landscape depends, in part, on livestock density and that livestock density will vary between farms. Incorporating this information into land cover classes associated with agriculture is possible through use of Agricultural Census data which provides information on livestock density in 2km^2 grid squares. However, there are potential issues with using this information (Winter et al., 2010). For example, livestock are assumed to be located in the 2km^2 associated with a farm address and yet it is likely that farm managers will graze livestock, or produce silage or hay, on land that is either owned or rented beyond this 2km^2

area. Nevertheless, this information may provide an adequate compromise in terms of understanding the variation of FIO pollution risk across catchments resulting from variable stocking densities.

In addition, the improved grassland land cover class is likely to encompass many different management regimes that are likely to represent different availabilities of FIOs, which may have influenced modelled outputs from SCIMAP. For example, improved grassland can be managed for livestock grazing and the source of FIOs will come in the form of faecal deposits from livestock. Improved grasslands can also be managed for silage production where spreading of slurry is likely to present a risk of FIO pollution. Across a catchment it is likely that there will be differences between farm management (Winter et al., 2010). For example, some dairy farms are now opting to house dairy cows on a permanent basis, particularly in wetter regions of the UK, whereas others continue to adopt a more traditional split between summer pasture grazing and winter housing for cows. The environmental risks that these contrasting management systems pose will differ (Harmel *et al.* 2010). Permanent housing of dairy cattle can result in the production of more slurry and/or farmyard manure, putting pressure on storage infrastructure and requiring more frequent application of organic fertiliser to the landscape, though on farms with sufficient storage those applications can be timed better to reduce coincidence with heavy rainfall. By contrast, on those farms that adopt grazing regimes there will inevitably be deposition of fresh livestock faeces on pasture, leading to an accumulation of untreated faeces containing high concentrations of FIOs (Muirhead 2009; Oliver 2014). The concentration of FIOs and dynamics of their mobilisation will vary between faecal, slurry and manure matrices driving variability in their respective risk to watercourse microbial quality (Hodgson *et al.* 2009; Guber *et al.* 2013; Blaustein *et al.* 2015b). Thus, augmenting the improved grassland land cover with management regime and livestock density may improve SCIMAP's characterisation of the spatial variability of FIO risk.

At present SCIMAP's prediction of diffuse pollution risk is time integrated and an annual average risk is predicted. An approach which considers seasons separately may be more appropriate when considering diffuse FIO pollution because the extent to which watercourses receive FIO pollution is likely to vary between seasons (Kay *et al.* 2008). For example, the persistence of FIOs in the landscape is dependent on abiotic conditions such as temperature (Martinez *et al.* 2013) and moisture (Moriarty and Gilpin 2014), which will vary between seasons (Oliver & Page 2016). Additionally, mobilisation of FIOs from landscape reservoirs will vary depending on patterns of rainfall (Blaustein *et al.* 2015 b) and, from a UK perspective, the regulatory end-point receptors, i.e. bathing waters, are monitored seasonally over the summer. The seasonal differences in SCIMAP's performance were more pronounced in the Wyre catchment. Therefore, it may be possible that an improvement in model performance can be achieved through accounting for characteristics of FIO fate and transfer that vary seasonally.

An interesting and surprising observation in this study was that for the Yealm catchment improved grassland was assigned a low risk value. This was unexpected because this land cover type can be associated with activities that might produce a high availability of FIOs (McGrane *et al.* 2014). For example, improved grassland is used to graze livestock, which deposit fresh faeces into the landscape or it can be amended using slurry and manure for silage production creating a source of FIOs in the environment (Hodgson *et al.* 2009; Blaustein *et al.* 2015b). This has been further supported by regression (Kay *et al.* 2010, Tetzlaff *et al.* 2012, McGrane *et al.* 2014) and export coefficient (Kay *et al.* 2008) approaches that have suggested an association between FIO pollution and land covers linked with the management of livestock and their manure. Our study assigned risk to land cover types relative to all other land cover types in the catchment. It is possible that, for this particular catchment, another land cover type was more risky than improved grassland. In the Yealm catchment, areas of rough grazing were assigned a high risk value and an optimum mean value of 0.78. Perhaps rough grazing in this catchment provides more FIO pollution to the river network than improved grassland and therefore inputs from extensive grazing, most likely via sheep, are more important than inputs from areas of improved pasture for this particular catchment. Similarly, in the Wyre catchment, a medium risk value (0.58) was assigned to areas associated with rough grazing but our confidence in the interpretation of this finding is low given the covariance of land cover types in the catchment. Issues concerning the many availabilities of FIO in one land cover that have been discussed may also be a factor in the assignment of a low risk weighting to improved grassland areas. Of course, the fact that land cover information is derived from remote sensing techniques may also influence results. There is some overlap in the spectral properties of improved grassland and rough grazing, neutral, acid and calcareous grasslands meaning in some cases it can be difficult to delineate between these land covers (Morton et al. 2011), which are likely to vary in their susceptibility to FIO contamination. Additional ground- truthing may improve model performance but is restricted by time and resource constraints.

In this experiment the SCIMAP modelling framework was applied to diffuse FIO pollution in two catchments in England. Model performance was variable with better agreement between modelled outputs and observed data in the Yealm catchment than for the Wyre catchment. For the Yealm catchment, where model performance was good, there was uncertainty involved with the assignment of optimum risk weightings. This would suggest that, in its current form, SCIMAP is not yet optimised for mapping FIO risks. It would, however, be surprising for a model developed to describe an inert pollutant such as fine sediment to perfectly describe the fate and transfer of a living organism and these results should not be viewed as a failure of a modelling framework, but rather as a learning process in which the development of new hypotheses can be framed, and further developments of SCIMAP for predicting FIO risks can occur (Beven 2007). There are significant differences between the processes of fine sediment and FIO fate and transfer in catchment systems. For example, sediment is a conservative pollutant that persists indefinitely in the environment where FIOs, once excreted from the alimentary canal, die-off over time (Stocker *et al.* 2014). In addition, there is an unlimited store of sediment in the landscape where the FIO reservoir will be finite; sediment pollution is therefore transport limited and FIO pollution source limited (Sigua et al., 2010). This study has provided the necessary evidence to highlight that the adaptation of SCIMAP to account for FIO fate and transfer will likely mark a significant departure from previous iterations of this risk-based framework.

3.6. Conclusion

This research has provided a 'bench-marking' modelling experiment to determine how well the current SCIMAP framework for diffuse fine sediment pollution can be applied to map diffuse FIO pollution risk in catchment systems. Overall performance was variable with reasonable performance of the model for the Yealm catchment but poor outputs when tested in the Wyre catchment. In addition, assignment of risk weightings to land cover types exhibited uncertainty for all land covers, excluding woodland in both catchments and improved grassland in the Yealm catchment. However, a number of opportunities for the development of SCIMAP to account for diffuse FIO pollution risks have been identified, paving the way for a roadmap of future research needs. First, the fitted approach developed by Milledge *et al.* (2012), which was used to train SCIMAP land cover risk weightings to individual catchments, can be hampered by the mosaic of land covers across a catchment and therefore, when using this approach, targeted effort must be made to ensure that there is variation in the composition of land cover in the contributing catchments of in-stream monitoring points. Second, management of livestock and their faeces is likely to vary across catchments and the risk to microbial water quality is likely to differ between different management regimes. Therefore, a single land cover class encompassing all of these management regimes is likely to be inappropriate and multiple land cover classes discriminating animal stocking rates and housing regimes should be developed. Finally, watercourse FIO pollution has been shown to vary seasonally and the SCIMAP framework should recognise this and taking account of the varying potential for FIO survival and mobilisation through the year is a priority in this respect. SCIMAP has proven useful for the targeting of interventions for conservative nutrient and fine sediment pollution and the framework shows promise in its consideration of FIOs. However, the un-conservative nature of FIOs undoubtedly provides a different set of challenges for this model. Opportunities for addressing these challenges exist and optimising SCIMAP for prediction of diffuse FIO pollution risk will mark a significant departure from previous versions of SCIMAP and provide a useful tool for those attempting to reduce the impact of faecal contamination of watercourses at the catchment scale.

4. High resolution characterisation of *E. coli* proliferation profiles in dairy cattle, beef cattle and sheep faeces

4.1 Abstract

An increased demand for food resulting from a growing population has contributed to the intensification of livestock management practices. Grazing livestock contribute faeces to the landscape and, once mobilised, pathogens present in fresh deposits of faeces have the potential to impact upon ecosystem services related to clean water such as bathing, drinking and shellfish harvesting waters. The extent to which these ecosystem services may be impacted upon can, in part, be attributed to the survival of pathogens in the landscape reservoir. Therefore, understanding the survival of pathogens in the environment is key if catchment management practices that reduce impacts to services provided by clean water are to be developed. Here the survival of *E. coli* (an internationally accepted faecal indicator organism (FIO)) in the faeces of dairy cattle, beef cattle and sheep is investigated using a controlled environment facility simulating diurnal variation of temperatures typically experienced during a British spring and summer. The experiment, conducted under controlled conditions, reproduced *E. coli* regrowth in livestock faeces which has previously been observed in field trials. This allowed for the development of a non-linear description of FIO survival dynamics in faeces 30 days post defecation and the relative risk of faeces from different livestock contributing *E. coli* to the environment to be investigated. A new framework for predicting *E. coli* regrowth in livestock faeces immediately following defecation has been developed. This new knowledge regarding the proliferation of *E. coli* in faeces can provide input into tools that are designed to inform catchment scale

assessments of watercourse *E. coli* contamination. Such assessments may help catchment managers and policy makers to manage health risks originating from grazing livestock on pasture.

Keywords: diffuse pollution, faecal indicator organism, microbial die-off, survival curves

4.2 Introduction

Increased demand for food production has led to approaches that aim to deliver sustainable intensification in agricultural systems (Rockström *et al.* 2017). Despite best efforts to promote sustainable intensification, the need to feed a growing population can still lead to poor management of livestock, and the unsustainable use of agricultural land and (in)organic fertilisers, with the potential to impact negatively on wider environmental quality (Yang *et al.* 2016). For example, increased livestock numbers on-farm could lead to higher volumes of livestock faeces being applied to land, either as manure, slurry or via direct defecation, introducing large quantities of faecal indicator organisms (FIOs) to agricultural land. Importantly, the mobilisation and delivery of FIOs to receiving waters following rainfall threatens important ecosystem services related to clean and safe drinking, bathing and shellfish harvesting water (Clements *et al.* 2015; Murphy *et al.* 2015; Wu *et al,* 2016).

E. coli is the most routinely monitored FIO in environmental samples, though its detection does not imply the presence of pathogenic microorganisms in the same sample (Bradshaw *et al*. 2016; Pachepsky *et al.* 2016). However, detection of *E. coli* in soil or water does indicate faecal contamination of the environment. The magnitude of *E. coli* burden contributed to land from agriculture is therefore a useful index when assessing the vulnerability of nearby watercourses to microbial pollution risk (Dymond *et al*., 2016). Understanding how the landscape burden of *E. coli* varies in space and time is challenging, due to the complex survival dynamics of *E. coli* under different abiotic conditions (Oliver *et al.* 2018). A particularly important source of *E. coli* in agricultural landscapes is freshly excreted livestock faeces which, unlike most slurry and farmyard manure, does not undergo any storage or treatment prior to land application and therefore often contains a higher concentration of FIOs (Chadwick *et al.* 2008).

Controlled laboratory studies, under constant temperature regimes, have been used extensively to determine the impact of specific environmental factors on *E. coli* persistence in livestock faeces. Outputs from such studies have been deterministic first-order decay functions that describe the exponential die-off of the target population under different temperatures (Wang *et al.* 2004), dry matter content of protective media (Ishii *et al.* 2010) or contrasting soil types (Lau & Ingham 2001). To complement the mechanistic understanding delivered via controlled laboratory studies, field-relevant investigations have profiled *E. coli* persistence in livestock faeces exposed to combinations of variable and interacting environmental factors (e.g. Oliver & Page, 2016; Moriarty et al., 2011; van Kessel et al., 2007). This field-relevant research has identified significant deviations from the first-order decay functions observed under controlled conditions, with *E. coli* cell growth and protracted survival leading to much longer persistence than that predicted from first-order die-off models (Brouwer *et al.* 2017).

Whether research has opted for a field or laboratory focus, there has been little direct comparison of persistence profiles in multiple faecal types under contrasting conditions. Most research has focussed on bovine faeces (e.g. Oladeinde et al. 2014; Martinez *et al*., 2013) with relatively little information currently available for ovine faeces (Moriarty et al. 2011; Hodgson et al 2009). Furthermore, the observed growth phase of *E. coli,* commonly identified in field-relevant studies, represents an interesting shift in our understanding of *E. coli* survival; however knowledge of what governs the rate and magnitude of post-defecation *E. coli* cell growth is lacking (Oliver *et al.* 2016b). This problem is compounded by observations of *E. coli* growth occurring at the beginning of long-term studies, leading to inferences of *E. coli* growth being based on only a few data points. Therefore, the modelling of *E. coli* persistence in faecal deposits may justify a piecewise approach whereby the initial growth phase is described separately from the subsequent decay phase. For the initial growth phase, non-linear modelling approaches may provide a better approximation of the system than the linear approaches employed previously (Oliver *et al.* 2010). For example, we might expect regrowth to be most rapid in fresh faeces, reducing through time as conditions within the deposit become less favourable. Thus, fitting an asymptotic model to describe the *E. coli* population growth rate, which declines as it moves toward a maximum, could provide an opportunity to improve upon linear modelling approaches.

While field-relevant studies are useful for investigating *E. coli* behaviour in the landscape, the complex mix of interacting environmental factors makes it difficult to identify the dominant drivers that govern *E. coli* persistence and growth in livestock faeces. Yet controlled static-temperature laboratory studies oversimplify real world conditions and rarely, if ever, capture growth as observed in the field. The use of a more advanced controlled environment facility (CEF) offers the potential to minimise uncertainty from variable interacting factors but elevate the quality of simulated controlled conditions, e.g. by allowing diurnal temperature regimes and varying daylight hours, but have yet to be exploited for exploring *E. coli* persistence in the context of environmental management. The aim of this study was to investigate *E. coli* persistence in beef, dairy and sheep faeces, using a CEF. Our objectives were to: (i) use high-resolution sampling to investigate and model the potential for *E. coli* regrowth previously unaccounted for by controlled laboratory studies; (ii) determine whether *E. coli* regrowth profiles vary in different livestock faeces across contrasting seasonally-defined conditions; (iii) investigate the temperature sensitivity of *E. coli* persistence within different faecal types; and (iv) characterise differences in the *E. coli* hazard associated with faecal deposits from different livestock types.

4.3 Materials and Method

4.3.1 Experimental climate chambers

All experiments were carried out in climate cabinets, which were designed to allow multifactorial climate manipulation (Snijders Microclima 1750E, Tilburg, Netherlands). Cabinets were set up to mimic diurnal temperature variation experienced during a typical British spring or summer. Two temperature treatments were used: (i) typical seasonal temperatures for spring and summer based on long term average datasets; and (ii) scenarios to test climate sensitivities for spring and summer. Temperature settings were derived from 30 year averages available from the Met Office MIDAS dataset (Met Office, 2012). For the climate sensitivity experiment, temperatures of 2°C more than the MIDAS seasonal averages were used. For the spring treatment this was equivalent to UKCP09 projections for temperature increases for 2070, 2070, and 2060 for the low, medium, and high emissions scenarios, respectively. For the summer treatment a temperature increase of 2°C is equivalent to UKCP09 projections for 2040 for the low, medium and high emissions scenarios. These data were acquired from the

Scottish Climate Projections App (2017) with the Eastern Scotland region selected. The probability level used was 50% representing an equal chance of UKCP09 climate model realisations resulting in a temperature either above or below the temperature specified. Temperature in the CEF varied from an average minimum and maximum following a sinusoidal wave mimicking diurnal variation of temperature (Table 4.1). In order to simulate solar irradiance, timers were set to mimic periods of daylight and night time with UV strengths typical for the UK during the seasons of interest. UV activation periods were centred over the time of maximum temperature. Monthly means of solar irradiance were acquired from the SoDa Service (2013) and converted to a seasonal mean (Table 4.1).

Table 4.1. Controlled environment facility settings

Season	**Minimum temperature (°C)**	**Maximum temperature (°C)**	**Temperature variation (°C)**	**Hours of daylight**	**UV ($J/cm^2/day$)**
Spring	3.97	11.86	7.89	13	34.41
Spring +2°C	5.97	13.86	7.89	13	34.41
Summer	10.2	18.34	8.14	15.5	36.3
Summer +2°C	12.2	20.34	8.14	15.5	36.3

4.3.2 Experimental design

To ensure that the faeces used in the experiments was representative of the livestock diet typical for the season of interest, fresh faeces was collected during the respective season, and then transferred to the CEF. Faeces was collected from farms in Stirlingshire, Scotland and from the same herd/flock for the spring and summer treatments. For each temperature treatment, five intact replicates of dairy and beef cattle faeces, which were less than 12 h old, were collected. Dairy cattle faeces were collected from an area where cows were held prior to milking, which was cleaned twice a day. Beef cattle and sheep were grazing on pasture and the freshness of faeces was ensured by collecting deposits from the area immediately surrounding livestock. In order to collect enough faeces for the sheep experiment, each replicate was made up of pellets from five fresh deposits which were collected and homogenised. Faecal deposits had an average fresh weight of 1516 g (sd= 350 g) and 1766 g (sd = 633 g) for dairy and beef cattle, respectively. The average fresh weight for groups of five faecal deposits from sheep was 116 g (sd = 58 g). All faeces were transferred into the climate cabinets on the day of collection. Every two days the faecal deposits were misted with sterile distilled water at a rate of 1ml/100cm^2 to mimic a 'morning dew' effect and avoid a complete dehydration of the faeces under CEF conditions. Bovine faecal samples were collected for microbial analysis on a daily basis, and every other day for sheep faeces (as dictated by the smaller faecal volume associated with ovine faeces). Sampling was undertaken over a period of 20 to 30 days, depending on the volume of source material available. A small sample of faeces representing a cross section of the deposit, approximately 0.5 cm in diameter, was retrieved using a sterile spatula. This was carried out daily for 15 to 20 days after which the sampling frequency was decreased in order to retain enough material to lengthen the complete duration of the experiment to 30 days. Sampling ceased when further samples could not be taken without intersecting with areas previously sampled. The spatula used to sample the faeces was sterilised between replicates, and the faeces transferred into a sterile sample pot. Repeated sampling was used over destructive sampling; sampling repeatedly from the same faecal deposit assumed homogenisation of the generic *E. coli* population within the faecal matrix during passage through the livestock gut, as previously demonstrated by repeated spatial sampling of faecal material (Oliver, 2014).

2.3 Sample analysis

On each sampling occasion approximately 2 g of faeces were removed from each deposit; 1 g was used to determine moisture content by oven drying the sample at 105^{o}C for 24 h, with the remainder used for quantifying the concentration of *E. coli*. The number of viable *E. coli* cells in faeces was determined using standard culture-based methods and carried out within 30 minutes of the faecal samples being collected. Briefly, approximately 1 g of faeces was added to 9 mL of phosphate buffer saline (PBS) prior to shaking at 130 rpm for 30 minutes. The resulting slurry mix was then vortex mixed and serially diluted prior to inoculation onto membrane lactose glucuronide agar (MLGA) (CM1031, Oxoid; Basingstoke, UK) using the spread plate method. Agar plates were inverted and incubated for 24 hours at 37°C. All colonies counted represented presumptive *E. coli* and all sample analysis was performed in duplicate. Membrane filtration of samples was also used to complement the spread plate method and improve the limit of detection. Briefly, 1 mL of each serially diluted sample was mixed with approximately 20 mL of sterile PBS and filtered through sterile cellulose acetate membranes of 0.45 µm pore size (Sartorius Stedim Biotech; Goettingen, Germany) using a vacuum filtration unit (Sartorius). Membrane filters were then aseptically transferred to plates containing MLGA, inverted and incubated for 24 h at 37 °C. The limit of detection was 50 cells / g of wet faeces. Method blanks of PBS were used to ensure no contamination occurred during sample processing.

2.4 Statistical analysis

All *E. coli* counts underwent $\log_{10}$ transformation prior to statistical analysis. Distributions of *E. coli* were not log normally distributed as determined using the Anderson – Darling normality test and this was accounted for in subsequent data analysis.

