AN ABSTRACT OF THE DISSERTATION OF

Lauren Brooks for the degree of Doctor of Philosophy in Microbiology presented on August 3, 2016.

Title: Novel Experimental Designs and Mathematical Models to Study Fecal Indicator Bacteria Persistence in Surface Water.

Abstract approved:

_____

Katharine G. Field

High counts of fecal indicators, used to signal the potential presence of pathogens associated with untreated waste, result in the classification of water bodies throughout the United States as impaired. Nonpoint sources of unknown origin that contribute to fecal contamination make management of impaired waters challenging, as they are difficult to distinguish, and it is thus problematic to correctly target mitigation efforts. Genetic markers used for microbial source tracking provide valuable information by identifying hosts that contribute to fecal loading, but do not provide a method to detect specific sources that contribute to impairment of water bodies. Spatial modelling efforts have been proposed for use in conjunction with fecal indicators and host-specific markers, but have been limited by a lack of adequate modeling for the complex processes that cause indicator decay.

We conducted a quantitative meta-analysis of published decay rate estimates for several common indicators using Bayesian hierarchical linear modeling. The meta-analysis revealed a large amount of variability across studies, including in findings of

significance for environmental parameters that impact persistence. Additionally, the meta-analysis revealed gaps in the data for genetic markers, while sufficient data was available for the traditional, culture-based indicators. We determined that temperature was consistently a significant predictor of decay rate estimates for all indicators, but light was only significant for culture-based indicators. We provided synthesized estimates for the selected indicators, but recommend caution in their application for source tracking or quantitative risk assessment due to high variability in parameter estimates and uncertainty in their extension beyond artificial settings.

We compared the decay profiles for general fecal indicators and markers associated with ruminants and cattle. We determined best fitting non-linear models based on information theory and used global model fitting to test for differences in curves for each combination of indicators. Additionally, we investigated the potential of the selected ruminant markers for use in source allocation using the ratio method, based on difference in the observed decay profiles. We found statistical differences between the decay curves of *E. coli* and all but one genetic marker. The differences across decay profiles suggest caution is necessary when interpreting microbial source tracking results using these markers, as differential decay may result in different findings depending on the marker selected.

We assessed the possibility of studying fecal indicator persistence in a truly open system using simulations. Using the concept of a Continuous-flow Stirred Tank Reactor, we developed an adjustment that can be applied to observed fecal indicator concentrations from an open system so that only loss due to decay is considered. The simulations showed that this adjustment is an effective way to account for loss for this

system. However, implementation of this system has limitations, as the removal of indicators through flow contributes to a decreased period of observations before a given indicator drops below the limits of detection.

We used the results from the simulations to design and implement an open system for decay studies. We compared decay profiles generated for several indicators from two open systems with different flow rates to those of closed and partially closed systems that have been previously used in decay studies. We used the results of these comparisons to investigate the effects of artificial settings used to study decay for fecal indicators. We found that the systems used in decay studies significantly influence the results for all indicators used.

Novel Experimental Designs and Mathematical Models to Study Fecal Indicator Bacteria Persistence in Surface Water

by
Lauren Brooks

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented August 3, 2016
Commencement June 2017

Doctor of Philosophy dissertation of Lauren Brooks presented on August 3, 2016

APPROVED:

Major Professor, representing Microbiology

Chair of the Department of Microbiology

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Lauren Brooks, Author

ACKNOWLEDGEMENTS

The author expresses sincere appreciation to the Oregon State University Department of Microbiology faculty, staff, and students for helping me immeasurably on this journey. In particular, I would like to thank my committee members, Joy Waite-Cusic, Orin Shanks, Thomas Sharpton, and Karl Schilke, for their guidance and support. Additionally, I would like to acknowledge Sarah Emerson for providing numerous impromptu statistical consultations.

I would also like to recognize my friends and family who have provided contributed greatly to my success and happiness. Particularly, thank you to Benjamin Brintz, who has been a wonderful source of support when times have been challenging.

Last but not least, I would like to say thank you to Kate Field, who has provided me with countless hours of intellectual, professional, and personal support. Words cannot express my gratitude for the guidance and encouragement Kate has provided me over the last few years.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

GENERAL INTRODUCTION, LITERATURE REVIEW, AND OBJECTIVES

Lauren E. Brooks

Fecal Contamination, Fecal Indicators, and Microbial Source Tracking

Fecal contamination is the number one cause of water impairment in the United States (USEPA, 2015). Responsible for the impairment of over 10,000 streams, fecal contamination poses a great challenge to water quality managers, health departments, and other concerned parties. Fecal matter can contain numerous pathogens leading to infections or gastrointestinal illnesses. These pathogens are typically low in abundance in water samples, making it challenging to detect individual pathogens of concern. Additionally, as a large number of pathogens need to be monitored in a body of water, it is not feasible to directly monitor for pathogens associated with fecal contamination.