We hypothesised that *E. coli* population growth within faecal deposits would likely be most rapid immediately following deposition, slowing as conditions within the deposit become less favourable for *E. coli* population growth. As linear modelling approaches assume a constant growth rate they would not be appropriate here; and whilst quadratic terms within a linear model can be used to address this problem, this approach can lead to predictions with negative values. Therefore, the use of more complex non-linear models is justified (Paine *et al.* 2012). The asymptotic exponential model provides an opportunity to investigate the magnitude and duration of *E. coli* growth. The asymptotic exponential form (equation 1) predicts growth rate to be fastest initially, slowing to a stationary maximum and has three parameters: an intercept (initial *E. coli* concentration); a horizontal asymptote (maximum *E. coli* concentration); and a rate constant (speed of *E. coli* population growth).

$$y = \mathrm{a} + (\mathrm{b}-\mathrm{a})e^{-\mathrm{c}x}$$

Equation 1: Where a is the horizontal asymptote, b is the intercept and *c* is the rate constant.

Classical regression approaches require data to exist under the assumptions that errors are uncorrelated and independent. However, repeated measurements of *E. coli* concentration from a given faecal deposit are not independent and are likely to be serially related. Therefore, a mixed effects approach, which incorporates a random effect allowing a model to vary between individual deposits and a temporal dependence structure between measurements, was required (Davidian & Giltinan 1995; Pinheiro & Bates 1995). Three temporal autocorrelation structures were tested: auto regressive order 1; compound symmetry; and autoregressive moving average. The Akaiki Information Criterion (AIC) was used to compare competing models and a reduction of > 2 was deemed to be an improvement in model performance. Confidence intervals were derived from an ordinary non-parametric bootstrap procedure because this conservative method makes no *a priori* assumptions about the distribution of the data (Carpenter & Bithell 2000). Where growth was not observed, a linear mixed effects model was fitted to the data incorporating the temporal autocorrelation structures described above. For the linear models, goodness of fit was quantified by calculating marginal and conditional R^2 values, as described by Nakagawa and Schielzeth (2013).

Where *E. coli* concentration growth was observed, the day of maximum *E. coli* concentration was determined for each replicate individually. A $\log_{10}$ transformation was required to normalise the data after which a one-way ANOVA and a Tukey test was applied to investigate whether the day of maximum *E. coli* concentration differed significantly between livestock/temperature treatments.

Moisture content of faeces was measured as a percentage and was thus bounded. Therefore, a logit transformation was applied and an Anderson – Darling normality test used to confirm the transformed data were from a normal distribution. A one-way ANOVA and a Tukey test was applied to determine any differences in the moisture content of faeces from the three livestock types collected during different seasons.

Data processing and analysis was implemented in the R statistics package utilising a number of third party plugins (R Core Team 2015; Wickham & Francois 2016; Graves *et al.* 2015; Pinheiro *et al.* 2015; Ogle 2015; Neuwirth 2014).

4.4 Results and discussion

The high resolution monitoring in this study has, for the first time, provided data that has enabled the development of a non-linear, asymptotic description of *E. coli* proliferation in livestock faeces immediately following deposition. Results are based on a total of 364, 383 and 255 faecal samples taken from beef, dairy and sheep faeces, respectively. This represents the most sustained period of high frequency sampling of three livestock faecal types thus far reported, providing an unparalleled evidence base with which to characterise *E. coli* growth patterns in livestock faeces.

Initial concentrations of *E. coli* in livestock faeces are shown in table 4.2 and, with the exception of the spring beef treatment, were in line with previously published research for all livestock and season combinations; the concentration of *E. coli* in the spring beef experiment were approximately 0.5 $\log_{10}$ CFU g^{-1} of dry faeces lower than values commonly reported in the literature (e.g. Hodgson *et al.* 2009; Muirhead *et al* 2009; Oliver *et al.* 2010; Oladeinde *et al.* 2014). Both temperature regimes applied to dairy faeces for the summer treatment resulted in maximum

Table 4.2. Average (n=5) initial, maximum, and day of maximum *E. coli* concentration for faeces from three different livestock types under four temperature regimes. All values are given as log_{10} CFU/g dry faeces.

Livestock	Temperature	Mean initial *E. coli* concentration (log_{10} CFU g^{-1} dry wt. faeces)	Mean maximum *E. coli* concentration (log_{10} CFU g^{-1} dry wt. faeces)	Mean day of maximum *E. coli* concentration	Mean initial moisture content (%)
Beef	Spring	3.87	5.02	6.28	87.2
	Spring +2°C	5.22	5.96	1.90	85.5
	Summer	6.03	8.51	15.8	91.2
	Summer +2°C	5.70	8.65	13.2	90.4
Dairy	Spring	6.19	8.11	20.4	84.9
	Spring +2°C	6.18	8.28	11.8	85.0
	Summer	6.60	8.87	10.7	86.5
	Summer +2°C	6.58	9.10	10.5	85.6
Sheep	Spring	6.79	8.20	11.5	66.9
	Spring +2°C	6.41	8.59	16.9	67.3
	Summer	7.03	8.73	10.4	71.9
	Summer +2°C	7.18	8.88	14.0	72.0

concentrations of *E. coli* that were greater than maximum concentrations reported in previous published studies (table 4.2). The summer and summer +2°C treatments exceeded the previously reported maximum by 0.32 and 0.49 log_{10} CFU *E.coli* g^{-1} dry faeces, respectively (Muirhead *et al.* 2005; Soupir *et al.* 2008a; Oladeinde *et al.* 2014; Oliver 2010; Van Kessel *et al.* 2007). While efforts were made to simulate field conditions these inflated maxima may be due to the experiment being carried out inside a CEF where faecal deposits were isolated from stressors present in the field, and did not encounter, for example, cell wash-out following rainfall. Furthermore, under field conditions soil macrofauna such as beetles and earthworms break up faeces, which can affect the survival of *E. coli* (Ryan *et al.* 2011; Hénault-Ethier *et al* 2007; Pedersen & Hendriksen 1993). Given that the faecal material in this experiment may have been protected from some factors experienced in the field, extrapolation of our regrowth model to field conditions must be done with a degree of caution, with recognition that the experiment was undertaken to develop greater insight into what drives patterns of *E. coli* regrowth.

The day of maximum *E. coli* concentration is shown in table 4.2 and, where *E. coli* growth was observed, did not differ significantly between temperature/stock type combinations ($p \geq 0.05$). However, a large variation was apparent. Day 13 (sd=6) was, on average, the timing of maximum *E. coli* concentration which is similar to previous studies with an average of 9 (sd = 9) days (Muirhead *et al.* 2005; Soupir *et al.* 2008a; van Kessel *et al.* 2007; Oladeinde *et al.* 2014; Oliver *et al.* 2010). An earlier day of maximum concentration was observed for beef cattle faeces in the spring treatment; little or no *E. coli* growth was associated with these faeces and early maximum *E. coli* concentrations arise due to a small deviation from a static phase of *E. coli* persistence. These data suggest that livestock type and temperature do not affect the time taken to reach a maximum *E. coli* concentration during regrowth; this is important in that it might present an opportunity to simplify the parameterisation of *E. coli* persistence. However, our experiment was conducted under moderate temperatures typical of spring and summer in the UK. Regions of the world where temperatures are higher, and closer to the optimum for *E. coli* replication (37^{o}C), may promote further *E. coli* regrowth.

An asymptotic model form provided good fit to the data for all instances of *E. coli* regrowth. Model results are shown in figure 4.1 and asymptotic model parameters for the models associated with each of the livestock types are given in table 4.3.

Table 4.3. Table of model parameters associated with asymptotic models for each of the livestock types. Numbers in parentheses indicate lower and upper bounds of a 95% confidence interval.

Livestock	**Temperature**	**Model parameters** Intercept ($\log_{10}$ CFUg^{-1})	Rate constant	Asymptote ($\log_{10}$ CFUg^{-1})	**Magnitude of growth (Asymptote - Intercept) ($\log_{10}$ CFUg^{-1})**
Beef	Spring	-	-	-	-
	Spring +2°C	-	-	-	-
	Summer Summer +2°C	5.88 (5.51, 6.20)	0.18 (0.13, 0.28)	7.72 (7.35, 8.09)	1.84
Dairy	Spring	6.35 (6.19, 6.49)	0.23 (0.17, 0.32)	7.02 (5.41, 8.62)	0.67
	Spring +2°C			7.00 (6.26,7.71)	0.65
	Summer			7.95 (6.35,9.43)	1.6
	Summer +2°C			8.68 (7.05,10.29)	2.33
Sheep	Spring	-	-	-	-
	Spring +2°C	-	-	-	-
	Summer Summer +2°C	7.10 (6.63, 7.52)	0.41 (0.24, 0.91)	7.91 (7.70, 8.28)	0.81

The asymptotic form contains three parameters: a starting value, rate constant and an asymptote. A fixed effect of temperature category was applied to the asymptote parameter only because: (i) the observed data showed no significant difference in the time taken to reach maximum concentration between temperature treatments, and (ii) the starting value is expected to be similar because, for each model, the source of faeces is the same. For the data associated with beef and sheep faeces, a solution for an asymptotic model was not achieved when the data from the spring temperature treatment were included (i.e. no growth was observed for those treatments). A solution was achieved when the summer data were considered separately. For these models, including an effect of temperature sensitivity (present / present +2°C) on the asymptote did not improve model performance (i.e. the reduction in AIC was < 2). For the dairy faeces data, an asymptotic model was fitted to both the spring and summer data with the inclusion of the temperature sensitivity effect on the asymptote improving model performance (AIC reduced by 11.98). A plot of the autocorrelation function associated with the models for the beef and dairy cattle experiment showed some temporal autocorrelation between residuals at different time points. The best performing autocorrelation structures were a compound-symmetry (AIC reduced by 12.87) and auto-regressive moving average autocorrelation structure (one auto-regressive parameter and one moving average parameter) (AIC reduced by 12.30) for the beef and dairy treatments, respectively. No *E. coli* growth was observed during the spring treatments for beef and sheep faeces; therefore, linear mixed effects models were fitted to these data. The *E. coli* concentrations in beef faeces decreased at a rate of -0.04 (-0.04, -0.03) $\log_{10}$ CFU g^{-1} dry faeces per day and had an intercept of 4.78 (4.19, 5.30) $\log_{10}$ CFU g^{-1} dry faeces (numbers in parentheses show lower and upper 95% confidence intervals). Marginal and conditional R^2 values associated with this model were 0.13 and 0.76, respectively. For the sheep data, a linear model showed a negative relationship between *E. coli* concentration and days

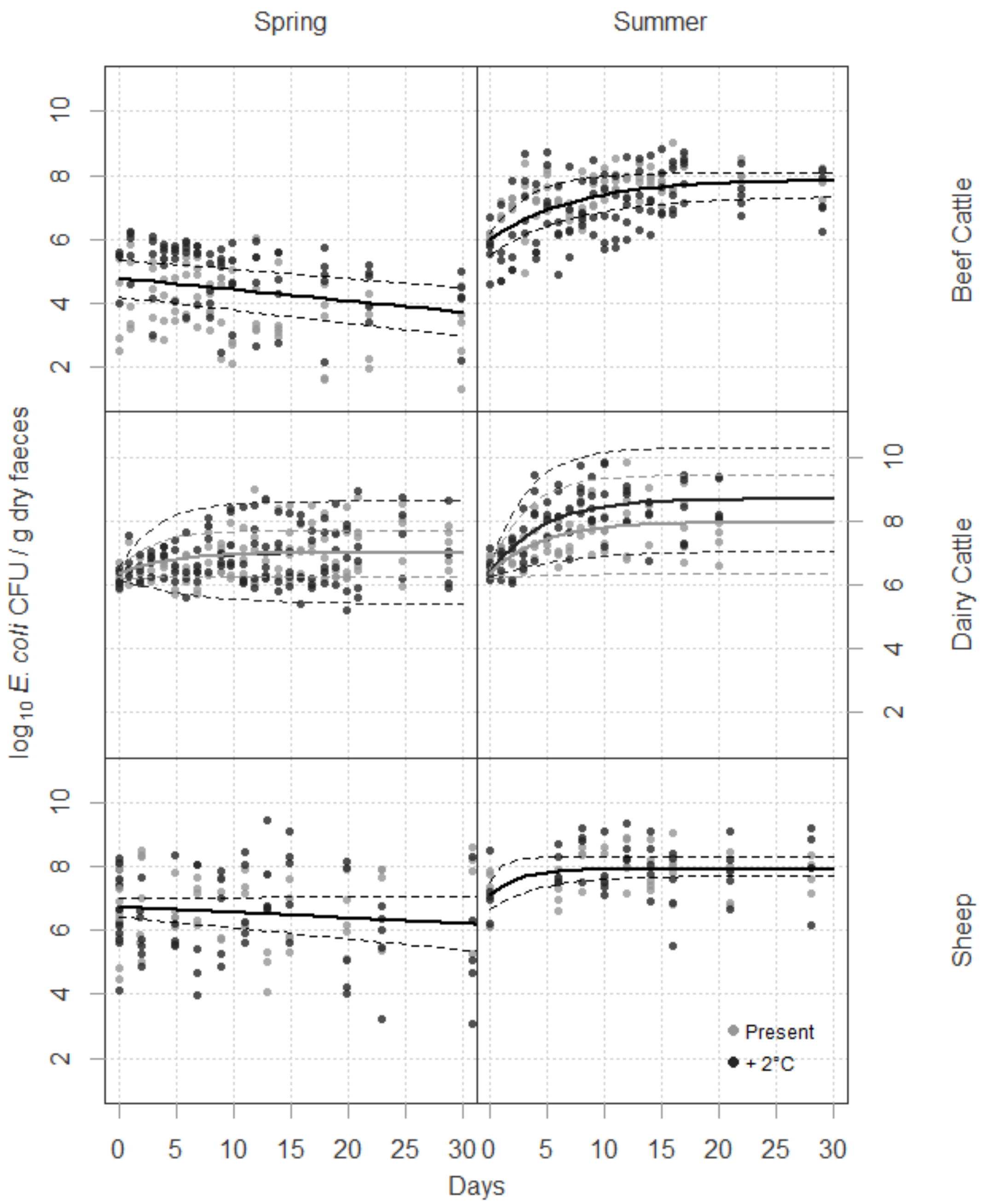

Figure 4.1. Scatter plot of *E. coli* concentrations through time. Where growth occurred lines are predictions of nonlinear (asymptotic exponential) mixed effects modelling. Where no growth was apparent lines illustrate linear mixed effects models. Dashed lines indicate 95% confidence intervals derived from normal non parametric bootstrap. Where the prediction is coloured black there was no improvement in model performance (change in AIC <2) when the present average/present average +2°C treatment was incorporated.

since defecation (slope parameter = -0.02 (-0.03, 0.00) $\log_{10}$ CFU g^{-1} per day, intercept = 6.72 (6.41, 7.01) $\log_{10}$ CFU g^{-1}). However, this was associated with a large uncertainty and a gradient of 0 within the 95% confidence interval. Marginal and conditional R^2 values were 0.02 and 0.06 respectively, suggesting that the model does not explain the variance in the data (Fig 1). This weak relationship may be due to *E. coli* concentrations remaining largely static within sheep faeces under the spring temperature regime, with variation between individual deposits greater than the change over 30 days.

Where *E. coli* concentration growth was observed, separate non-linear, asymptotic models were fitted to the different livestock types because a model could not be fitted when all instances of growth across all livestock type/season combinations were included. Separate models are justified given the marked differences in the management of beef cattle, dairy cattle and sheep. For example, there will be differences in the diet and reproductive status of different livestock. In the future, it may be possible to develop a unified model of *E. coli* persistence in livestock faeces but this would require a large amount of supplementary information on different management regimes, which at present is unavailable.

In this study, relatively few individuals (i.e. faecal deposits) were investigated intensively in order to characterise a profile of *E. coli* regrowth through time. The new understanding presented here demonstrates that future studies can investigate more faecal deposits but less intensively; this may allow the mixed effects approach to identify contrasts between livestock type/temperature combinations more easily. The use of an asymptotic exponential form has limitations in that it is only applicable at temperatures where growth is observed and clearly it would be advantageous to develop a more flexible model that can account for temperatures where *E. coli* growth does not occur. Despite these potential limitations this new model provides a step change in our understanding of *E. coli* persistence in fresh livestock faeces.

Parameters derived from the models can be used to compare and contrast *E. coli* persistence in the faeces of the different livestock studied. For example, the magnitude of growth can be taken as the asymptote minus the intercept. For dairy faeces, the summer +2°C temperature treatment showed the highest level of *E. coli* growth, whilst the dairy faeces under the spring temperature treatments showed the lowest increase in *E. coli* concentrations. Model results showed that only dairy faeces under the summer temperature regime showed a difference in the magnitude of growth between the present and +2°C treatment, with the +2°C treatment showing an increase of 0.73 $\log_{10}$ $CFUg^{-1}$ dry faeces more *E. coli* growth relative to the standard summer temperature treatment. For beef faeces, *E. coli* concentration growth was only observed in the summer temperature treatments with only a small difference in *E. coli* concentration growth between the summer and +2°C temperature treatments. The magnitude of *E. coli* growth observed in the beef faeces within the summer temperature treatments was comparable to that recorded in dairy faeces. *E. coli* growth was only observed in sheep faeces under the summer temperature treatment with no difference in the magnitude of *E. coli* growth between summer and +2°C temperature treatments. For all livestock types, sheep faeces showed the lowest *E. coli* growth during the summer experiment. The rate constant of the asymptotic equation provided insight into the rate of *E. coli* growth, with larger values indicating faster growth. In order of fastest to slowest for rates of *E. coli* growth: sheep faeces >dairy faeces >beef cattle faeces (table 4.3).

Increased potential for *E. coli* growth during warmer temperatures for all three livestock types was observed. Moisture content also appeared to affect the rate of *E. coli* concentration change. The average initial moisture content of faeces is shown in table 4.2. The effect of moisture on the change in *E. coli* concentration through time was observed by taking the moisture content associated with an individual sample and the rate of *E. coli* concentration change between that observation and the corresponding previous observation. Figure 4.2 illustrates that *E. coli* growth in dairy and beef cattle faeces is more likely to be observed at higher moisture contents with beef cattle faeces showing some *E. coli* growth at lower moisture contents. No clear pattern was evident for sheep faeces and growth rates appeared to decrease as deposits dried over time. A one-way ANOVA revealed that beef cattle faeces collected during the spring were 4.5% drier than those collected in the summer (figure 4.3) and this reduction in moisture content associated with spring faeces may have influenced differences in the survival of *E. coli* in fresh beef cattle faeces that were observed between the spring and summer temperature treatments. Reductions in the moisture content of faeces are likely due to differences in the diet of beef cattle in the two seasons studied. For example, the diet of beef cattle in spring was more likely to be supplemented with hay, silage and concentrates, whereas in the summer the diet was dominated by fresh grass. The effect of diet on *E. coli* persistence in faecal deposits is likely to be multifaceted with moisture being one of many controlling factors; for example, Donnison *et al.* (2008) reported a reduced burden of FIOs in cattle that were fed silage

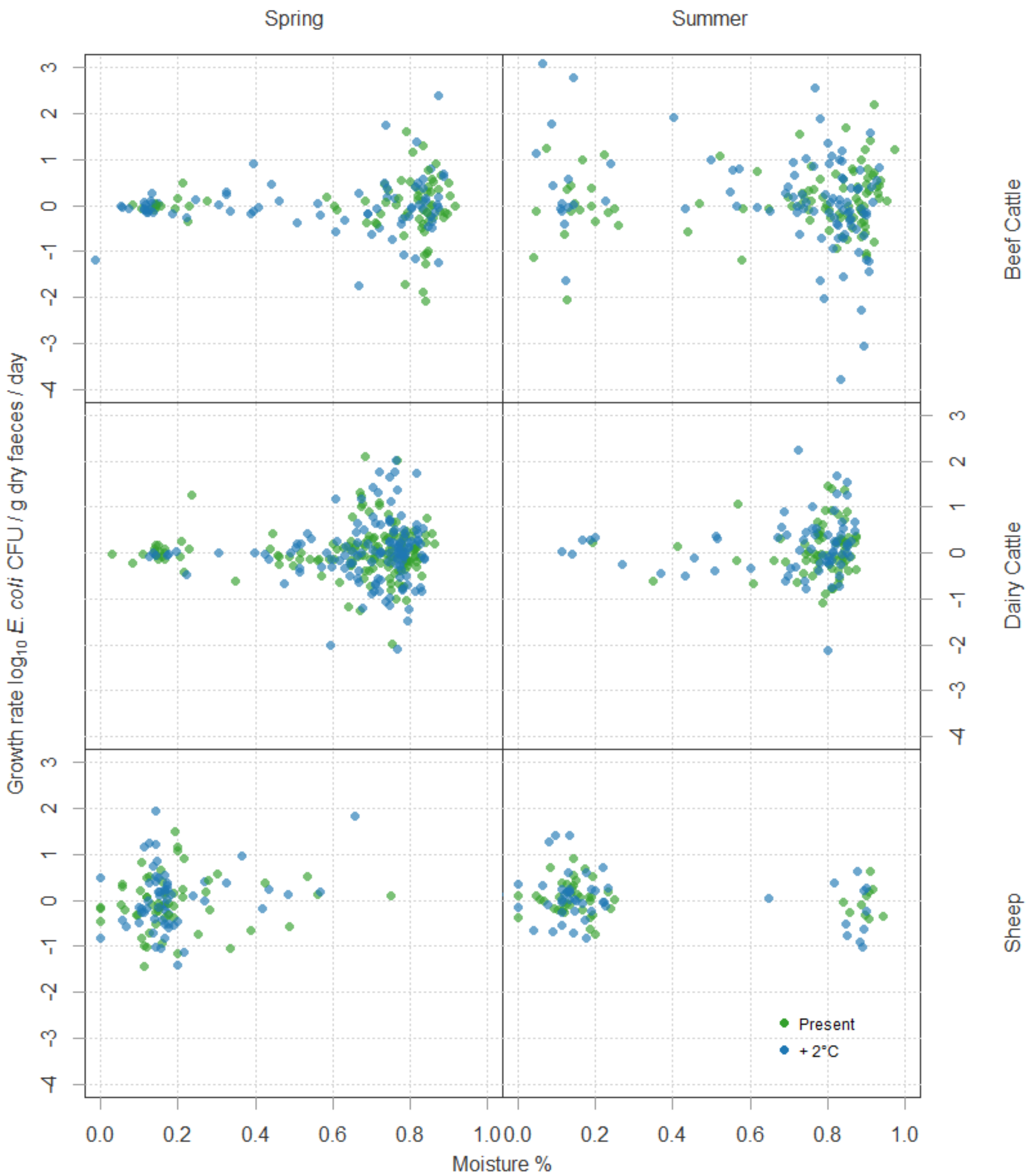

Figure 4.2. A scatter plot of *E. coli* concentration growth rate against moisture content. Colours indicate present average or present average +2°C temperature treatments.

compared to cattle grazing pasture and suggested reductions in rumen pH as a controlling variable. If the difference in *E. coli* concentrations in beef cattle faeces between seasons is extrapolated to a catchment scale the differences in the size of the landscape reservoir of *E. coli* through time are likely to be marked, reinforcing the importance of accounting for seasonal persistence profiles of *E. coli* in catchment-scale models (Oliver *et al*., 2018). Therefore, further investigation into the influence of cattle diet on FIO concentrations in faeces is warranted.

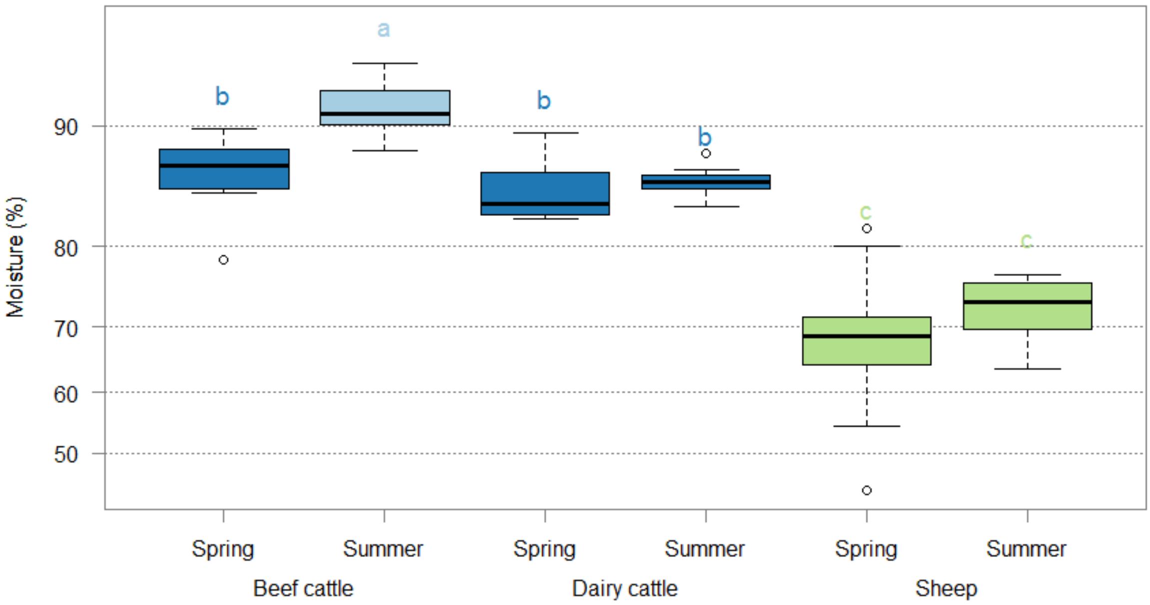

Figure 4.3. Boxplot showing logit transformed initial moisture content in the faeces of three livestock types from two seasons. Different letters and colours illustrate where results of a Tukey post-hoc test revealed differences between livestock and season combinations. Y axis labels have been back transformed to improve interpretability.

A relationship between moisture content and *E. coli* growth rate was not apparent in sheep faeces, which were significantly drier ($p<0.05$) than the faeces of the beef and dairy cattle. This reduced moisture content may have contributed to the lack of *E. coli* growth in sheep faeces exposed to the spring temperature treatments and the very limited growth relative to dairy and beef cattle observed in the summer temperature treatments. For sheep faeces in the summer temperature treatments there were a few observations with high moisture content. These observations occurred during the first five days of the experiment with rapid drying occurring over subsequent days. Despite higher moisture content, *E. coli* growth was not apparent suggesting that moisture content is not the only limiting factor for *E. coli* growth in sheep faeces. Variations in the survival of *E. coli* in the faecal reservoir due to changes in moisture content and temperature may contribute to observed seasonal variations of watercourse FIO pollution (Cho *et al.* 2016b).

This study, operating within a CEF, succeeded in replicating *E. coli* regrowth in livestock faeces, which has previously been observed under field conditions (Oliver *et al.* 2010; Oladeinde *et al.* 2014). A key difference in our study relative to other laboratory studies is that temperature was not statically held at a single value; it varied diurnally following a sinusoidal wave form. This would suggest that diurnal variation in temperature can somehow promote a mechanism to drive *E. coli* regrowth, which controlled experiments under a constant temperature cannot replicate. From our experiment it is unclear whether it is the size of diurnal variation or the absolute temperature that drives protracted *E. coli* survival because there was only a small difference (7.89 vs 8.14°C) in the diurnal variation between the seasons studied. Furthermore, little is known about the mechanisms that promote *E. coli* regrowth in faeces but it is possible morphological changes in

E. coli cells may promote more rapid growth of *E. coli* under varying temperature compared to static temperature regimes. For example, *E. coli* growth was observed when the cells were refrigerated below the minimum temperature for growth but were exposed to warm temperatures every 12 hours. The formation of filamentous *E. coli* at temperatures colder than the minimum for growth was suggested as a driver (Jones *et al.* 2004). Likewise, Mattick *et al.* (2003) showed how refrigerated filamentous *Salmonella* spp. rapidly multiply when temperature was increased. Thus, it is possible that the development of filamentous forms of *E. coli* and subsequent rapid division over a diurnal temperature variation contributes to protracted *E. coli* survival in livestock faeces under variable field conditions compared to static temperature conditions. Investigation into the influence of *E. coli* morphology on nuances in its survival in livestock faeces is therefore warranted.

Replicating *E. coli* regrowth, as seen under field conditions, in a laboratory setting demonstrates the potential for improvements in reductionist, laboratory-based studies. For example, embedding a more accurate (but controlled) representation of environmental drivers into mechanistic studies via more sophisticated CEF functionality can reveal new insight that would be overlooked by simplistic constant temperature regimes. Clearly, interactions between multiple environmental variables in the field make it difficult to identify variables that control profiles of *E. coli* persistence; however, our study demonstrates that CEFs can be used to control some environmental variables while varying others for a more detailed investigation relative to static-temperature laboratory studies.

Diffuse pollution mitigation measures are costly and occupy valuable productive land, and therefore measures must be targeted toward areas where they will contribute to the greatest improvement in water quality and the least disruption to catchment stakeholders (Beharry-Borg *et al.* 2013). Ultimately, the results from our study can be used to improve understanding of the relative contribution of different livestock types to microbial watercourse pollution. Field burden models have shown how total *E. coli* reaches an asymptote as the introduction of new *E. coli* via fresh faecal deposits equilibrates with die off of *E. coli* in existing faecal deposits (Oliver *et al.* 2010). Therefore, the peak concentration of *E. coli* within individual deposits is one way to characterise the hazard of faecal deposits and can be calculated by multiplying the asymptote of the models and the dry weight of the deposits. For deposits under the spring temperature treatment, dairy cattle contributed the most *E. coli* per deposit (9.35 $\log_{10}$ CFU) followed by sheep (7.80 $\log_{10}$ CFU) with beef cattle contributing the least *E. coli* per deposit (7.06 $\log_{10}$ CFU). For the summer experiment the peak *E. coli* hazard followed the order of dairy cattle (10.29 $\log_{10}$ CFU) > beef cattle (10.00 $\log_{10}$ CFU) > sheep (8.98 $\log_{10}$ CFU). Defecation rates and variable stocking densities would also influence the level of hazard associated with faecal loading of pasture and must also be accounted for when making landscape scale predictions of *E. coli* burden. While this provides a useful concept, national scale inventories of faecal deposit mass by livestock age, faecal deposit *E. coli* concentrations and defecation rates are needed to supplement the data collected here. This would help to make predictions of the relative hazard associated with different livestock faeces more robust.

Such characterisation of microbial hazards can be integrated into existing risk based models of diffuse pollution transfer, for example SCIMAP (Porter *et al.* 2017). Risk-based approaches may be especially useful in the study of catchment microbial dynamics because of the relative lack of understanding on the fate and transfer of FIOs in the landscape compared to other agricultural diffuse pollutants (Oliver *et al.* 2016b). Furthermore, risk-based approaches can often answer the management question regarding what actions to take without the expense of complex process-based models.

4.5 Conclusion

FIO survival at the landscape level is likely to be a key controlling factor on the extent to which river networks become contaminated following rainfall and is a key component of catchment scale predictions of FIO contamination. However, existing catchment scale modelling approaches often assume a simple linear decay function, which does not capture the complexity of *E. coli* persistence in fresh faeces (Coffey *et al.* 2010; Cho *et al.* 2016b). A linear approach is likely to underestimate the burden of *E. coli* in fresh livestock deposits because it does not account for *E. coli* proliferation, under favourable conditions, as observed in field studies (Oliver *et al.* 2010), and now also captured within a CEF mimicking fluctuating environmental conditions. The model developed as part of the current study provides a critical preliminary step towards a framework of accounting for seasonal variations in *E. coli* growth associated with livestock faeces at the catchment scale. However, management practices (for example diet and livestock housing), which vary throughout the year and between farms, are also likely to influence *E. coli* survival. The interaction of agricultural management practices and meteorological variables presents a contemporary challenge for the field of catchment microbial dynamics and further understanding is needed if the risks to ecosystem services related to clean and safe water are to be fully understood and predicted. The analysis presented here will prove beneficial for land managers and make catchment scale predictions of *E. coli* accumulation and persistence on land more robust, accurate and evidence-based, and thus more useful to the policy and decision-making community.

5. Evaluating E. coli mobilisation from fresh faeces of different livestock under simulated rainfall

5.1 Abstract

Pathogens associated with fresh deposits of faeces from grazing livestock have the potential to cause gastrointestinal illness in human beings. Once mobilised by rainfall, pathogens can be transported through the river network exposing sensitive receptors via ecosystem services associated with clean water; for example, shellfish harvesting areas, designated recreational bathing sites and drinking water reservoirs. Measuring changes in concentrations of *E. coli,* a faecal indicator organism (FIO), provides a means of assessing the extent to which water bodies have been contaminated by a faecal source. Release of *E. coli* from a faecal source under rainfall is an important factor controlling its subsequent transport to sensitive aquatic receptors. For the first time, this study measures the release of *E. coli* into overland flow from faecal deposits associated with three different stock types: dairy cattle, beef cattle and sheep. Comparing different stock types is important because it provides insight into which agricultural management practices provide the greatest risk to microbial water quality. To ensure the characteristics of rainfall were consistent throughout the experiment a rainfall simulator was constructed that allowed for calibration of rainfall rate and distribution. This allowed for robust comparisons of the experimental treatments. There was no difference in the profile of *E. coli* release over time from faeces associated with beef cattle, dairy cattle and sheep but there was a greater variation associated with sheep faeces. This finding suggests that variable prevalence and survival of *E. coli* in the faeces of different livestock types may be a greater driver of differences in risk to

microbial water quality compared to variable *E. coli* release under rainfall. However, this initial study requires support from national inventories of *E. coli* release profiles which capture spatial, temporal and between herd/flock variabilities. Despite these data being unavailable at present, the current research can support management of diffuse microbial pollution by providing a first approximation input for tools that predict *E. coli* transport or identify areas of the landscape for mitigation; which will ultimately protect important ecosystem services relying on clean and safe water.

5.2 Introduction

An increase in the demand for food production from existing agricultural land has resulted in the intensification of farming practices. Despite a push toward sustainable intensification of agriculture (Rockstrom *et al.* 2017) increases in production can require an increased use of (in)organic fertilisers and densities of livestock, which is likely to increase the landscape burden of faecally-derived human pathogens. During rainfall these pathogens can become mobilised and transferred to watercourses with subsequent delivery to sensitive receptors such as bathing, shellfish harvesting and drinking waters (Clements *et al.* 2015; Murphy *et al.* 2015; Wu *et al,* 2016) where they may pose a risk of gastrointestinal illness in exposed populations. In order to move toward a sustainable intensification of agriculture, catchment managers must be able to mitigate impacts associated with microbial pollutants which originate from agricultural sources. Pathogens themselves are not routinely monitored because they occur sporadically in the environment and are difficult to enumerate in the laboratory. Faecal indicator organisms (FIOs) such as *E. coli* provide an internationally accepted framework for assessing the extent of watercourse microbial pollution and the development of mitigation strategies requires understanding of the fate and transfer of *E. coli* through catchment systems; the diffuse pollution transfer continuum (Haygarth *et al.* 2005) provides a conceptual framework to aid this understanding. This transfer continuum, when applied to FIO pollution, requires knowledge of the spatial and temporal variation of: (i) FIO source loads, (ii) FIO mobilisation from sources and (iii) subsequent transfer through the landscape. Understanding the mobilisation of *E. coli* from agricultural sources is therefore key to developing mitigation strategies designed to reduce microbial diffuse pollution.

An important agricultural source of *E. coli* is the deposition of fresh faeces from grazing livestock. Fresh faecal deposits can be an especially potent source of *E. coli* because unlike slurry and manure they have undergone no treatment (Chadwick *et al.* 2008). Previous study considering the mobilisation of *E. coli* from faecal deposits has employed a variety of methods including: laboratory investigations, for example rolling of vials containing faeces and a buffer solution (Hodgson *et al.* 2009); observing concentrations in run off from natural rainfall (Tate *et al.* 2000); and simulated rainfall (e.g. Soupir *et al.* 2010; Blaustein *et al.* 2015a). Rainfall simulator design has varied in complexity from drips formed through needles (Moriarty & Gilpin 2014) to the use of spray nozzles with pressure and flow rate gauges to carefully control application rates (Muirhead *et al.* 2005, Fergusson *et al.* 2007, Soupir *et al.* 2010, Blaustein *et al.* 2015a). The use of rainfall simulators allows control of rainfall rates which is impossible in experiments utilising natural rainfall because attributes associated with rain vary greatly between meteorological events. Therefore, controlled rainfall rates from rainfall simulators allow for more robust comparisons of experimental treatments (Davies *et al.* 2004). Studies utilising rainfall simulators have suggested that rainfall is able to mobilise and deliver *E. coli* derived from faeces to the stream network (Collins *et al.* 2005) and have highlighted a number of key factors that determine the rate of mobilisation. For example, while the concentration of *E. coli* in run-off decreases with time (Moriarty and Gilpin 2014) percentage release of *E. coli* from faecal deposits remains consistent (Muirhead *et al.* 2005) showing the importance of the influence of microbial persistence in the contamination of watercourses. Furthermore, depth of rainfall has been found to be a better predictor of *E. coli* mobilisation from faecal matrices than rainfall intensity (Blaustein *et al.* 2016).

Studies utilising rainfall simulators have previously considered mobilisation of *E. coli* from faecal deposits from a single livestock type; for example, Moriarty & Gilpin (2014) investigated ovine faeces and Blaustein *et al.* (2016) investigated dairy cow manure. It is therefore difficult to compare the mobilisation of *E. coli* from the faeces of different animals and to understand variations in impacts to microbial water quality associated with grazing different livestock. Understanding the relative differences in the contribution of *E. coli* to watercourses from deposits of faeces from different livestock can be especially useful in the development of risk based approaches to predicting watercourse microbial contamination. Risk-based approaches are likely to be helpful in the field of catchment microbial dynamics due to some of the current limitations in our understanding regarding the fate and transfer of microbial pollutants relative to other diffuse pollutants such as sediment and nutrient pollution, which potentially limits the parameterisation of fully process-based models (Oliver *et al.* 2016b).

The aim of this study was to develop new understanding on the release of *E. coli* from the faeces of different livestock under rainfall. The objectives were to: (i) develop *E. coli* release profiles from faecal deposits under simulated rainfall; and (ii) investigate whether the profile of *E. coli* release from faecal deposits under simulated rainfall varies between the faeces of dairy cattle, beef cattle and sheep.

5.3 Materials and Methods

5.3.1 Rainfall simulator

The rainfall simulator (shown in figure 5.1) was based on the design of Kibet *et al.* (2014). To ensure a consistent rate of rainfall throughout the experiment a wide angle full cone spray nozzle (TeeJet FL-5VS) was used with flow and pressure gauges for calibration.

Figure 5.1. From the top left moving clockwise: placement of soil boxes underneath a spray nozzle; water pressure and flow gauges; rainfall simulator construction to prevent interference from wind.

Operating at a pressure of 1 bar this gave a flow rate of 1.19 L/min. Figure 5.2 shows the spatial variation in rainfall intensity across the area of the rainfall simulator, placement of soil boxes and the distribution of drop size given by the spray nozzle. Drop size was measured by placing a filter paper in a

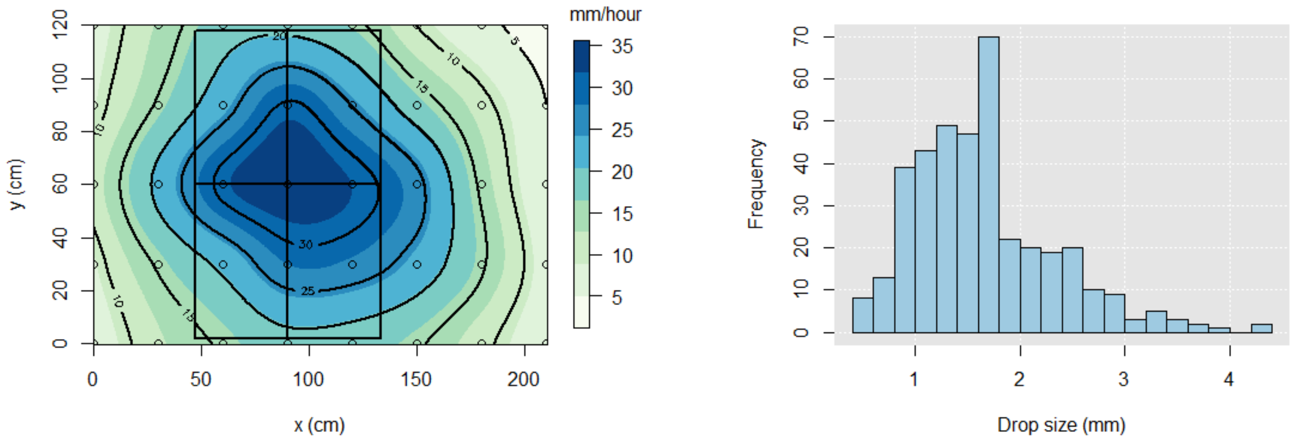

Figure 5.2. The left plot shows the spatial variation of simulated rain within the rainfall simulator; boxes indicate placement of soil-runoff boxes and the circles indicate the placement of rainfall collection pots. The distribution of droplet size is shown in the right hand figure.

petri dish which was momentarily passed through the simulated rain. A photo of the filter paper and a ruler was taken immediately. The photos were used within WebPlotDigitizer (Rohatgi 2018) to measure the diameter of water droplets. The rainfall simulator was enclosed within tarpaulin to prevent the disruption of simulated rain by wind.