The presence of fecal matter is detected through the use of non-pathogenic, abundant, and easily cultured fecal indicator bacteria (FIB), such as *E. coli* and *Enterococcus* spp. FIB have been used for decades for water quality monitoring (Geldreich, 1970), but with the rise of molecular tools, methods for enumeration of genetic markers associated with fecal indicators have been developed. Genetic markers have been identified for traditional FIB (Chern et al., 2011; Haugland et al., 2005; Ludwig and Schleifer, 2000), additional groups of indicator organisms like *Bacteroidales* (Bernhard and Field, 2000; Siefring et al., 2008, and host specific markers that provide information regarding sources contributing to the loading. Host specific genetic markers target a range of organisms, but typically rely on viruses (Fong and Lipp, 2005; Jofre et al., 2014; Ley et al., 2002), specific groups of bacteria (Bernhard and Field, 2000; Green et al., 2012; Kreader, 1995; Mieszkin et al., 2010; Schriewer et al., 2013), or functional genes (Shanks et al., 2007).

The ability to identify sources contributing to fecal contamination can help efforts to assess human-health risks. Poor correlations between traditional FIB and gastrointestinal illnesses have been reported (Colford et al., 2007; Wade et al., 2003), while stronger correlations

have been found between GI and human specific sources of fecal contamination (Schoen and Ashbolt, 2010; Soller et al., 2010). These findings support the idea that human health risks may not be as high for other animal sources of contamination (World Health Organization, 1999). Fecal contamination can be broadly divided into point and non-point sources. Point sources are identifiable, stationary sources of a contaminant, such as a wastewater treatment plant discharge or a storm drain, and are heavily regulated. In contrast, non-point sources are more challenging to manage, as they are difficult to identify and may come from a combination of many sources, including, but not limited to, wildlife, leaking septic tanks, agricultural fields, and storm water. Additionally, contributions of non-point sources are hard to quantify, as they may be transient across time and space.

A number of approaches have been developed to identify contributors of nonpoint sources of fecal matter. Traditional approaches, such as sanitary surveys, are time and labor intensive, and rely on repeated sampling of potential sources of contamination. Attempts to use microbial source tracking (MST) markers for source identification have primarily centered on listing the host species that contribute to fecal loading (Anderson et al., 2005; Chase et al., 2012; Gordon et al., 2013; Harwood et al., 2009; Lee et al., 2010; McQuaig et al., 2012; Rodriguez, 2012). Additionally, some effort has been made to address the topic of source allocation, which would rank the proportional contribution of sources found to contribute to fecal loading (Wang et al., 2013), but this method has not been applied for genetic markers outside of experimental testing.

The coupling of MST markers with spatial modeling has shown promise for identifying sources contributing to pathogen or FIB loading. Spatial models used to simulate watershed scale hydrological processes, such as nutrient or sediment transport, have been implemented to attempt

to simulate the transport of FIB or pathogens (Baffaut and Sadeghi, 2010; Benham et al., 2006; Cho et al., 2012). However, the model most often used, a variation on the Soil and Water Assessment Tool, has failed to adequately simulate FIB loading (Frey et al., 2013). If FIB and MST marker loading were modeled effectively, not only could risk assessments based on the sources of contamination be improved, but also mitigation efforts could be targeted at sources that are most likely contributing to the fecal loading.

The survival of selected indicators or pathogens in the water column has been largely neglected or overly simplified in spatial modeling attempts (Cho et al., 2016). Spatial models that incorporate a decay term into their watershed based modeling effort, assume a single decay rate for all modeled FIB (Baffaut and Sadeghi, 2010; Bougeard et al., 2011; Coffey et al., 2013; Collins and Rutherford, 2004; Gautam et al., 2006). Cho and colleagues improved on this effort by incorporating a correction term for temperature, which has been shown to improve modeling results (Cho et al., 2016), but neglected a number of other parameters that have demonstrated the potential to impact decay rates in the environment. Thus, an improved understanding of the "in stream parameters" that affect FIB and pathogen survival is necessary to further improve modeling efforts (Cho et al., 2016).

Current understanding of FIB and MST marker fate in waters

Numerous studies have identified factors that affect the fate of fecal indicators in the environment, using artificial systems intended to model the environments of interest. As fecal bacteria exit the host and enter secondary environments, they experience a dramatic change from the nutrient rich primary environment of the gut. In general, measurable fecal bacteria decline over time in secondary environments. This decline is often referred to as decay, although the exact interpretation depends on the indicator selected, as it is used both to refer to loss of

cultivability for culture based indicators, and to simply a decline in measurable markers for culture independent approaches. The extent and rate of decay in secondary environments depends largely on a combination of factors intrinsic to the bacteria, and extrinsic factors of the environment. Some environments have been found to support growth (Gerba and McLeod, 1976; Haller et al., 2009; Muirhead et al., 2005), but typically in aquatic environments decay is observed.