The rainfall simulator was fed by mains tap water which was tested to ensure that any constituents of the tap water did not affect *E. coli* survival. Briefly 0.1 ml of an *E. coli* stock solution of known concentration was added to 40ml samples of tap water and deionised water; this was replicated five times for each water type. After refrigeration at 4°C for 24 hours *E. coli* concentration change was assessed by inoculating membrane lactose glucuronide agar (MLGA) plates with a sample from each replicate and incubating inverted for 24 hours at 37°C. Deviance from the expected *E. coli* concentration given the concentration of the stock solution was taken as *E. coli* die off and any difference in deviation between tap water and deionised water taken as a tap water effect. The difference between the two water types was assessed using a t test. There was some uncertainty in the *E. coli* concentration of the stock solution. Therefore, there was uncertainty in the number of cells added to each water sample prior to refrigeration. In order to understand the potential for this uncertainty to lead to a type 2 error (that is accepting there is no difference in *E. coli* die off between the two water types when there is a difference) an expected cell count was drawn from a normal distribution determined by the mean and standard deviation of the cell counts associated with the stock solution. Subsequently the difference in deviance from the expected count and observed count in the two water types was assessed using a t test. Using computer simulation this was repeated 10,000 times with the percentage of significant differences taken as the probability of a type 2 error given the uncertainty in the *E. coli* concentration of the stock solution. For the stock solution, river water was filtered through sterile cellulose acetate membranes which were subsequently placed on MLGA and incubated inverted for 24 hours. Three *E. coli* colonies were transferred to sterilised Luria-Bertani (LB) broth and incubated at 37°C for 24 hours. To clean the cells of LB broth the solution was centrifuged, supernatant removed and pellet re-suspended in sterile phosphate buffer solution. This was repeated three times to remove all residual LB broth. The *E. coli* concentration of the resulting solution was estimated by inoculating MLGA with a small sample and incubating inverted for 24 hours.

5.3.2 Soil boxes

Soil boxes were 50cm long, 35cm wide, 15cm deep, constructed from marine plywood and lined with plastic. An aluminium gutter was placed at the end of each soil box to aid collection of run off. Dilution of the sample by simulated rain in the gutter was prevented by fitting a splash guard over the gutter. Soil boxes were filled with 31.2 kg of A-horizon, mineral gley soil, in layers of 5.2 kg to ensure uniform bulk density. Each layer was compacted with a plywood board and grooved with a hand cultivator tool before a subsequent layers were added. The surface of the soil was left bare as the focus of the experiment was on mobilisation of *E. coli* from faecal deposits and not on microbial transfer over vegetated surfaces. The soil boxes were placed into the rainfall simulator at a slope of 20%.

5.3.3 Faecal deposits

Five intact fresh faecal deposits of each stock type were collected from farms in Stirlingshire, Scotland on the day the experiment took place. Dairy cattle faeces were collected from an area where cattle were held prior to milking. This area was cleaned twice a day so freshness of the collected faeces was ensured. Beef cattle and sheep were grazing on pasture so faeces were collected immediately following deposition.

5.3.4 Experimentation

Immediately before the experiment, soil boxes were placed into the rainfall simulator for 60 minutes to ensure that the moisture content was consistent across all of the soil boxes. Prior to simulating rainfall, faecal deposits were weighed and samples of faeces were collected from each faecal deposit in order to estimate the total *E. coli* burden of the faeces. To enable this, composite samples of each faecal deposit were made up of five 1cm diameter cores for the bovine faeces and 3 faecal pellets for the ovine faeces. Samples were placed into sterile tubes for subsequent microbiological analysis. Faecal deposits were then placed at the top of the tilted soil box. Samples of run-off were collected in sterile pots with the length of time for a sample to be taken recorded; this allowed for the volume of run off being discharged and the *E. coli* load associated with the run-off to be calculated. Rainfall collection pots were placed at the corners of each soil box in order to determine a rainfall application rate for each soil box separately and to confirm no differences in the rainfall rate throughout the experiment.

5.3.5 Microbiological analysis

Samples of faeces were suspended in 9 ml of phosphate buffer solution (PBS) and mixed for 30 mins at 130 rpm. The resulting slurry underwent serial dilution prior to inoculation onto Membrane Lactose Glucuronide agar (MLGA) (CM1031, Oxoid; Basingstoke, UK) via the spread plate and membrane filtration method. Samples of run-off were filtered through sterile cellulose acetate membranes of 0.45 µm pore size (Sartorius Stedim Biotech; Goettingen, Germany) using a filtration unit (Sartorius), prior to the filters being aseptically transferred to MLGA. The plates were then inverted and incubated for 18 to 24 hours at 37°C. All sample analysis was carried out in duplicate and counted colonies represent presumptive *E. coli*.

5.3.6 *E. coli* release modelling

Based on flow rate, *E. coli* counts were converted to a bacterial load and expressed as *E. coli* per minute. Following linear interpolation between time points these data were used to calculate the *E. coli* release as an accumulating percentage of the total *E. coli* load applied to each soil box. Subsequently the potential for a relationship between percentage *E. coli* release and increasing rainfall depth was investigated. Rainfall depth rather than time elapsed was used as the dependant variable because Blaustein *et al.* (2015) suggests a stronger effect of the former on *E. coli* release. To normalise the data and residuals from subsequent modelling it was necessary to carry out a logit transformation on the percentage release of *E. coli*. A scatterplot of these data suggested that there was an asymptotic relationship between logit transformed *E. coli* release and increasing rainfall depth. An asymptotic model was fitted to the data using a mixed effects approach because the experimental design consisted of repeated samples from individual replicates that are not independent and may be serially related (Pinheiro & Bates 1995).

A single mixed effects model was developed with rainfall depth specified as the predictor variable and a unique identifier for each replicate specified as a random effect. To account for potential serial relatedness three temporal autocorrelation structures were trialled: auto regressive order 1; compound symmetry; and autoregressive moving average. The Akaiki Information Criterion (AIC) was used to determine the best performing auto correlation structure with an AIC reduction of ≥ 2 taken as an improvement in model performance. All data processing and statistical analysis was carried out in the R statistics package (R Core Team 2015) utilising a number of third party plugins (Wickham & Francois 2016; Pinheiro *et al.*2015; Neuwirth 2014).

5.4 Results

Die off of *E. coli* in the tap water feeding the rainfall simulator and in de-ionised water did not differ significantly ($p > 0.05$, figure 5.2). The *E. coli* concentration of the stock solution used for this comparison was 7.82 (sd = 0.07) $\log_{10}$ cells / ml. The probability of finding a significant difference between the two water types given the uncertainty in the *E. coli* concentration of the stock solution was 4.3%.

Total deposit *E. coli* content, moisture content and weight of faecal deposits from each stock type are described in table 5.1. *E. coli* concentrations were 2.99 (s = 0.51), 5.99 (s =1 .22) and 4.54 (s = 1.13) $\log_{10}$ CFU g^{-1} (dry weight) for beef cattle, dairy cattle and sheep faeces respectively.

The rainfall simulator constructed for this experiment performed consistently throughout the study. An analysis of variance confirmed that there was no difference in the rainfall rate applied to individual soil boxes ($p > 0.05$). Variation in the discharge of surface run-off from the soil boxes was similar across the treatments (figure 5.3). Discharge increased initially, stabilising after between 15 and 30 mins. An increased discharge of surface run off was observed for two of the replicates used for the dairy cattle faeces treatment.

The pattern of variation in the load of *E. coli* in the surface run off was similar between the stock types while the absolute numbers of *E. coli* load were determined by the concentration of *E. coli* in the fresh faeces (figure 5.5). An initial increase in *E. coli* load was seen over the first two to three time points after which a consistent decline was observed.

Table 5.1. Summary information for faecal deposits; values in parentheses are standard deviations. Model random effect is the deviation of an individual faecal deposit from the population model. Here the standard deviation of the random effects associated with each stock type are shown to demonstrate individual variability associated with different stock types.

Stock type	**Mean total faecal load (log$_{10}$ *E.coli* CFU)**	**Mean faecal deposit weight (g)**	**Mean moisture content (%)**	**Model random effect (standard deviation)**		
				Asymptote (maximum *E. coli* release (%))	**Intercept**	**Shape**
Sheep	6.26 (1.33)	61.99 (37.21)	64.35 (4.79)	2.55	8.17	0.51
Dairy	8.91 (0.22)	866.62 (234.87)	85.30 (1.94)	0.60	0.45	0.28
Beef	6.11 (0.49)	1323.17 (16.40)	85.68 (3.26)	0.61	0.92	0.44

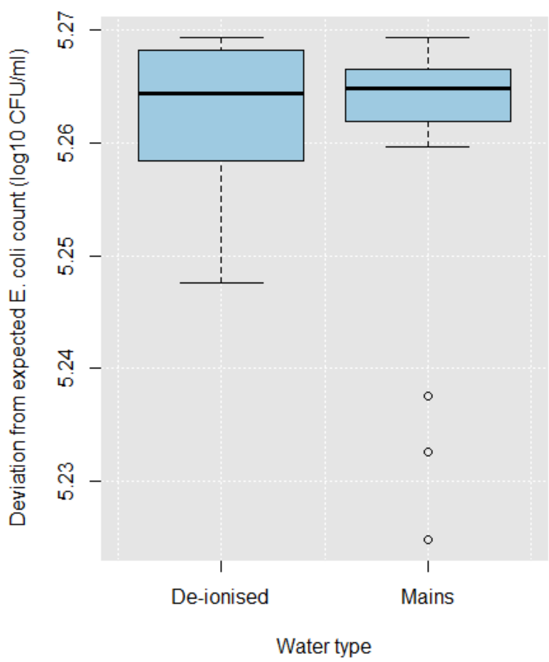

Figure 5.3. Boxplot of *E. coli* die off in the mains water feeding the rainfall simulator and in de-ionised water.

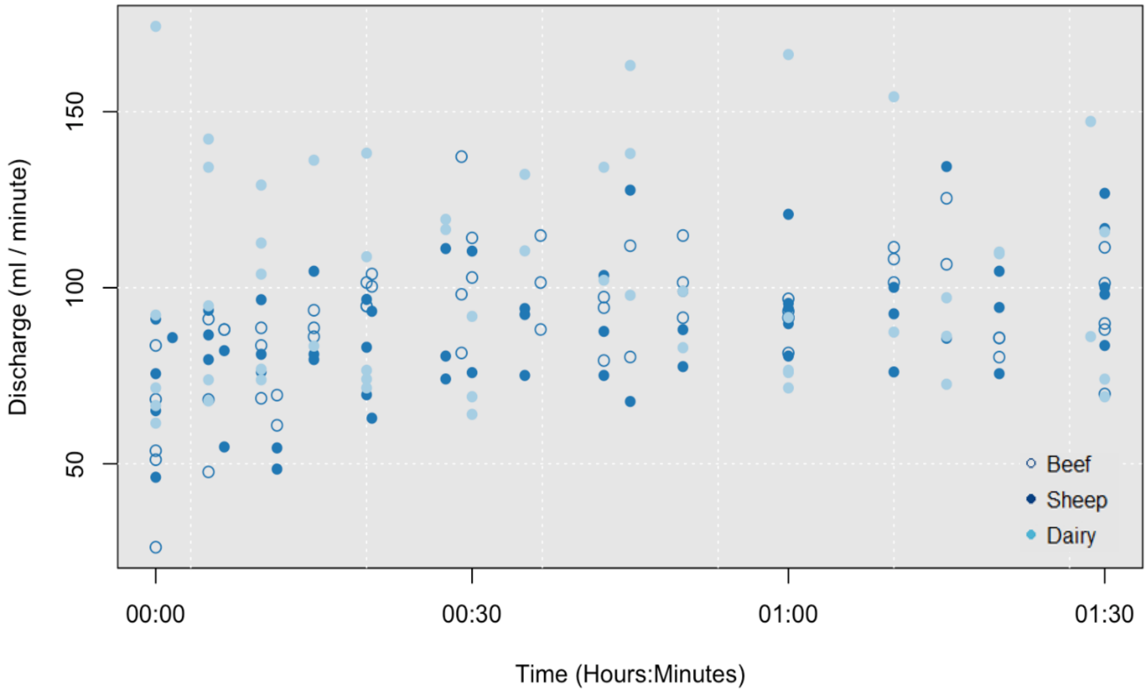

Figure 5.4. Discharge of surface runoff from soil boxes across stock type treatments and through time.

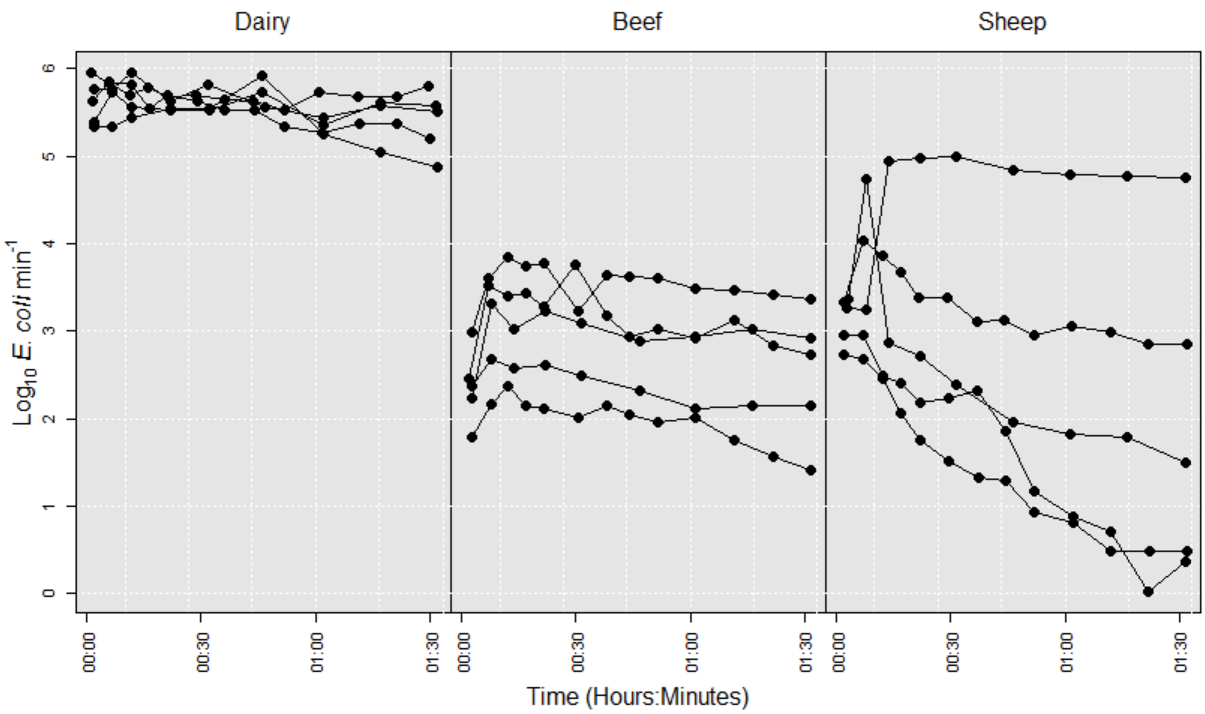

Figure 5.5. *E. coli* load in surface run off from soil boxes through time.

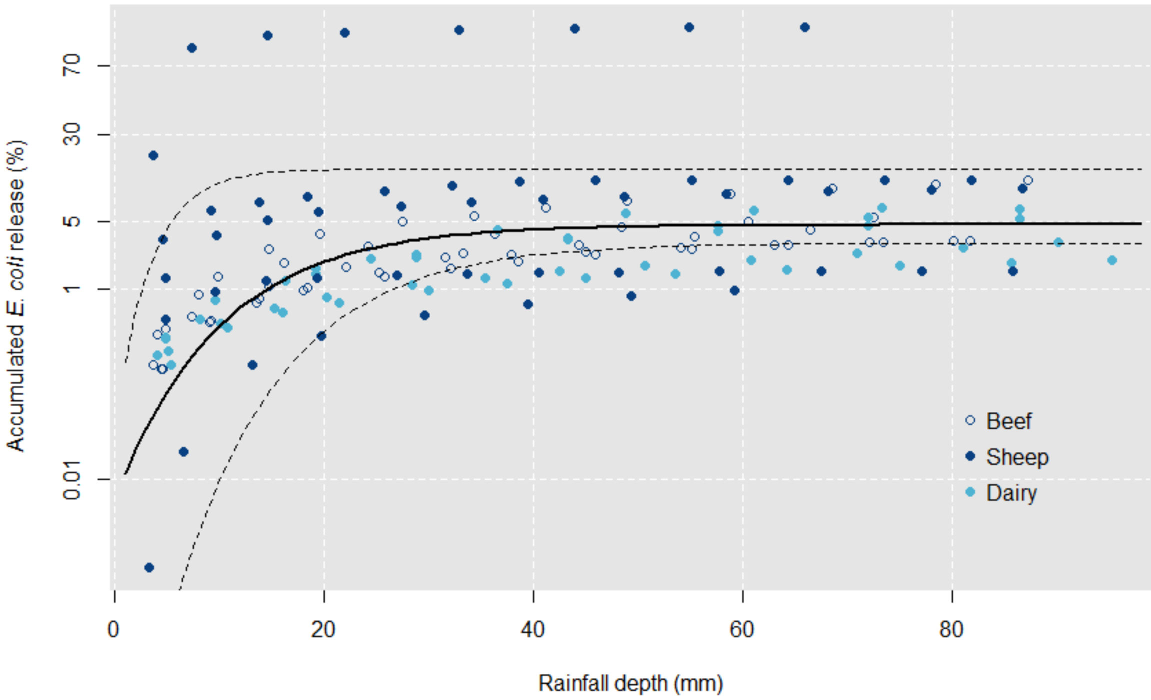

Figure 5.6. Logit transformed *E. coli* release as a percentage of the total faecal burden applied to each soil-run off box. The y axis has been back transformed to enhance interpretability. The solid line illustrates the results of a asymptotic mixed effect model. Dashed lines are 95%, bias-corrected and accelerated (BCa), bootstrap prediction intervals.

The relationship between *E. coli* release and rainfall depth fitted an asymptotic model (Figure 5.5). The asymptotic model suggests a law of diminishing returns with a phase of rapid *E. coli* release for the first 20 mm of rain followed by a slowing of *E. coli* release at higher values of rainfall depth. The model predicted a maximum *E. coli* release of 4.87% associated with deposits from all stock types. For the logit transformed data, the predicted value of the model parameters were -2.97 (-3.48,-1.69), -9.60 (-19.10, -7.99), -1.93 (-2.31, -1.22) for the asymptote, intercept and shape parameter respectively (parentheses show 95% BCa confidence intervals). A plot of the autocorrelation function suggested some potential for temporal autocorrelation in the model. The best performing temporal autocorrelation structure was auto-regressive moving average (zero auto-regressive parameters and four moving average parameters) (AIC reduced by 134.63).

A model differentiating between stock types did not converge on a solution. However, a dot plot of the random effects against stock type (figure 5.7) shows a potential difference in the shape parameter between stock types. This suggests that the rainfall depth at which the maximum *E. coli* release is reached may differ between stock types. The dot plot also illustrates variability in the model parameters associated with each individual deposit. There was increased variability in all three parameters associated with sheep faeces; the standard deviation of the random effect is shown in table 5.1.

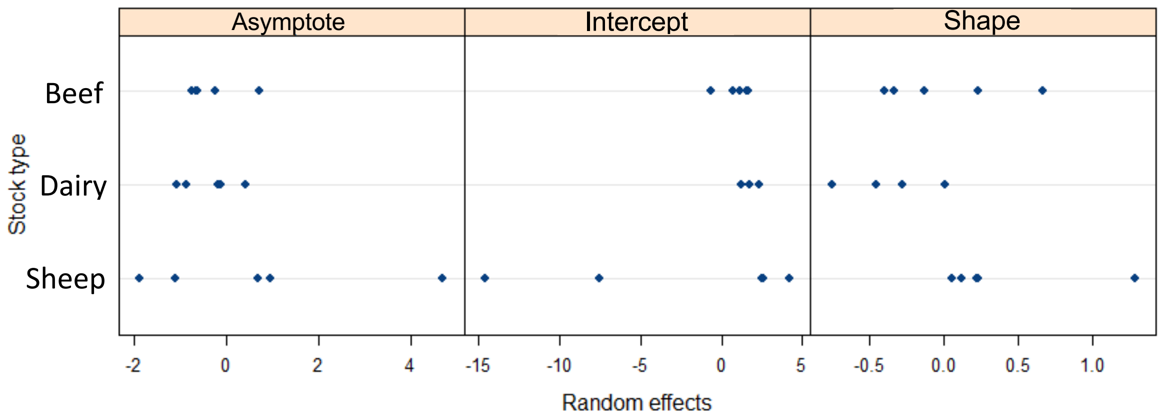

Figure 5.7. Dot plot of model random effects (deviation of individual faecal deposits from the population model) against stock type.

5.5 Discussion

In this study *E. coli* release from faeces associated with three different livestock types has been profiled in parallel and under controlled conditions for the first time. *E. coli* release from faecal matrices associated with sheep, beef cattle and dairy cattle showed a two stage release profile, which is consistent with existing observations associated with dairy cattle faeces (for example Blaustein *et al.* 2016; Ferguson *et al.* 2007). Initially the release of *E. coli* was rapid, with cell emergence slowing as rainfall continued. This suggests an initial flushing of easily mobilised *E. coli* that may exist as freely living organisms on the surface of the faecal deposit and has been described previously (Tate *et al.* 2000). As the supply of cells that are easily mobilised depletes *E. coli* emergence slows. The observed deceleration in the release of *E. coli* from the faeces may be due to better protection of cells within deeper layers of the faecal matrix, thus limiting the effectiveness of raindrop impact and associated detachment processes on the remaining *E. coli* population. The strength of association of *E. coli* with manure particles is likely to be important here, unless the raindrops physically dislodge organic material that the cells are attached to. The attachment of *E. coli* to particles may reduce the speed at which cells are transported over the soil surface and into the collection gutter further reducing the speed at which cells emerge in run-off as well as allowing more time for the overland flow to be diluted by rainfall (Blaustein *et al.* 2016).

There is limited experimental data considering the transport state of *E. coli* in run off. *E. coli* can be transported as free living cells, flocs of multiple cells or attached to soil and manure particles (Muirhead *et al.* 2005). Evidence suggests that >25% of the *E. coli* in overland flow is transported as free living cells (Muirhead *et al.*

2005). However, attachment rates are reported as averages in the run off of a rain event. The two stage release profile here may suggest that the attachment of *E. coli* varies throughout individual rainfall events. In the later stages of a rainfall event, as the source of readily mobilised cells depletes, the much slower mobilisation of cells attached to particles and flocs of cells may make up a greater proportion of the *E. coli* in run off. The distinction between attached and freely living cells is important because cells that exist as flocs or attached to particles are more likely to be filtered by vegetation and soil matrices (Fiener & Auerswald 2003). Therefore, attached cells that are released toward the end of rainfall events may present a reduced hazard to watercourse microbial quality compared to free living cells released at the beginning of an event. Furthermore, although rainfall depth rather than intensity has been highlighted as a useful predictor of *E. coli* mobilisation, increased rainfall intensity may increase flow velocity which has been shown to decrease cell attachment (Guber *et al.* 2005) and potentially increase downstream microbial hazards because a greater proportion of the *E. coli* load is made up of highly mobile free living organisms. Understanding variabilities of *E. coli* attachment to particles and other bacteria within individual rainfall events should, therefore, be a focus of future investigation.

The percentage of total *E. coli* released from livestock faeces under rainfall has been shown to be very small (<5%) (Moriarty & Gilpin, 2014; Ferguson et al. 2007). However, the *E. coli* load of faecal deposits on pasture is large (up to ~ 8 $\log_{10}$ CFU g^{-1} dry faeces) (Oliver *et al.* 2016a; Moriarty *et al.* 2011a) so a relatively small percentage release can create significant hazards for microbial water quality when mobilised cells are successfully delivered to a receiving water. The small release also suggests that, despite rainfall, *E. coli* removal is relatively minor and faecal deposits on pasture can persist as a significant source of *E. coli* after rainfall events.

The study presented here focussed on *E. coli* release from fresh deposits during a single rainfall event. Under subsequent rainfall events it is unclear what profile *E. coli* release may take. A potential theory is that during subsequent rainfall events, *E. coli* release continues through the two-stage profile where the last rainfall event finished. Alternatively, the release profile may 'reset' with *E. coli* mobilisation during subsequent rainfall events accommodating the same two-stage profile illustrated here for fresh deposits. Rainfall simulation studies operating at larger scales show time since last grazing as a significant factor in determining *E. coli* loads and concentrations in rainfall induced overland flow. This may be attributed to cell wash out during rainfall before the experiment, perhaps, supporting the prior theory. However, *E. coli* die off is also likely to contribute to this effect (Collins *et al.* 2005). Given the high concentration of *E. coli* remaining in faecal deposits after rainfall the effect of multiple rainfall events on the mobilisation of *E. coli* at the individual cow pat scale should be a focus of further study.

The age a faecal deposit receives its first rainfall may influence *E. coli* release. For example, drying of faeces has been shown to reduce the rate at which *E. coli* is mobilised from beef cattle and sheep faeces (Hodgson *et al.* 2009). Alternatively, UV may inactivate easily released cells existing on the surface of a faecal deposit. However, there is little variation in *E. coli* release between fresh bovine faeces and week old bovine faeces (Ferguson *et al.* 2007); fresh and 30 day old dairy cattle faeces (Muirhead *et al.* 2005) and between sheep faeces aged 0, 1, 4, 7, 14 and 21 days (Moriarty & Gilpin 2014). Despite little evidence of total *E. coli* mobilised varying with age up to 30 days, it has been suggested that an aged faecal deposit may require rehydration before releasing *E. coli* which could lead to a delay in the emergence of cells in run off (Moriarty & Gilpin 2014) producing a different profile of *E. coli* release.

Repacked soil boxes may not be an accurate representation of soil conditions in the environment because the soil structure has been compromised. The focus of this study was on a single part of the diffuse pollution transfer continuum, mobilisation, and the transfer and attenuation of *E. coli* through soils and in overland flow was not a focus of this experiment. However, in order to make valid comparisons it was important to ensure that the soil conditions were consistent throughout the experiment and the use of repacked soil boxes as a medium on which to place faecal deposits allowed for control of soil conditions. Similar justification is given to the use of a rainfall simulator in place of natural rainfall. Natural rainfall is highly variable, and the attributes associated with rain during one event are unlikely to be similar to subsequent events. A rainfall simulator ensured that rainfall conditions were consistent throughout the experiment (Davies *et al.* 2004), thus making comparisons between treatments more robust.

Similar rainfall simulation studies have partitioned surface run off and leachate (Blaustein *et al.* 2016). In the present study we chose to focus on surface run off because this rapidly responding pathway has been highlighted as a key contributor to high *E. coli* concentrations in river samples following rainfall (Collins *et al.* 2005). Furthermore, soil has been shown to act as an efficient filter of pathogens (Morales *et al.* 2014) which may reduce the relative contribution of subsurface flow pathways to microbial watercourse pollution. However, it is important to note that at the hillslope scale, field drainage might increase the importance of subsurface pathways (Oliver *et al.* 2005).

An aim of this study was to determine how hazards to microbial water quality vary with livestock type. There were no contrasts in the release of *E. coli* from the faeces of the three livestock types investigated for the particular rainfall intensity studied. However, values of the model parameters for individual faecal deposits suggest that there may be a small difference in the shape parameter between livestock types. Based on figure 5.6, sheep faeces reaches the maximum *E. coli* release at smallest values of rainfall depth followed by beef cattle faeces and finally dairy cattle faeces. However, the between stock type variability was not strong enough to distinguish predictive models specific to livestock type. The lack of a clear difference between livestock types is in contrast to previous literature which suggest *E. coli* are mobilised more readily from beef cattle faeces than sheep faeces (Hodgson *et al.* 2009). However, the previous laboratory study simulated raindrop interactions by rotating a mixture of rain water and faeces in a vial which may not be representative of rain/faeces interactions in the environment. No difference in the mobilisation of *E. coli* between livestock types in the present study suggests that relative hazards to microbial water quality associated with grazing different livestock are likely to be driven by the variability in the prevalence and survival of *E. coli* in faecal deposits from different livestock (Oliver *et al.* 2016a; Moriarty *et al.* 2011b). However, the extent to which *E. coli* associated with deposits from different stock types attach to faecal colloids and soil particles may influence variabilities in microbial water quality hazards associated with grazing different livestock types.

The profile of *E. coli* release was similar between all livestock types however there was a much greater variability in the release of *E. coli* associated with sheep faeces. Most of this variation is attributed to the asymptote and intercept parameter of the predicted model with one outlier showing a percentage release of up to 80%. Further data would need to be collected to confirm whether this variability is genuine or as a result of a single outlier. However, even discounting this single outlier, mobilisation of *E. coli* from sheep faeces was variable. This provides a challenge for incorporating this new knowledge into catchment scale tools because the uncertainty will be propagated through the catchment continuum contributing to large confidence intervals associated with predictions of downstream *E. coli* concentrations.

Decisions about diffuse pollution management are often made at the catchment scale so extrapolation of findings at smaller scales is required before these findings can be useful to water quality stakeholders. Our observations have potential to inform process-based models of diffuse pollution transfer that aim to predict downstream concentrations of *E. coli*; for example, Soil and Water Assessment Tool (SWAT) (Cho *et al.* 2016) and Hydrologic Simulation Program in FORTRAN (HSPF) (Pandey *et al.* 2016) (see Cho *et al.* (2016) for a full list). The findings may also inform risk-based approaches, for example SCIMAP (Porter *et al.* 2017), that aim to target locations for management based on the likelihood of part of the landscape contributing *E. coli* to watercourses. However, the data presented here and in similar studies (Blaustein *et al.* 2016) provide only an indication of the release profile of microbial contaminants from faecal sources and national-scale inventories of *E. coli* release profiles from multiple herds/flocks and locations is required before these release kinetics can be robustly captured within decision making tools that operate at larger scales. In addition, it is unknown what uncertainty is incorporated into large scale models when knowledge from small scales (individual faecal deposit or field-scale) is scaled up to the catchment scale (Cho *et al.* 2016b).

5.6 Conclusion

For the first time *E. coli* release into overland flow has been profiled for sheep, beef cattle and dairy cattle faeces in parallel and under controlled conditions. This allowed for an assessment of the relative contributions that different livestock make to hazards associated with microbial water quality. This first assessment revealed that *E. coli* release profiles of the three stock types studied are similar, suggesting that differences in microbial water quality hazard come primarily from variable *E. coli* prevalence and persistence in the faeces of different livestock; although further research investigating different rates of rainfall is required to support this finding. The results presented here can be used to inform models operating at larger scales however a much larger inventory of *E. coli* release profiles is needed to determine spatial, temporal, between herd/flock and within herd/flock variabilities and make the use of the knowledge regarding *E. coli* release from faecal deposits within catchment scale tools more robust. Despite current limitations associated with data availability this study has provided valuable new knowledge on the fate and transfer of *E. coli* originating from the faeces of different livestock types and can inform further study and provide starting points for models predicting watercourse *E. coli* contamination.

6. Synthesis: using new knowledge to adapt SCIMAP for diffuse microbial pollution; next steps and challenges for the field of catchment microbial dynamics

6.1 Opportunities for development

The Sensitive Catchment Integrated Mapping Analysis Platform (SCIMAP) has been widely accepted in the fields of diffuse sediment and nutrient pollution (Reaney *et al.* 2012). It is therefore timely that it should be further developed to account for diffuse FIO pollution. A core aim of this project has been to assess SCIMAP's suitability as a tool for mapping diffuse FIO pollution risk and identify and act upon opportunities for improvement. The SCIMAP approach has its foundations in the source, mobilisation, delivery, impact (SMDI) conceptualisation of diffuse pollution transfer. This approach describes how a source of pollution is only converted to an impact if it can be released from its source and transferred to a sensitive receptor (Haygarth *et al.* 2005). Therefore, in the language of risk assessment, the source is akin to a hazard with mobilisation and delivery determining the likelihood of a sensitive receptor becoming exposed to the hazard. The SMDI continuum has provided a narrative for the adaptation of SCIMAP to diffuse FIO pollution with investigations into how its treatment of source, mobilisation and delivery should be modified to better represent this novel pollutant.

SCIMAP is at its most powerful when it is combined with the SCIMAP fitted approach demonstrated in chapter 3 which trains the land cover risk weightings to represent the spatial pattern of sources within an individual catchment. However, this approach requires a spatial dataset of information on FIO contamination levels that is distributed both spatially and temporally (Porter et al. 2017). These datasets need to capture the variability in the mosaic of landcovers across a catchment and also account for seasonal variations in environmental variables such as temperature (Martinez *et al.* 2013) and rainfall (Blaustein *et al.* 2015b) that influence FIO survival and transfer, and land and stock managements (Oliver *et al.*2016b). A requirement for this level of information may provide a barrier to the use of SCIMAP at a wider scale for predicting FIO risk because these datasets are difficult and expensive to develop.

A set of default land cover risk weightings can be developed for situations where catchment scale FIO information is not available. The evidence gathered in chapters 4 and 5 can be exploited to create a set of default land cover risk weightings which are required for the SCIMAP approach. An initial investigation into the performance of SCIMAP when applied to diffuse FIO pollution highlighted opportunities for improvement in the representation of the source and mobilisation phases of the SMDI continuum. At present SCIMAP's treatment of diffuse pollution transport is time-integrated because diffuse pollution management strategies are often permanent features that once installed cannot be moved and should therefore be placed to optimise pollution management throughout the year (Porter *et al.* 2017). However, a time-integrated approach to predicting diffuse FIO pollution may not be appropriate because, unlike conservative sediment and nutrient pollution, once introduced into the environment FIOs may proliferate or die off. It is widely recognised that meteorological variables that vary throughout the year impact the survival and persistence of FIOs in the landscape (Oliver *et al.* 2016a) and the resulting temporal variations in pollution risk should be characterised. This problem is especially pertinent in the case of designated recreational bathing sites where a sensitive receptor is only present during part of the year. Therefore, SCIMAP’s treatment of FIO source hazard may be improved by capturing temporal variations in the FIO loads of faecal sources that are a result of variable rates of FIO concentrations in fresh faeces and post defecation FIO growth. Furthermore, SCIMAP can better predict temporal variations in FIO risk if it captures temporal variations in mobilisation of FIOs from sources under rates of rainfall that vary throughout the year.

Another opportunity for development highlighted after an initial investigation of SCIMAP’s performance was that a key part of SCIMAP relies on land cover classifications to assign source hazard weightings to parts of the landscape (Porter et al., 2017). Two such classifications were improved pasture and rough grazing. Grazing different animals at varying densities will result in these land cover classifications covering a wide range of FIO availabilities due to the variable prevalence, persistence (source) and release (mobilisation) of FIOs associated with faeces from different animals. This creates uncertainty in the risk weighting that should be assigned to a land cover type (Milledge *et al.* 2012). In the case of diffuse FIO pollution, spatial variation in source hazard may be better approximated by understanding spatial variations in the types and densities of livestock that are contributing FIOs to the landscape.

6.2 Spatial and temporal variations in sources of microbial hazards

The chapter ‘*high resolution characterisation of E. coli proliferation profiles in sheep, dairy cattle and beef cattle faeces*’ was designed to gather information that would inform the development of SCIMAP to better account for temporal and between stock type variations in source hazard. Two elements that contribute to a faecal deposit's overall hazard are: the starting concentration of *E. coli* in fresh faeces and the extent (magnitude and duration) to which *E. coli* can grow within the faecal matrix. Overall source hazard weightings for faecal deposits from different livestock at different times of year that capture both aspects of hazard can be derived by ranking deposits by initial concentration and plotting against a corresponding rank of the magnitude of *E. coli* growth (figure 6.1). The overall risk value is then the distance from the origin rescaled 0 to 1. Parameters from the model developed in the current study were used to inform this, with the intercept used to describe initial concentration and asymptote – intercept taken as the magnitude of *E. coli* growth. Where no *E. coli* growth occurred, the initial concentration was taken as the intercept and growth specified as 0. Capturing the influence of spatial variation in stocking density can be achieved by combining these new hazard weightings with agricultural census data. This spatial information can easily be incorporated into SCIMAP as it exists by replacing the land cover hazard weightings.

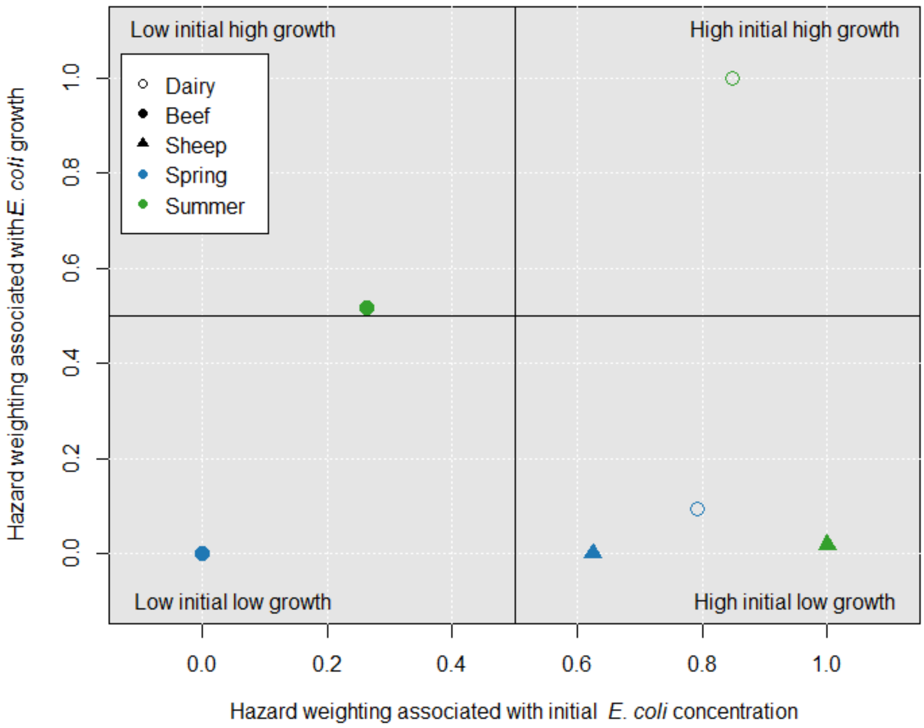

Figure 6.1. Scatter plot illustrating how initial concentration of *E. coli* and the extent to which *E. coli* can grow in a feacal deposit contributes to an overall source hazard weighting. Stock season combinations occuring in the top right of the plot will have the highest source hazard; stock season combinations at the bottom left of the plot have the lowest source hazard weighting. An overall source hazard weighting is therefore the distance form the origin.

6.3 Temporal variations in mobilisation hazard associated with seasonal variations in rainfall

The SMDI continuum suggests that a source of hazard must be mobilised and delivered before a sensitive receptor can become exposed to the hazard. Therefore, capturing variations in FIO mobilisation that result from seasonally variable rates of rainfall may improve SCIMAP's consideration of diffuse FIO pollution. Chapter 5 was designed to investigate the release of *E. coli* from faeces with increasing rainfall depth. The relationship derived from this study can be combined with routinely collected rainfall information to predict the extent to which microbial sources release *E. coli* in different months.

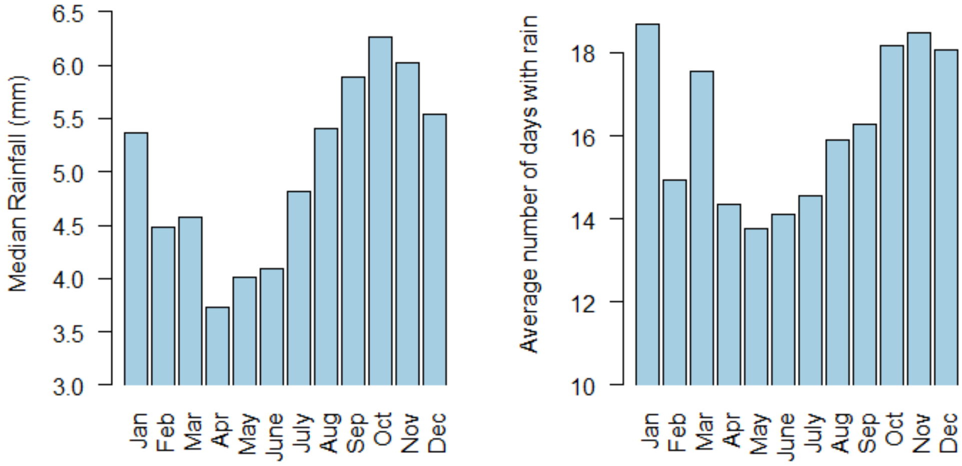

Figure 6.2. From left to right: bar plot of median daily rainfall amount; bar plot of average number of days with rain. Met Office rainfall data for the period 1961 to 2009.

The hazard associated with a faecal deposit releasing *E. coli* depends on the relative likelihood of rain occurring on a day in a given month compared to all other months; and the average daily rainfall during a given month compared to all other months. Rainfall data were acquired from the Met Office MIDAS dataset 'UK daily rainfall data'. Data for the period 1961 to 2009 were downloaded and for each month the average number of days with rain and median rainfall amount (median were used because the data were skewed to the left) was calculated (Figure 6.2.).