Variables that impact indicator decay begin in the host gut, where a number of potential differences can shape the fate of the fecal microorganisms. The concentration of indicators used to signal fecal contamination can vary within and between host species (Dick et al., 2005; Haugland et al., 2010; Layton et al., 2010; Savichtcheva and Okabe, 2006; Shanks et al., 2007) . Additionally, other factors such as diet or health of the host may influence concentrations of indicators in the gut (Claesson et al., 2012; Wu et al., 2011), thus influencing their detection and decay in the environment.

The source of the fecal bacteria has been suggested to lead to significantly different decay rates, but the reports on what fecal sources decay the fastest are conflicting. For instance, Liang and colleagues (2012) found that *E. coli* from cattle decayed faster than *E. coli* from sewage, while the opposite was found in another study for both *E. coli* and enterococci (Korajkic et al., 2013a). Anderson and colleagues (2005) reported enterococci from soils as having the lowest decay rates, followed by those sourced from sewage. In that study, enterococci sourced from dog feces had the highest decay rate. In addition to these contradictory results, no differences between enterococci decay sources from sewage influent, effluent, or urban runoff have also been reported (Noble et al., 2004).

Environmental factors that affect the decay of fecal indicators largely depend on the intrinsic properties of the selected indicator. For instance, factors affecting the detection of *Bacteroidales* survival in the environment differ from those that affect *E. coli.* Members of the order *Bacteroidales* are obligate anaerobes, meaning they are less likely to multiply in the environment, although this has been reported (Walters and Field, 2009). Some strains of FIB may be better adapted for survival outside the host, further complicating the use of indicators, as there is variation even within a single Operational Taxonomic Unit (Maraccini et al., 2012). Measurement techniques can also impact decay observations. While cell counts detect only live, cultivable cells, qPCR detects genetic markers, which can vary in concentration within the cell based on the physiological conditions of the bacterium (Ludwig and Schleifer, 2000; Muela et al., 1999).

Properties of the secondary habitat also have been reported to impact decay, but again, differences across findings make the interpretations of these results challenging. Most investigators have used artificial, controlled environments, such as in micro- or mesocosms, to examine these properties. The type of water into which the fecal bacteria are released has been shown by some to increase decay of *E. coli* (Darakas et al., 2009; Jeanneau et al., 2012; Korajkic et al., 2013a; Solecki et al., 2011) and *Enterococcus* (Anderson et al., 2005; Jeanneau et al., 2012; Solecki et al., 2011), while others have found the salinity of the water to have no effect (Ahmed et al., 2014; Noble et al., 2004). Conversely, longer persistence of human associated markers has been reported in marine waters (Green et al., 2011; Jeanneau et al., 2012), while again, others have reported no differences (Ahmed et al., 2014). Potential mechanisms for differences in observed decay in different water types include the effects of osmotic pressures on the cell membranes (Carlucci and Pramer, 1960) or potential effects of different microbial

communities (Green et al., 2011), but these mechanisms have not been tested in decay studies using natural waters.

Findings on the effects of differing temperatures on indicator decay have been much more consistent both for specific markers and across markers. In general, decreases in temperature seem to be associated with lowered decay for *E. coli* (Craig et al., 2004; Noble et al., 2004; Sokolova et al., 2012), *Enterococcus* spp. (Noble et al., 2004), and genetic markers (Kreader, 1998; Sokolova et al., 2012). However, Dick and colleagues reported no differences for the tested genetic markers or *E. coli* when comparing decay rates at 15°C and 25°C (Dick et al., 2010).

Another environmental variable that has been demonstrated to impact survival of FIB is ultraviolet radiation from sunlight. Light can directly damage DNA, resulting in loss of cellular functions, and thus increasing decay (Blatchley et al., 2001). While the effect of light on decay rates has been studied in great detail, there is little agreement across indicators. For culture based indicators *E. coli* and *Enterococcus*, light increased decay (Bae and Wuertz, 2009a; Boehm et al., 2009; Davies-Colley et al., 1994; Dick et al., 2010; Kay et al., 2005; Korajkic et al., 2013b; Noble et al., 2004; Sinton et al., 2007, 2002, 1999; Walters and Field, 2009; Whitman et al., 2004), or, in one study, had no effect (Walters and Field, 2009). However, light reportedly increased the decay of some molecular markers (Bae and Wuertz, 2009a; Green et al., 2011; Korajkic et al., 2014; Walters et al., 2009; Walters and Field, 2009), while others reported that sunlight exposure had no effects on some markers (Bae and Wuertz, 2009a; Boehm et al., 2009; Dick et al., 2010; Green et al., 2011; Sokolova et al., 2012; Walters and Field, 2009).