The relationship between rainfall amount and *E. coli* release developed in chapter 5 was used to predict the percentage of a deposit's *E. coli* that may be released under rainfall in different months given the average daily rainfall of that month. For example, daily rainfall for a month is taken and read from the X axis of figure 5.6; percentage of *E. coli* release is then shown on the associated Y axis. These data were then rescaled 0 to 1 with values of 1 illustrating most hazardous months and 0 least hazardous. Developing an overall mobilisation hazard is then similar to the treatment of source hazard.

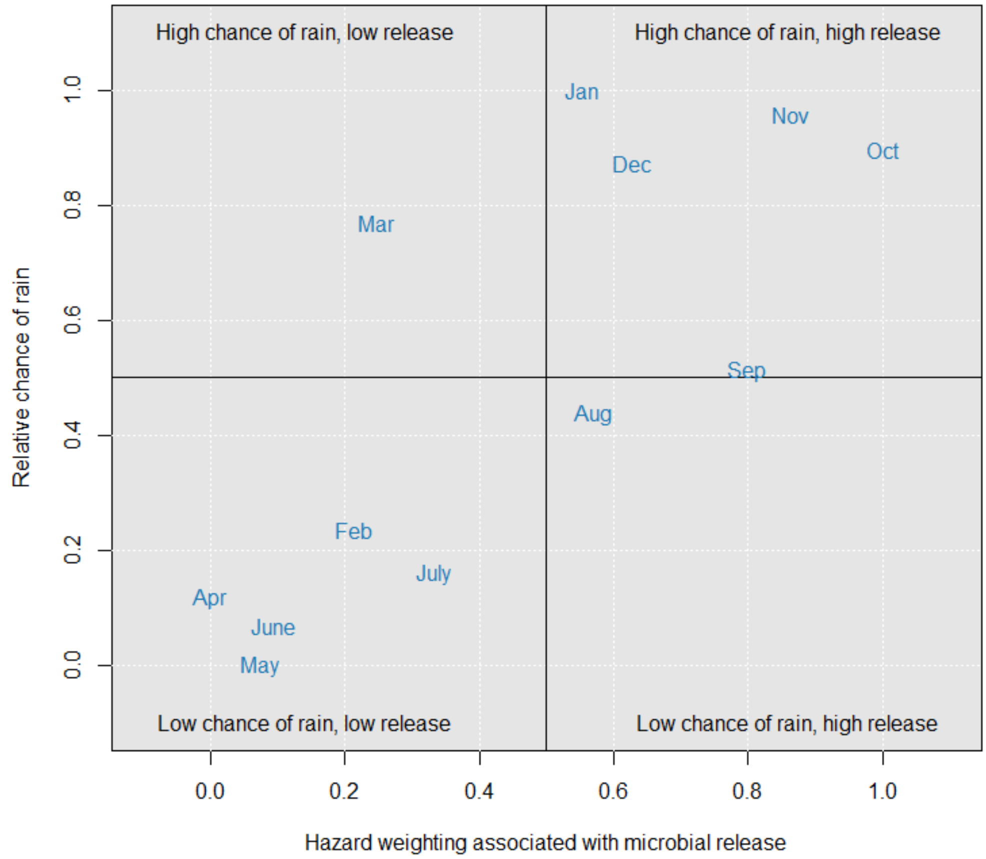

Figure 6.3. Scatter plot of a hazard weighting reflecting the relative likelihood of rain on a day in a given month against a hazard weighting associated with the amount of rain that occurs on a single day in a given month. A hazard weight that reflects both theses aspects of mobilisation hazard is then the distance from the origin.

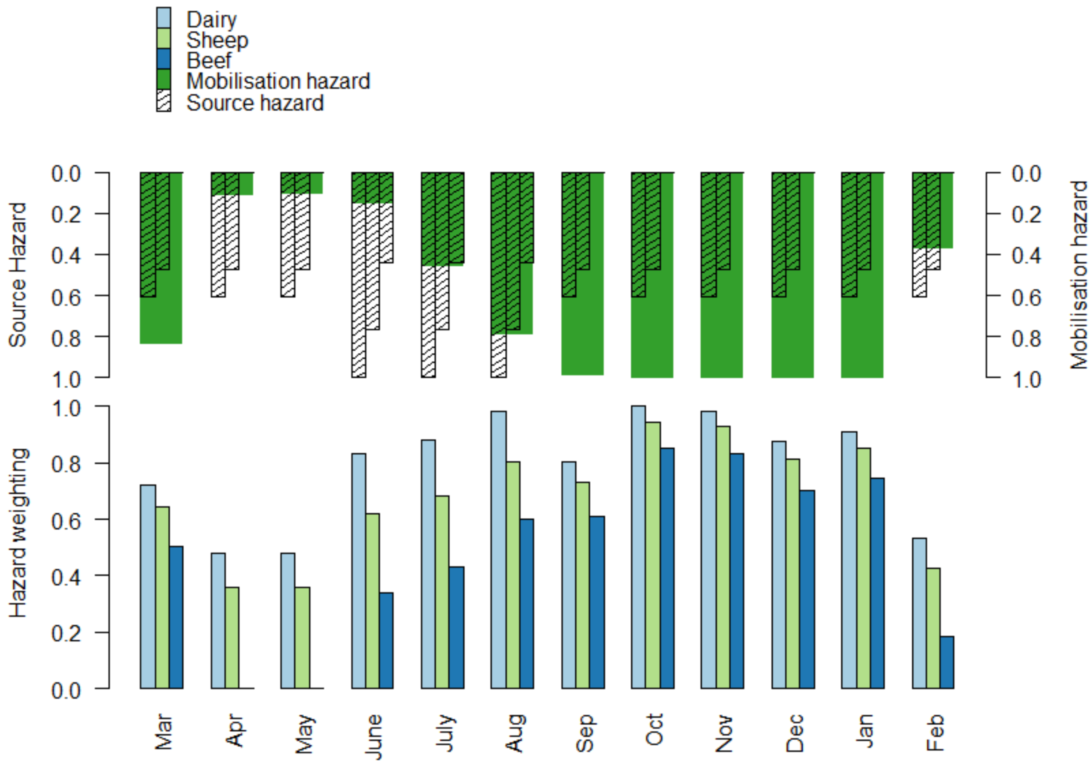

Figure 6.4. The lower bar plot illustrates stock type and month differentiated hazard weightings that capture aspects of hazard associated with the concentration of *E. coli* in fresh faeces; the extent to which *E. coli* grows in faeces; the relative chance of rain on a day of a given month; and the amount of rain that falls in a single day of a given month. The upper bar plot shows how source hazard and mobilisation hazard contribute to the overall hazard weight in different months.

The number of days with rain in each month is rescaled 0 to 1 and plotted against the *E. coli* release hazard weightings. The distance from the origin is then the overall mobilisation hazard due to rainfall on different months (figure 6.3).

6.4 Combining source and mobilisation hazards

A plot of mobilisation hazard against source hazard would illustrate a new hazard weighting that encompasses both source and mobilisation terms. The source and mobilisation encompassing hazard weighting would be the distance from the origin of this plot. Therefore, the hazard posed by faeces from a given stock type and month could be calculated as:

$$\sqrt{A^2 + B^2 + C^2 + D^2}$$

Where *A* is the rank of initial *E. coli* concentration; *B* is the rank of *E. coli* growth; *C* is the rank of number of days with rain; *D* is the rank of rain amount. Figure 6.4 illustrates the stock type/month differentiated source/mobilisation hazard weightings based on the data collected within this thesis. In this example two seasons of high risk can be identified. There is a higher risk of *E. coli* transfer associated with months June, July and August. This risk is driven by greater source hazard; there are likely to be warmer temperatures in these months driving a greater magnitude of *E. coli* growth. The second period of high risk occurs September to January and this risk is driven, largely, by increased chance of rainfall. This highlights the importance of considering interacting meteorological factors when assessing the risk of *E. coli* transfer to watercourses. The weightings developed using this approach can be combined with agricultural census information to develop spatial information that can replace the land cover weightings presently used in SCIMAP. This adaptation of the way SCIMAP treats sources and mobilisation of diffuse FIO pollution may address the challenges associated with temporal and between stock type variations in *E. coli* source and likelihood of mobilisation outlined in chapter 3: '*Predicting diffuse microbial* *pollution risk across catchments: the performance of SCIMAP and recommendations for future development*'.

While the method outlined provides a proof of concept, the workflow needs to capture information from additional existing research. Information capturing variations in hazard from a variety of locations, animals and management practices should be gathered. This consolidation of existing research should be carried out in a strictly quantitative manner using contemporary meta-analysis techniques like those described by Gurevitch *et al.* (2018) and is unfortunately not within the scope of this project.

The development of a default set of FIO risk weightings for the SCIMAP framework allows for the mapping of FIO risk in catchments that are not associated with large amounts of FIO data. However, there will be more uncertainty associated with these risk weightings than those developed using the SCIMAP fitted method and distributed datasets. Collection of FIO data is therefore a priority for managing diffuse FIO pollution in sensitive catchments. Currently available data collection techniques provide a significant barrier in this respect. The field of catchment microbial dynamics is still largely confined to the use of culture-based methods which are relatively expensive (compared to nutrient analyses), labour intensive and time consuming; these issues prevent the collection of the amount of data required for testing and specifying models that operate at large spatial and temporal scales.

Within the field of diffuse sediment and nutrient pollution, technologies that acquire near real-time information have been developed which allows for the efficient, and often remote, collection of large quantities of data (Owen *et al.* 2012). This allows for robust training and testing of catchment scale models of pollution transfer. The development of new technologies that are able to measure FIO concentrations remotely and near real-time should be an ambition for catchment microbial dynamics. Devices such as ColiMinder (VWM Solutions) provide opportunities for data collection but to this author's knowledge has not been applied to spatially distributed sample collection for catchment-scale model development and specification. These systems exploit the presence of a metabolite β-glucuronidase as an indicator of faecal pollution and has an advantage over culture-based methods in that viable but non-culturable *E. coli* are captured (Joensen *et al.* 2014).

The SCIMAP framework provides opportunities for mapping diffuse pollution risk and planning the deployment of mitigation measures, and has been the focus of this thesis. However, SCIMAP's use case is relatively specific and concerns the targeting of areas for diffuse pollution mitigation. The information gathered within this thesis could have wider impact in that it can be useful for developing other modelling approaches with different objectives; for example, predicting *E.* coli contamination in response to rainfall (SWAT (Cho *et al.* 2012)) or risk of health implications following contamination of watercourses (QMRA (McBride *et al.* 2013)).

The information presented throughout the current work could also provide opportunities for developing an Agent Based Modelling approach (Reaney 2008). Agent based modelling (ABM) is yet to be applied within the field of catchment microbial dynamics but could be useful for modelling the fate, transfer and impact of diffuse microbial contaminants. ABMs describe a system via a group or multiple groups of entities (agents) and an environment. Simple rules then describe how these agents interact with each other and their environment. Through simulation, agents are allowed to interact with other agents and their environment with a global level effect emerging as a result. For describing complex systems this approach has significant benefits over traditional mathematical modelling approaches which are limited by mathematical tractability. This simulation-based approach is also useful for understanding uncertainty in highly stochastic systems.

In the context of catchment microbial dynamics, agents could be livestock, faecal deposits and/or parcels of contaminated run off. Environment variables might include topographic, land use, stream network and meteorological information. Experiments like those presented in the current work which are developed under the source, mobilisation, delivery, impact framework can then inform rules determining how agents interact with each other and their environment to result in downstream impacts to ecosystem services that rely on clean water.

7. Conclusion:

Faecal contamination of watercourses has been reduced through engineering solutions associated with point sources of contamination. As point sources have been addressed the relative importance of diffuse sources from the landscape have become more important. The distributed nature of diffuse sources of FIO pollution makes it difficult to determine where sources are, and which sources contribute most to microbial water quality impacts. The objective of the work described in this thesis was to develop a solution to the problem of determining where diffuse FIO pollution originates and where mitigation measures should be deployed for the greatest improvement in water quality. SCIMAP was identified as a tool which shows promise, and prior to this work, had not been considered for application to diffuse FIO pollution. The research here has shown that:

- SCIMAP as it exists has limitations in its application to diffuse FIO pollution in that:
    - It's approximation of source risk using landcover information does not appropriately reflect the spatial and temporal distribution of FIO source risks;
    - SCIMAP's time integrated approach may not be appropriate for a non-conservative pollutant.
- *E. coli* populations are able to grow in livestock faeces for up to 30 days post defecation.
- The extent of *E. coli* growth is influenced by temperature conditions that vary seasonally.

- The release of *E. coli* from faecal matrices exposed to simulated rainfall is relatively small and can be approximated using rainfall depth. No livestock type variation was discovered.

These findings were consolidated in a proposed workflow that outputs a set of default hazard weightings that can be applied within the SCIMAP framework. Below, these findings are described in more detail.

Assessment of SCIMAP's performance when applied to diffuse FIO pollution:

SCIMAP, which has been optimised for diffuse sediment and nutrient pollution, has been highlighted as a tool with potential to inform the spatial targeting of diffuse FIO mitigation pollution. However, an initial investigation into its performance highlighted a number of weaknesses: (i) SCIMAP's treatment of source hazard does not account for the variability in FIO availability that arises from grazing different livestock at varying densities; (ii) the existing 'time integrated' approach does not capture seasonal variations in the proliferation, survival and release of *E. coli* in/from faecal matrices.

Post defecation proliferation of E. coli in livestock faeces:

Chapter 3 demonstrated how the growth of *E. coli,* often observed in field trials, can be replicated under controlled conditions. This shows experiments conducted under controlled but environmentally relevant conditions can inform understanding on the fate and transfer of *E. coli* in the environment. Furthermore this experiment showed an increased potential for *E. coli* growth in livestock faeces during warmer seasons in the U.K. These results informed the development of seasonally differentiated hazard weightings that can inform the SCIMAP approach.

Mobilisation of E.coli under simulated rainfall:

In chapter 4, the release of *E. coli* from faecal deposits was investigated. It was demonstrated that the release of *E. coli* can be approximated using accumulated depth of rainfall. This investigation showed a law of diminishing returns with release of *E. coli* becoming more difficult as rainfall continued. This experiment also suggested no difference in the release of *E.coli* between three different livestock types. The model of *E. coli* release developed here was combined with routinely collected rainfall information to create monthly differentiated hazard weightings for grazing livestock on pasture.

Integrating new knowledge within the SCIMAP approach:

An initial investigation into the performance of SCIMAP when applied to diffuse FIO pollution highlighted a key weakness in that the approach did not consider the highly seasonal nature of diffuse FIO pollution or the variation in FIO source loading due to grazing different livestock types. The experiments conducted subsequently aimed to inform an adapted SCIMAP approach that may better account for the spatial and temporal variation in diffuse FIO pollution. Rates of *E. coli* growth varied with seasonal variations in temperature and also between livestock types; rate of *E. coli* release varied with rainfall that changes monthly. A workflow for capturing these variations within SCIMAP was developed.

Overall the research here has developed an approach for applying SCIMAP to diffuse FIO pollution. The implementation of these recommendations would provide a tool for those who aim to reduce FIO contamination through the mitigation of diffuse inputs from agricultural landscapes. Furthermore, this research has shown the usefulness of controlled experiments to inform modelling approaches operating at large scales. Field studies will remain vital in improving knowledge of the fate and transfer of FIOs in the environment; however, variable and interacting factors making it difficult to disentangle the effects of individual drivers of diffuse pollution. The experiments investigating the proliferation of *E. coli* in faeces within a CEF and the release of *E. coli* from faeces under simulated rain demonstrate how new knowledge derived from controlled experiments can inform decision-support tools operating at field scales. While such extrapolations should be treated with caution and tested against environmentally derived data, such approaches are required for understanding how faecal pollution reacts to varying meteorological conditions and agricultural management interventions, which in turn allows for the development of operationally useful decision support tools.

7.1 Recommendations for future research.

SCIMAPs potential for supporting diffuse FIO pollution mitigation planning has been illustrated for the first time. However, there remains scope for further research considering the application of SCIMAP to diffuse FIO pollution; development of fate and transfer modelling approaches for a variety of objectives; and to develop knowledge regarding the fate and transfer of FIO pollution more generally. Key areas for further investigation arising from this research are:

- Validation of an updated version of SCIMAP encompassing the work flow described in chapter 3 that captures seasonal and between livestock type variation in post defecation *E. coli* proliferation and rainfall induced release.
- Relative to other pollutants, there remains a significant lack of knowledge regarding the fate and transfer of FIOs. In terms of adapting SCIMAP for application to FIO, data regarding the persistence and release of FIOs under a greater variety of meteorological conditions is needed. Understanding the impact of different agricultural practices, for example variable diets and housing regimes, on the fate and transfer of FIOs is also required.
- A strictly quantitative meta-analysis following stringent rules of research synthesis should be used to consolidate existing knowledge on the fate and transfer of *E. coli.* Results from this should feed into the workflow outlined in chapter 6 to inform the development of default hazard weightings for SCIMAP.
- The development of data collection tools and protocols that allow for the efficient and ideally remote collection of spatially distributed samples across catchments will facilitate the application of the SCIMAP fitted approach, thus reducing uncertainty in the framework as a whole. For example, a network of remote monitoring stations in the River Eden catchment, located in Cumbria, England has been used to investigate phosphorus and fine sediment pollution (Perks *et al.* 2015).

Overall, this project has provided new information on the fate and transfer of FIOs in the landscape and has provided a framework for a new model of FIO transfer.

SCIMAP FIO will support catchment managers by supporting decision making regarding the targeting of diffuse FIO mitigation measures, ultimately reducing contamination of recreational bathing areas, shellfish harvesting operations and drinking water supplies.

8 References

Adhikari, H., Barnes, D.L., Schiewer, S. and White, D.M. (2007) Total coliform survival characteristics in frozen soils. *Journal of Environmental Engineering-Asce,* 133 (12), pp. 1098-1105.

Aislabie, J., McLeod, M., Ryburn, J., McGill, A. and Thornburrow, D. (2011) Soil type influences the leaching of microbial indicators under natural rainfall following application of dairy shed effluent. *Soil Research,* 49 (3), pp. 270-279.

Auguie, B. (2015). gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.0.0. https://CRAN.R-project.org/package=gridExtra.

Avery, S.M., Moore, A. and Hutchison, M.L., (2004). Fate of *Escherichia coli* originating from livestock faeces deposited directly onto pasture. *Letters in Applied Microbiology*, **38**(5), pp.355-359.

Bedri, Z., Corkery, A., O'Sullivan, J.J., Alvarez, M.X., Erichsen, A.C., Deering, L.A., Demeter, K., O'Hare, G.M., Meijer, W.G. and Masterson, B. (2014) An integrated catchment-coastal modelling system for real-time water quality forecasts. *Environmental Modelling & Software,* 61, pp. 458-476.

Beharry-Borg, N., Smart, N.B.J.C.R., Termansen, M., Hubacek, K. (2013) Evaluating farmers ’ likely participation in a payment programme for water quality protection in the UK uplands. *Reg. Environ. Change.* **13**, pp. 633-647.

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology,* **320**, 18-36.

Beven, K. (2007). Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process. *Hydrology and Earth System Sciences Discussions*, **11**, 460-467.

Beven, K. and Binley, A. (2014) GLUE: twenty years on. *Hydrological Processes,* 28 (24), pp. 5897-5918.

Beven, K., & Binley, A. (1992). The future of distributed models: model calibration and uncertainty prediction. Hydrological processes, 6(3), 279-298.

Beven, K. & Kirkby, M. (1979). A physically based variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, **24**, 43-69.

Blaustein, R.A., Hill, R.L., Micallef, S.A., Shelton, D.R. and Pachepsky, Y.A., (2016) Rainfall intensity effects on removal of fecal indicator bacteria from solid dairy manure applied over grass-covered soil. *Science of the Total Environment*, **539**, pp.583-591.

Blaustein, R.A., Pachepsky, Y.A., Hill, R.L. & Shelton, D.R. (2015a). Solid Manure as a source of fecal indicator microorganisms: release under simulated rainfall. *Environmental Science & Technology,* **49**, 7860-9.

Blaustein, R.A., Pachepsky, Y.A., Shelton, D.R. and Hill, R.L., (2015b). Release and removal of microorganisms from land-deposited animal waste and animal manures: a review of data and models. *Journal of environmental quality*, **44**(5), pp.1338-1354.

Bloodworth, J., Holman, I., Burgess, P., Gillman, S., Frogbrook, Z. and Brown, P. (2015) Developing a multi-pollutant conceptual framework for the selection and targeting of interventions in water industry catchment management schemes. *Journal of Environmental Management,* 161, pp. 153-162.

Bottcher, D. and Hiscock, J.G. (2002) WAMView–A Point and Non-point Source GIS Watershed Assessment Model. *Proceedings of the Water Environment Federation,* 2002 (2), pp. 1-10.

Bougeard, M., Le Saux, J. C., Perenne, N., Baffaut, C., Robin, M., & Pommepuy, M. (2010). Modeling of *Escherichia Coli* fluxes on a catchment and the impact on coastal water and shellfish quality. *Journal of the American Water Resources Association*, JAWRA-09.

Bradshaw, J.K., Snyder, B.J., Oladeinde, A., Spidle, D., Berrang, M.E., Meinersmann, R.J. *et al.* (2016). Characterizing relationships among fecal indicator bacteria, microbial source tracking markers, and associated waterborne pathogen occurrence in stream water and sediments in a mixed land use watershed. *Water Research,* **101**, 498-509.

Bray, J.R. & Curtis, J.T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs,* **27**, 325-349.

Brennan, F.P., Grant, J., Botting, C.H., O'Flaherty, V., Richards, K.G. and Abram, F. (2013) Insights into the low-temperature adaptation and nutritional flexibility of a soil-persistent Escherichia coli. *FEMS Microbiology Ecology,* 84 (1), pp. 75-85.

Brouwer, A.F., Eisenberg, M.C., Remais, J.V., Collender, P.A., Meza, R. and Eisenberg, J.N. (2017). Modeling biphasic environmental decay of pathogens and implications for risk analysis. *Environmental Science & Technology*, **51**(4), pp.2186-2196.

Bucci, A., Allocca, V., Naclerio, G., Capobianco, G., Divino, F., Fiorillo, F. and Celico, F. (2015) Winter survival of microbial contaminants in soil: An in situ verification. *Journal of Environmental Sciences-China,* 27, pp. 131-138.

Buchanan, B.P., Falbo, K., Schneider, R.L., Easton, Z.M. and Walter, M.T. (2013) Hydrological impact of roadside ditches in an agricultural watershed in Central New York: implications for non-point source pollutant transport. *Hydrological Processes,* 27 (17), pp. 2422-2437.

Carpenter, J. & Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine,* **19**, 1141-1164.

Carr, D., ported by Nicholas Lewin-Koh, Martin Maechler and contains copies of lattice function written by Deepayan Sarkar (2014). hexbin: Hexagonal Binning Routines. Package version 1.27.0. https://CRAN.R-project.org/package=hexbin.

Chadwick, D.; Fish, R.; Oliver, D.M.; Heathwaite, L.; Hodgson, C.; Winter, M. (2008). Management of livestock and their manure to reduce the risk of microbial transfers to water–the case for an interdisciplinary approach. Trends in *Food Science & Technology*, **19**, pp.240-247.

Characklis, G.W., Dilts, M.J., Simmons, O.D., Likirdopulos, C.A., Krometis, L.A.H. and Sobsey, M.D. (2005) Microbial partitioning to settleable particles in stormwater. *Water Research,* 39 (9), pp. 1773-1782.

Cho, K.H., Pachepsky, Y.A., Kim, J.H., Guber, A.K., Shelton, D.R. and Rowland, R. (2010) Release of *Escherichia coli* from the bottom sediment in a first-order creek: Experiment and reach-specific modeling. *Journal of Hydrology,* 391 (3-4), pp. 322-332.

Cho, K.H., Pachepsky, Y.A., Kim, J.H., Kim, J.W. and Park, M.H., (2012). The modified SWAT model for predicting fecal coliforms in the Wachusett Reservoir Watershed, USA. *Water research*, **46**(15), pp.4750-4760.

Cho, K.H., Pachepsky, Y.A., Kim, M., Pyo, J., Park, M., Kim, Y.M. *et al.* (2016a). Modeling seasonal variability of fecal coliform in natural surface waters using the modified SWAT. *Journal of Hydrology,* **535**, 377-385.

Cho, K.H., Pachepsky, Y.A., Oliver, D.M., Muirhead, R.W., Park, Y., Quilliam, R.S. and Shelton, D.R., (2016b). Modeling fate and transport of fecally-derived microorganisms at the watershed scale: state of the science and future opportunities. *Water Research*, 100, 38-56.

Cizek, A.R., Characklis, G.W., Krometis, L., Hayes, J.A., Simmons, O.D.,III, Di Lonardo, S., Alderisio, K.A. and Sobsey, M.D. (2008) Comparing the partitioning behavior of *Giardia* and *Cryptosporidium* with that of indicator organisms in stormwater runoff. *Water Research,* 42 (17), pp. 4421-4438.

Clements, K.; Quilliam, R.S.; Jones, D.L.; Wilson, J.; Malham, S.K. (2015) Spatial and temporal heterogeneity of bacteria across an intertidal shellfish bed: Implications for regulatory monitoring of faecal indicator organisms. *Science of the Total Environment*., **506**, pp.1-9.

Coffey, R., Cummins, E., Bhreathnach, N., Flaherty, V. & Cormican, M. (2010). Development of a pathogen transport model for Irish catchments using SWAT. *Agricultural Water Management,* **97**, 101-111.

Collins, R., Elliott, S. and Adams, R. (2005) Overland flow delivery of faecal bacteria to a headwater pastoral stream. *Journal of Applied Microbiology,* 99 (1), pp. 126-132.

Collins, R., Mcleod, M., Hedley, M., Donnison, A., Close, M., Hanly, J., Horne, D., Ross, C., Davies-Colley, R. and Bagshaw, C. (2007) Best management practices to mitigate faecal contamination by livestock of New Zealand waters. *New Zealand Journal of Agricultural Research,* 50 (2), pp. 267-278.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., and Böhner, J. (2015): System for Automated Geoscientific Analyses (SAGA) v. 2.1.4, *Geosci. Model Dev*., 8, 1991-2007.

Davidian, M. & Giltinan, D.M. (1995). *Nonlinear models for repeated measurement data*. CRC press.

Davies, C.M., Ferguson, C.M., Kaucner, C., Krogh, M., Altavilla, N., Deere, D.A. and Ashbolt, N.J., (2004). Dispersion and transport of Cryptosporidium oocysts from fecal pats under simulated rainfall events. *Applied and environmental microbiology*, **70**(2), pp.1151-1159.

de Brauwere, A., Ouattara, N.K. and Servais, P. (2014) Modeling fecal indicator bacteria concentrations in natural surface waters: a review. *Critical Reviews in Environmental Science and Technology,* 44 (21), pp. 2380-2453.

Deepayan, S. (2008) Lattice: Multivariate Data Visualization with R. Springer, New York. ISBN 978-0-387-75968-5.

Donnison, A., Ross, C. and Clark, D. (2008) Escherichia coli shedding by dairy cows. *New Zealand Journal of Agricultural Research,* 51 (3), pp. 273-278.

Dorner, S.M., Anderson, W.B., Slawson, R.M., Kouwen, N. and Huck, P.M. (2006) Hydrologic modeling of pathogen fate and transport. *Environmental Science & Technology,* 40 (15), pp. 4746-4753.

Droppo, I.G., Krishnappan, B.G., Liss, S.N., Marvin, C. and Biberhofer, J. (2011) Modelling sediment-microbial dynamics in the South Nation River, Ontario, Canada: Towards the prediction of aquatic and human health risk. *Water Research,* 45 (12), pp. 3797-3809.

Dymond, J.R., Serezat, D., Ausseil, A.G.E. and Muirhead, R.W. (2016). Mapping of *Escherichia coli* sources connected to waterways in the Ruamahanga catchment, New Zealand. *Environmental science & technology*, **50**(4), pp.1897-1905.

E.U. (2006a). Directive 2006/07/EC of the European parliament and of the council of 15th February 2006 concerning the management of bathing water quality and repealing directive 76/160/EEC. *Official Journal of the European Union* L 64, 37e51.04/03/2006.

E.U. (2006b). Directive 2006/113/EC of the European parliament and of the council of 12th December 2006 on the quality required of shellfish water. *Official Journal of the European Union* L 376, 14e20.27/12/2006.

Engström, E., Thunvik, R., Kulabako, R. and Balfors, B., (2015). Water transport, retention, and survival of *Escherichia coli* in unsaturated porous media: a comprehensive review of processes, models, and factors. *Critical Reviews in Environmental Science and Technology*, *45*, 1-100.

Environment Agency (2016). River flow and water quality information for sites in the Catchment Sensitive Farming Enhanced Water Quality Monitoring Programme (CSF EWQMP) provided by the Environment Agency © Environment Agency copyright and database right 2016. All rights reserved. Licensed under the Open Government Licence v2.0.

Evanson, M. and Ambrose, R.F. (2006) Sources and growth dynamics of fecal indicator bacteria in a coastal wetland system and potential impacts to adjacent waters. *Water Research,* 40 (3), pp. 475-486.

Eze, J. I., Scott, E. M., Pollock, K. G., Stidson, R., Miller, C. A., & Lee, D. (2014). The association of weather and bathing water quality on the incidence of gastrointestinal illness in the west of Scotland. *Epidemiology and infection*, 142(06), 1289-1299.

Falbo, K., Schneider, R.L., Buckley, D.H., Walter, M.T., Bergholz, P.W. and Buchanan, B.P. (2013) Roadside ditches as conduits of fecal indicator organisms and sediment: Implications for water quality management. *Journal of Environmental Management,* 128, pp. 1050-1059.

Falkenmark, M. (2004). Towards integrated catchment management: opening the paradigm locks between hydrology, ecology and policy-making*. International Journal of Water Resources Development*, **20**, no. 3, pp. 275-281.Fraser, R.H., Barten, P.K. and Pinney, D.A. (1998) Predicting stream pathogen loading from livestock using a geographical information system-based delivery model. *Journal of Environmental Quality,* **27** (4), pp. 935-945.

Ferguson, C.M., Davies, C.M., Kaucner, C., Ashbolt, N.J., Krogh, M., Rodehutskors, J. *et al.* (2007). Field scale quantification of microbial transport from bovine faeces under simulated rainfall events. *Journal of Water and Health,* **5**, 83-95.

Fiener, P. and Auerswald, K., (2003). Effectiveness of grassed waterways in reducing runoff and sediment delivery from agricultural watersheds. *Journal of Environmental Quality*, **32**(3), pp.927-936.

Freestone, P. P., Sandrini, S. M., Haigh, R. D., & Lyte, M. (2008). Microbial endocrinology: how stress influences susceptibility to infection. Trends in microbiology, **16**(2), 55-64.

Garzio-Hadzick, A., Shelton, D.R., Hill, R.L., Pachepsky, Y.A., Guber, A.K. and Rowland, R. (2010) Survival of manure-borne *E. coli* in streambed sediment: Effects of temperature and sediment properties. *Water Research,* **44** (9), pp. 2753-2762.

Goss, M. & Richards, C. (2008). Development of a risk-based index for source water protection planning, which supports the reduction of pathogens from agricultural activity entering water resources. *Journal of Environmental Management,* **87**, 623-632.

Graves, Piepho, Selzer & Dorai-Raj (2015). multcompView: Visualizations of Paired Comparisons. R package version 0.1-7. https://CRAN.R-project.org/package=multcompView

Guber, A., Pachepsky, Y., Yakirevich, A., Shelton, D., Sadeghi, A., Goodrich, D. and Unkrich, C. (2011) Uncertainty in modelling of faecal coliform overland transport associated with manure application in Maryland. *Hydrological Processes,* **25** (15), pp. 2393-2404.

Guber, A.K., Fry, J., Ives, R.L. and Rose, J.B. (2015) *Escherichia coli* Survival in, and Release from, White-Tailed Deer Feces. *Applied and Environmental Microbiology,* **81** (3), pp. 1168-1176.

Guber, A.K., Pachepsky, Y.A., Dao, T.H., Shelton, D.R. & Sadeghi, A.M. (2013). Evaluating manure release parameters for nonpoint contaminant transport model KINEROS2/STWIR. *Ecological Modelling,* **263**, 126-138.

Guber, A.K., Shelton, D.R. and Pachepsky, Y.A., (2005). Transport and retention of manure-borne coliforms in soil. *Vadose Zone Journal*, **4**(3), pp.828-837.

Gurevitch, J., Koricheva, J., Nakagawa, S. and Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature*, **555**(7695), p.175.

Harmel, R.D., Karthikeyan, R., Gentry, T. and Srinivasan, R. (2010). Effects of agricultural management, land use, and watershed scale on *E. coli* concentrations in runoff and streamflow. *Transactions of the ASABE,* **53**(6), pp.1833-1841.

Harwood, V.J., Staley, C., Badgley, B.D., Borges, K. and Korajkic, A. (2014) Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiology Reviews,* **38** (1), pp. 1-40.

Haygarth, P.M., Condron, L.M., Heathwaite, A.L., Turner, B.L. & Harris, G.P. (2005). The phosphorus transfer continuum: Linking source to impact with an interdisciplinary and multi-scaled approach. *Science of the Total Environment,* **344**, 5-14.

Heathwaite, A.L., Quinn, P.F. & Hewett, C.J.M. (2005). Modelling and managing critical source areas of diffuse pollution from agricultural land using flow connectivity simulation. *Journal of Hydrology,* **304**, 446-461.

Hénault-Ethier, L., (2007). Vermicomposting: from microbial and earthworm induced effects in bacterial sanitation to the chemistry of biodegradation under batch or continuous operation (Doctoral dissertation, Concordia University).

Hodgson, C.J., Bulmer, N., Chadwick, D.R., Oliver, D.M., Heathwaite, A.L., Fish, R.D. *et al.* (2009). Establishing relative release kinetics of faecal indicator organisms from different faecal matrices. *Letters in Applied Microbiology,* **49**, 124-130.

Hudson, P.F. (2003) Event sequence and sediment exhaustion in the lower Panuco Basin, Mexico. *Catena,* 52 (1), pp. 57-76.

Ishii, S., Yan, T., Vu, H., Hansen, D.L., Hicks, R.E. & Sadowsky, M.J. (2010). Factors Controlling Long-Term Survival and Growth of Naturalized *Escherichia coli* Populations in Temperate Field Soils. *Microbes and Environments,* **25**, 8-14.

Jenkins, M.B., Adams, M.P., Endale, D.M., Fisher, D.S., Lowrance, R., Newton, G.L. *et al.* (2015). Storm flow dynamics and loads of fecal bacteria associated with ponds in southern piedmont and coastal plain watersheds with animal agriculture. *Agricultural Water Management,* **148**, 97-105.

Jenkins, M.B., Endale, D.M., Fisher, D.S., Adams, M.P., Lowrance, R., Newton, G.L. and Vellidis, G. (2012) Survival dynamics of fecal bacteria in ponds in agricultural watersheds of the Piedmont and Coastal Plain of Georgia. *Water Research,* 46 (1), pp. 176-186.

Joensen, K.G., Scheutz, F., Lund, O., Hasman, H., Kaas, R.S., Nielsen, E.M. and Aarestrup, F.M. (2014). Evaluation of real-time WGS for routine typing, surveillance and outbreak detection of verotoxigenic Escherichia coli. *Journal of clinical microbiology*, pp.JCM-03617.

Jones, T., Gill, C.O. & McMullen, L.M. (2004). The behaviour of log phase *Escherichia coli* at temperatures that fluctuate about the minimum for growth. *Lett Appl Microbiol* **39,** 296–300

Kay, D., Crowther, J., Stapleton, C.M. and Wyer, M.D. (2018) Faecal indicator organism inputs to watercourses from streamside pastures grazed by cattle: Before and after implementation of streambank fencing. *Water Research*

Kay, D., Aitken, M., Crowther, J., Dickson, I., Edwards, A., Francis, C. *et al.* (2007a). Reducing fluxes of faecal indicator compliance parameters to bathing waters from diffuse agricultural sources: the Brighouse Bay study, Scotland. *Environmental Pollution,* **147**, 138-149.

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

Kay, D., Crowther, J., Fewtrell, L., Francis, C.A., Hopkins, M., Kay, C., McDonald, A.T., Stapleton, C.M., Watkins, J., Wilkinson, J. and Wyer, M.D. (2008) Quantification and control of microbial pollution from agriculture: a new policy challenge? *Environmental Science & Policy,* **11** (2), pp. 171-184.

Kay, D., Crowther, J., Stapleton, C., Wyer, M., Fewtrell, L., Anthony, S. *et al.* (2008). Faecal indicator organism concentrations and catchment export coefficients in the UK. *Water Research,* **42**, 2649-2661.

Kay, D., Edwards, A.C., Ferrier, R.C., Francis, C., Kay, C., Rushby, L., Watkins, J., McDonald, A.T., Wyer, M., Crowther, J. and Wilkinson, J. (2007b) Catchment microbial dynamics: the emergence of a research agenda. *Progress in Physical Geography,* 31 (1), pp. 59-76.

Kay, D., McDonald, A., Stapleton, C., Wyer, M. and Crowther, J. (2009) Catchment to coast systems–managing microbial pollutants for bathing and shellfish harvesting waters. *Handbook of Catchment Management,* , pp. 183-210.

Kibet, L.C., Saporito, L.S., Allen, A.L., May, E.B., Kleinman, P.J., Hashem, F.M. and Bryant, R.B., (2014). A protocol for conducting rainfall simulation to study soil runoff. *Journal of visualized experiments: JoVE*, **86**.

Kiefer, L.A., Shelton, D.R., Pachepsky, Y., Blaustein, R. and Santin-Duran, M. (2012) Persistence of *Escherichia coli* introduced into streambed sediments with goose, deer and bovine animal waste. *Letters in Applied Microbiology,* 55 (5), pp. 345-353.

Koskey, A.M., Fisher, J.C., Traudt, M.F., Newton, R.J. and McLellan, S.L. (2014) Analysis of the gull fecal microbial community reveals the dominance of *Catellicoccus marimammalium* in relation to culturable Enterococci. *Applied and Environmental Microbiology,* **80** (2), pp. 757-765.

Krometis, L.-., Characklis, G., Sobsey, M. and Noble, R. (2009) Quantification of *E. coli* Partitioning in Surface Waters via Culture-based Techniques and qPCR. *Abstracts of the General Meeting of the American Society for Microbiology,* **109**.

Kruskal, J. B. (1964b). Multidimensional scaling by optimizing goodness-of-fit to a non-metric hypothesis. Psychometrika, **29**, 1-27.

Kruskal, J.B. (1964a). Nonmetric multidimensional scaling: a numerical method. *Psychometrika,* **29**, 115-129.

Lane, S.N., Reaney, S.M. and Heathwaite, A.L., (2009) Representation of landscape hydrological connectivity using a topographically driven surface flow index. *Water Resources Research*, **45**(8).

Lau, M. & Ingham, S. (2001). Survival of faecal indicator bacteria in bovine manure incorporated into soil. *Letters in Applied Microbiology,* **33**, 131-136.

Liang, X., Soupir, M.L., Rigby, S., Jarboe, L.R. and Zhang, W. (2014) Flow cytometry is a promising and rapid method for differentiating between freely suspended *Escherichia coli* and E. coli attached to clay particles. *Journal of Applied Microbiology,* **117** (6), pp. 1730-1739.

Liu, P., Soupir, M.L., Zwonitzer, M., Huss, B. & Jarboe, L.R. (2011). Association of Antibiotic Resistance in Agricultural *Escherichia coli* Isolates with Attachment to Quartz. *Applied and Environmental Microbiology,* **77**, 6945-6953.

Martinez, G., Pachepsky, Y.A., Shelton, D.R., Whelan, G., Zepp, R., Molina, M. *et al.* (2013). Using the $Q_{(10)}$ model to simulate *E. coli* survival in cowpats on grazing lands. *Environment International,* **54**, 1-10.

Mattick, K.L., Phillips, L.E., JØRGENSEN, F., Lappin-Scott, H.M. & Humphrey, T.J., (2003). Filament formation by Salmonella spp. inoculated into liquid food matrices at refrigeration temperatures, and growth patterns when warmed. *Journal of food protection*, **66**(2), 215-219.

McBride, G.B., Stott, R., Miller, W., Bambic, D. and Wuertz, S. (2013) Discharge-based QMRA for estimation of public health risks from exposure to stormwater-borne pathogens in recreational waters in the United States. *Water Research,* 47 (14), pp. 5282-5297.

McGonigle, D.F., Harris, R.C., McCamphill, C., Kirk, S., Dils, R., Macdonald, J. and Bailey, S. (2012) Towards a more strategic approach to research to support catchment-based policy approaches to mitigate agricultural water pollution: A UK case-study. *Environmental Science & Policy,* 24, pp. 4-14.

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.

McKergow, L.A. & Davies-Colley, R.J. (2010). Stormflow dynamics and loads of *Escherichia coli* in a large mixed land use catchment. *Hydrological Processes,* **24**, 276-289.

Met Office (2012): Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853-current). NCAS British Atmospheric Data Centre, Accessed: 10 February 2016.

Met Office (2014) UKCP09: 5 km gridded data - annual averages. Rainfall intensity on days of rain ≥1 mm (mm/day) 1971 – 2000. Downloaded November 2014.

Retrieved from:
http://www.metoffice.gov.uk/climatechange/science/monitoring/ukcp09/download/longterm/fivekm_annual.html

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.

Monaghan, R., De Klein, C. and Muirhead, R. (2008) Prioritisation of farm scale remediation efforts for reducing losses of nutrients and faecal indicator organisms to waterways: A case study of New Zealand dairy farming. *Journal of Environmental Management,* 87 (4), pp. 609-622.

Morales, I., Atoyan, J.A., Amador, J.A. and Boving, T., (2014). Transport of pathogen surrogates in soil treatment units: Numerical modeling. *Water*, **6**(4), pp.818-838.

Morató, J., Codony, F., Sánchez, O., Pérez, L.M., García, J. & Mas, J. (2014). Key design factors affecting microbial community composition and pathogenic organism removal in horizontal subsurface flow constructed wetlands. *Science of the Total Environment,* **481**, 81-89.

Moriarty, E.M. & Gilpin, B.J. (2014). Leaching of *Escherichia coli* from sheep faeces during simulated rainfall events. *Letters in Applied Microbiology,* **58**, 569-575.

Moriarty, E.M., Mackenzie, M.L., Karki, N. and Sinton, L. W., (2011b). Survival of Escherichia coli, Enterococci, and Campylobacter spp. in sheep feces on pastures. *Applied and environmental microbiology*, **77**(5), pp.1797-1803.

Moriarty, E.M., Sinton, L.W., Mackenzie, M.L., Karki, N. and Wood, D.R. (2008) A survey of enteric bacteria and protozoans in fresh bovine faeces on New Zealand dairy farms. *Journal of Applied Microbiology,* 105 (6), pp. 2015-2025.

Moriarty, E.M., Weaver, L., Sinton, L.W. and Gilpin, B. (2012) Survival of *Escherichia coli,* Enterococci and *Campylobacter jejuni* in Canada Goose Faeces on Pasture. *Zoonoses and Public Health,* 59 (7), pp. 490-497.

Moriarty, E.M.; McEwan, N.; Mackenzie, M.; Karki, N.; Sinton, L.W. (2011a) Incidence and prevalence of microbial indicators and pathogens in ovine faeces in New Zealand. *New Zealand Journal of Agricultural Research*. **54**, pp. 71-81.

Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G. *et al.* (2011). Final Report for LCM2007-the new UK land cover map. Countryside Survey Technical Report No 11/07.

Muirhead, R. (2015). A Farm-Scale Risk-Index for Reducing Fecal Contamination of Surface Waters. *Journal of Environmental Quality,* **44**, 248-255.

Muirhead, R., Davies-Colley, R., Donnison, A. and Nagels, J. (2004) Faecal bacteria yields in artificial flood events: quantifying in-stream stores. *Water Research,* 38 (5), pp. 1215-1224.

Muirhead, R.W. & Littlejohn, R.P. (2009). Die-off of *Escherichia coli* in intact and disrupted cowpats. *Soil Use and Management,* **25**, 389-394.

Muirhead, R.W., Collins, R.P. & Bremer, P.J. (2005). Erosion and subsequent transport state of *Escherichia coli* from cowpats. *Applied and Environmental Microbiology,* **71**, 2875-2879.

Muirhead, R.W., Elliott, A.H. and Monaghan, R.M. (2011) A model framework to assess the effect of dairy farms and wild fowl on microbial water quality during base-flow conditions. *Water Research,* 45 (9), pp. 2863-2874.

Murphy, S., Jordan, P., Mellander, P.E. and O'Flaherty, V. (2015). Quantifying faecal indicator organism hydrological transfer pathways and phases in agricultural catchments. *Science of The Total Environment*, **520**, pp.286-299.

Nakagawa, S. and Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**(2), 133-142.

Neuwirth (2014). RColorBrewer: ColorBrewer Palettes. R package version 1.1-2. https://CRAN.R-project.org/package=RColorBrewer

Ogle, D.H. (2015) FSA: Fisheries Stock Analysis. R package version 0.8.1.

Oladeinde, A., Bohrmann, T., Wong, K., Purucker, S.T., Bradshaw, K., Brown, R. *et al.* (2014). Decay of fecal indicator bacterial populations and bovine-associated source-tracking markers in freshly deposited cow pats. *Applied and Environmental Microbiology,* **80**, 110-118.

Oliver DM, Page T, Zhang T, Heathwaite AL, Beven K, Carter H, McShane G, O' Keenan P & Haygarth PM (2012). Determining *E. coli* burden on pasture in a headwater catchment: Combined field and modelling approach, *Environment International,* **43**, 6-12.

Oliver, D.M. (2014). Seasonal and within-herd variability of *E. coli* concentrations in fresh dairy faeces. *Letters in Applied Microbiology,* **59**, 86-92.

Oliver, D.M. & Page, T. (2016a). Effects of seasonal meteorological variables on *E. coli* persistence in livestock faeces and implications for environmental and human health. *Scientific Reports,* **6**, 37101.

Oliver, D.M., Bartie, P.J., Heathwaite, A.L, Reaney, S.M., Parnell, J.A.Q & Quilliam, R.S. (2018). A catchment-scale model to predict spatial and temporal burden of *E. coli* on pasture from grazing livestock, *Science of the Total Environment*, 616-617, 678-687

Oliver, D.M., C.D. Clegg, P.M. Haygarth, and A.L. Heathwaite. (2005a) Assessing the potential for pathogen transfer from grassland soils to surface waters. *Adv. Agron*. 85:125–180.

Oliver, D.M., Haygarth, P.M., Clegg, C.D. & Heathwaite, L. (2006). Differential *E. coli* die-off patterns associated with agricultural matrices. *Environmental Science & Technology,* **40**, 5710-5716.

Oliver, D.M., Heathwaite, L., Haygarth, P.M. and Clegg, C.D. (2005) Transfer of *Escherichia coli* to water from drained and undrained grassland after grazing. *Journal of Environmental Quality,* 34 (3), pp. 918-925.

Oliver, D.M., Page, T., Heathwaite, A.L. & Haygarth, P.M. (2010). Re-shaping models of *E. coli* population dynamics in livestock faeces: Increased bacterial risk to humans? *Environment International,* **36**, 1-7.

Oliver, D.M., Page, T., Hodgson, C.J., Heathwaite, A.L., Chadwick, D.R., Fish, R.D. and Winter, M. (2010) Development and testing of a risk indexing framework to determine field-scale critical source areas of faecal bacteria on grassland. *Environmental Modelling & Software,* 25 (4), pp. 503-512.

Oliver, D.M., Porter K.D.H., Pachepsky, Y.A., Muirhead, R.W., Reaney, S.M., Coffey, R., Kay, D. Milledge, D.G., Anthony, S.G., Page, T., Bloodworth, J.W., Mellander, E., Carbonneau, P.E., McGrane, S.J., Quilliam, R.S. (2016b) Predicting microbial water quality with models: Over-arching questions for managing risk in agricultural catchments. *Science of the Total Environment*, **544**, 39-47.

Oliver, D.M., Porter, K.D.H., Heathwaite, A.L., Zhang, T. and Quilliam, R.S. (2015a) Impact of low intensity summer rainfall on *E. coli*-discharge event dynamics with reference to sample acquisition and storage. *Environmental Monitoring and Assessment,* 187 (7), pp. 426-426.

Oliver, D.M.; Page, T. (2016b) Effects of seasonal meteorological variables on E. coli persistence in livestock faeces and implications for environmental and human health. *Scientific Reports*. **6**.

OS MasterMap Topography Layer [GML geospatial data], Coverage: Lancashire and Devon, Ordnance Survey, GB. Using: EDINA Digimap Ordnance Survey Service, Downloaded: June 2015.

Owen, G.J., Perks, M.T., Benskin, C.M.H., Wilkinson, M.E., Jonczyk, J. and Quinn, P.F. (2012). Monitoring agricultural diffuse pollution through a dense monitoring network in the River Eden Demonstration Test Catchment, Cumbria, UK. Area, **44**(4), pp.443-453.

Pachepsky, Y., Shelton, D., Dorner, S. and Whelan, G. (2016). Can *E. coli* or thermotolerant coliform concentrations predict pathogen presence or prevalence in irrigation waters? *Critical reviews in microbiology*, **42**(3), pp.384-393.

Pachepsky, Y.A. & Shelton, D.R. (2011). *Escherichia coli* and Fecal Coliforms in Freshwater and Estuarine Sediments. *Critical Reviews in Environmental Science and Technology,* **41**, 1067-1110.

Paine, C.E., Marthews, T.R., Vogt, D.R., Purves, D., Rees, M., Hector, A. and Turnbull, L.A. (2012). How to fit nonlinear plant growth models and calculate growth rates: an update for ecologists. *Methods in Ecology and Evolution*, **3**(2), pp.245-256.

Pandey, A., Himanshu, S.K., Mishra, S.K. and Singh, V.P. (2016) Physically based soil erosion and sediment yield models revisited. *Catena*, **147**, pp.595-620.

Pandey, P.K. & Soupir, M.L. (2013). Assessing the Impacts of *E. coli* Laden Streambed Sediment on E. Coli Loads Over a Range of Flows and Sediment Characteristics. *Journal of the American Water Resources Association,* **49**, 1261-1269.

Pandey, P.K., Kass, P.H., Soupir, M.L., Biswas, S. & Singh, V.P. (2014). Contamination of water resources by pathogenic bacteria. AMB Express, vol. 4, pp. 51-014-0051-x. eCollection 2014.

Pandey, P.K., Soupir, M.L. and Rehmann, C.R. (2012) A model for predicting resuspension of *Escherichia coli* from streambed sediments. *Water Research,* **46** (1), pp. 115-126.

Pandey, P.K., Soupir, M.L., Ikenberry, C.D. and Rehmann, C.R., (2016). Predicting streambed sediment and water column Escherichia coli levels at watershed scale. *Journal of the American Water Resources Association*, **52**, pp.184-197.

Park, Y., Pachepsky, Y., Shelton, D., Jeong, J. and Whelan, G., (2016). Survival of manure-borne and fecal coliforms in soil: temperature dependence as affected by site-specific factors. *Journal of Environmental Quality*, **45**, pp.949-957.

Pedersen, J.C. & Hendriksen, N.B. (1993). Effect of passage through the intestinal tract of detritivore earthworms (Lumbricus spp.) on the number of selected Gram-negative and total bacteria. *Biology and Fertility of Soils,* **16**, 227-232.

Perks, M.T., Owen, G.J., Benskin, C.M.H., Jonczyk, J., Deasy, C., Burke, S., Reaney, S.M. and Haygarth, P.M., (2015). Dominant mechanisms for the delivery of fine sediment and phosphorus to fluvial networks draining grassland dominated headwater catchments. *Science of the Total Environment*, **523**, pp.178-190.

Pinheiro J, Bates D, DebRoy S, Sarkar D and R Core Team (2015). _nlme: Linear and Nonlinear Mixed Effects Models_. R package version 3.1-122, <URL: http://CRAN.R-project.org/package=nlme>.

Pinheiro, J.C. & Bates, D.M. (1995) Mixed effects models, methods, and classes for S and Splus. Department of Statistics, University of Wisconsin. Available from Statlib

Piorkowski, G., Jamieson, R., Bezanson, G., Hansen, L.T. and Yost, C. (2014) Reach specificity in sediment *E. coli* population turnover and interaction with waterborne populations. *Science of the Total Environment,* 496, pp. 402-413.

Porter, K.D., Reaney, S.M., Quilliam, R.S., Burgess, C. and 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**, pp.456-465.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rea, C.L., Bisesi, M.S., Mitsch, W., Andridge, R. and Lee, J. (2015) Human Health-Related Ecosystem Services of Avian-Dense Coastal Wetlands Adjacent to a Western Lake Erie Swimming Beach. *Ecohealth,* **12** (1), pp. 77-87.

Reaney, S.M., (2008) The use of agent based modelling techniques in hydrology: determining the spatial and temporal origin of channel flow in semi-arid catchments. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, **33**(2), pp.317-327.

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.

Retzbach, J., Otto, L. and Maier, M. (2015) Measuring the perceived uncertainty of scientific evidence and its relationship to engagement with science. *Public Understanding of Science (Bristol, England)*.

Rickson, R. (2014) Can control of soil erosion mitigate water pollution by sediments? *Science of the Total Environment,* **468**, pp. 1187-1197.

Rockström, J., Williams, J., Daily, G., Noble, A., Matthews, N., Gordon, L., Wetterstrand, H., DeClerck, F., Shah, M., Steduto, P. and de Fraiture, C. (2017). Sustainable intensification of agriculture for human prosperity and global sustainability. *Ambio*, **46**(1), p.4.

Rohatgi, A. (2018). WebPlotDigitizer. [ONLINE] Available at: https://automeris.io/WebPlotDigitizer. [Accessed 10 November 2017].

Ryan, U., Yang, R., Gordon, C. & Doube, B. (2011). Effect of dung burial by the dung beetle Bubas bison on numbers and viability of *Cryptosporidium* oocysts in cattle dung. *Experimental Parasitology,* **129**, 1-4.

Sarkar, D. and Andrews, F. (2013). latticeExtra: Extra Graphical Utilities Based on Lattice. R package version 0.6-26. https://CRAN.R-project.org/package=latticeExtra.

Schoumans, O., Chardon, W., Bechmann, M., Gascuel-Odoux, C., Hofman, G., Kronvang, B., Rubæk, G.H., Ulen, B. and Dorioz, J. (2014) Mitigation options to reduce phosphorus losses from the agricultural sector and improve surface water quality: A review. *Science of the Total Environment,* **468**, pp. 1255-1266.

Scottish Climate Projections App. (2017). [online] Available at: http://www.environment.scotland.gov.uk/get-interactive/data/scottish-climate-projections/ [Accessed 2 Jun. 2017]. Contains information licensed under the Non-Commercial Government Licence v1.0.

Shelton, D.R., Pachepsky, Y.A., Kiefer, L.A., Blaustein, R.A., McCarty, G.W. and Dao, T.H. (2014) Response of coliform populations in streambed sediment and water column to changes in nutrient concentrations in water. *Water Research,* 59, pp. 316-**324**.

Shen, J., Sun, S. and Wang, T. (2005) Development of the fecal coliform total maximum daily load using Loading Simulation Program C and tidal prism model in estuarine shellfish growing areas: A case study in the Nassawadox coastal embayment, Virginia. *Journal of Environmental Science and Health,* **40** (9), pp. 1791-1807.

Sigua, G.C., Palhares, J.C.P., Kich, J.D., Mulinari, M.R., Mattei, R.M., Klein, J.B., Muller, S. & Plieske, G. (2010). Microbiological quality assessment of watershed associated with animal-based agriculture in Santa Catarina, Brazil. *Water, Air, & Soil Pollution*, **210**, 307-316.

SoDa Service (2013). [online] Available at: http://www.soda-is.com [Accessed 18 March 2013]. Contains information licensed under the License Agreement for SoDa Products and Services available for free, available at: http://www.soda-pro.com/documents/10157/326300/SoDa-license.pdf

Soupir, M.L., Mostaghimi, S. & Lou, J. (2008a). Die-Off of *E. coli* and Enterococci in Dairy Cowpats. *Transactions of the Asabe,* **51**, 1987-1996.

Soupir, M.L., Mostaghimi, S. and Dillaha, T. (2010) Attachment of *Escherichia coli* and Enterococci to Particles in Runoff. *Journal of Environmental Quality,* **39** (3), pp. 1019-1027.

Soupir, M.L., Mostaghimi, S. and Love, N.G. (2008b) Method to Partition Between Attached and Unattached *E-coli* in Runoff From Agricultural Lands. *Journal of the American Water Resources Association,* **44** (6), pp. 1591-1599.

Spiegelhalter, D., Pearson, M. and Short, I. (2011) Visualizing uncertainty about the future. *Science (New York, N.Y.),* 333 (6048), pp. 1393-1400.

Stevens, C.J. and Quinton, J.N. (2009) Policy implications of pollution swapping. *Physics and Chemistry of the Earth, Parts A/B/C,* 34 (8), pp. 589-594.

Stocker, M.D., Pachepsky, Y.A. & Shelton, D.R. (2014). Performance of Weibull and Linear Semi-logarithmic Models in Simulating *Escherichia coli* Inactivation in Waters. *Journal of Environmental Quality,* **43**, 1559-1565.

Tate, K.W., Atwill, E.R., Bartolome, J.W. & Nader, G. (2006). Significant attenuation by vegetative buffers on annual grasslands. *Journal of Environmental Quality,* **35**, 795-805.

Tetzlaff, D., Cape, 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.

Texier, S., Prigent-Combaret, C., Gourdon, M.H., Poirier, M.A., Faivre, P., Dorioz, J.M. *et al.* (2008). Persistence of Culturable Fecal Contaminants in Dairy Alpine Grassland Soils. *Journal of Environmental Quality,* **37**, 2299-2310.

Van Kessel, J.S., Pachepsky, Y.A., Shelton, D.R. & Karns, J.S. (2007). Survival of *Escherichia coli* in cowpats in pasture and in laboratory conditions. *Journal of Applied Microbiology,* **103**, 1122-1127.

VanderZaag, A.C., Campbell, K.J., Jamieson, R.C., Sinclair, A.C. and Hynes, L.G. (2010) Survival of *Escherichia coli* in agricultural soil and presence in tile drainage and shallow groundwater. *Canadian Journal of Soil Science,* 90 (3), pp. 495-505.

Wade, T.J., Calderon, R.L., Sams, E., Beach, M., Brenner, K.P., Williams, A.H. *et al.* (2006). Rapidly measured indicators of recreational water quality are predictive of swimming-associated gastrointestinal illness. *Environmental Health Perspectives*, 24-28.

Walters, E., Kaetzl, K., Schwarzwaeder, K., Rutschmann, P., Mueler, E. and Horn, H. (2014) Persistence of fecal indicator bacteria in sediment of an oligotrophic river: Comparing large and lab-scale flume systems. *Water Research,* 61, pp. 276-287.

Wang, L., Mankin, K.R. & Marchin, G.L. (2004). Survival of fecal bacteria in dairy cow manure. *Transactions of the ASAE,* **47**, 1239-1246.

Whatmore, S.J., Lane, S.N., Odoni, N.A., Ward, N. and Bradley, S. (2011) Coproducing flood risk knowledge: redistributing expertise in critical participatory modelling. *Environment and Planning A,* 43 (7), pp. 1617-1633.

Wickham & Francois (2016). dplyr: A Grammar of Data Manipulation. R package version 0.5.0. https://CRAN.R-project.org/package=dplyr

Wickham, H. (2007). Reshaping Data with the reshape Package. Journal of Statistical Software, 21, 1-20.

Wickham, H. (2014). tidyr: Easily Tidy Data with spread() and gather() Functions.. R package version 0.2.0. https://CRAN.R-project.org/package=tidyr.

Wickham, H. (2015). lazyeval: Lazy (Non-Standard) Evaluation. R package version 0.1.10. https://CRAN.R-project.org/package=lazyeval.

Wilkinson, K.G., Tee, E., Tomkins, R.B., Hepworth, G. and Premier, R. (2011) Effect of heating and aging of poultry litter on the persistence of enteric bacteria. *Poultry Science,* 90 (1), pp. 10-18.

Wilkinson, R.J., McKergow, L.A., Davies-Colley, R.J., Ballantine, D.J. and Young, R.G. (2011) Modelling storm-event *E. coli* pulses from the Motueka and Sherry

Rivers in the South Island, New Zealand. *New Zealand Journal of Marine and Freshwater Research,* 45 (3), pp. 369-393.

Williams, G.P. (1989) Sediment concentration versus water discharge during single hydrologic events in rivers. *Journal of Hydrology,* 111 (1), pp. 89-106.

Williams, S. K. R., & Caldwell, K. D. (2014). Field-flow fractionation. *Analytical and Bioanalytical Chemistry*, 406(6), 1577-1578.

Winter, M., Oliver, D.M., Fish, R., Heathwaite, A.L., Chadwick, D. and Hodgson, C. (2011). Catchments, sub-catchments and private spaces: Scale and process in managing microbial pollution from source to sea. *Environmental Science & Policy*, *14*, 315-326.

Wu, J. & Jackson, L. (2016). Association of land use and its change with beach closure in the United States, 2004–2013. *Science of the Total Environment*, 571, 67-76.

Wu, J., Yunus, M., Islam, M.S. and Emch, M., (2016). Influence of climate extremes and land use on fecal contamination of shallow tubewells in Bangladesh. *Environmental science & technology*, **50**(5), pp.2669-2676.

Yang, Q., Tian, H., Li, X., Ren, W., Zhang, B., Zhang, X. and Wolf, J. (2016). Spatiotemporal patterns of livestock manure nutrient production in the conterminous United States from 1930 to 2012. *Science of the Total Environment*, **541**, pp.1592-1602.

Yearley, S. (1999) Computer Models and the Public's Understanding of Science A Case-Study Analysis. *Social Studies of Science,* 29 (6), pp. 845-866.