The effects of nutrient levels on *E. coli* and *Enterococcus* spp. have been studied previously, while data are lacking for genetic markers. Several studies have reported that

increasing nutrient concentrations in river waters did not alter decay (Lim and Flint, 1989; Noble et al., 2004), while others found that lower concentrations of nutrients such as carbon, nitrogen, and phosphorous led to decreased survival of *E. coli* (Carrillo et al., 1985; Haller et al., 2009; Hendricks, 1972). Differences in nutrient concentrations have been proposed as at least a partial explanation of different decay profiles observed for FIB from sewage versus cattle feces (Korajkic et al., 2013a), but this was not tested directly.

Interactions of the microbial community have been shown to significantly impact fecal indicator persistence. Menon and colleagues reported that grazing of predatory microorganisms was responsible for a majority of mortality in both fresh and marine waters (Menon et al., 2003). Additionally, studies that removed the indigenous water microbial community found a significant reduction in the decay of FIB (Bell et al., 2009; Korajkic et al., 2014). Thus, the activity of predators, a variable that is complex and difficult to quantify, plays a role in shaping decay in the environment as well.

A factor may only be significant in the presence or absence of another factor. For instance, Walters and Field (2009) reported that *E. coli* from cattle sources decayed faster in light, while sunlight exposure had no effect on *E. coli* sourced from sewage. Thus, while many decay studies only intentionally address one variable at a time, additional confounding measures resulting from the selection of experimental design may introduce unintentional variables into decay studies.

Statistical analyses for decay studies

Decay studies are labor-intensive and costly, making it difficult to collect and process a large number of samples. As a result, all FIB decay studies discussed above have been conducted with three or fewer biological replicates. With only three replicates, the statistical power to

compare differences in decay profiles is low, but can vary depending on the statistical tests used for comparison.

Many decay studies fit the observed data to a decay model, most commonly the Chick Watson model for log linear decay (Ahmed et al., 2014; Anderson et al., 2005; Bae and Wuertz, 2011; Bell et al., 2009; Jeanneau et al., 2012; Liang et al., 2012; Noble et al., 2004; Schulz and Childers, 2011; Sokolova et al., 2012; Tambalo et al., 2012; Walters et al., 2009; Walters and Field, 2009). Following curve fitting, some studies compare parameter estimates for different curves using a Student's t-test to test for statistical significance (Ahmed et al., 2014; Anderson et al., 2005; Bell et al., 2009; Jeanneau et al., 2012; Sokolova et al., 2012; Solecki et al., 2011) . Other studies, particularly those that find a non-log-linear decay shape, have tested for differences based on differences in individual time points through a repeated measures analysis of variance (Bae and Wuertz, 2011; Liang et al., 2012; Solecki et al., 2011). Both of these statistical methods are dependent on the selection of time points, which may influence the shape and parameter estimates if using curve fitting, or result in making comparisons on only a select few days if using repeated measures.

None of the decay studies in the literature have reported conducting power analyses to determine whether or not the methods used to assess differences between environmental variables were capable of detecting differences. Additionally, no uncertainty analyses have been reported; these would provide valuable information about the variability introduced as a result of limitations of experimental design, such as sampling frequency. Thus, it is possible that inadequate statistical methods for comparing decay across treatments have contributed to the discrepancies reported throughout the literature.

Ruminant sources of fecal contamination and ruminant associated markers

Despite the abundant literature available on FIB decay, studies using ruminant sources of fecal material are underrepresented in the available data. Ruminants are herbivorous mammals with a specialized digestive tract for processing plant-based food. Ruminants include a number of animals of agricultural significance, such as cattle, goats, and sheep, as well as wild animals, such as elk and deer. Extrapolation of decay rates estimated from other sources, such as sewage, to ruminant sources should be done with caution, as differences in *E. coli* and enterococci persistence from sewage and cattle have been reported (Korajkic et al., 2013a).

While it is believed that public health is most endangered by human sources of fecal matter, such as sewage, some studies have found similar or increased risk for illness with fecal matter from cattle. Soller and colleagues (2010) found that fecal material from cattle poses similar threats as sewage, although this varies by region and cattle type (Zhao et al., 1995). Differences in pathogen levels may coordinate with shifts in the microbial community, correlated with cattle age (Shanks et al., 2014) and diet (Shanks et al., 2011). Given the prevalence of ruminant sources of fecal contamination, and the high human health risks that result from exposure, it is essential to use reliable methods to detect ruminant sources contributing to water contamination.

A number of genetic markers have been identified to identify fecal contamination from ruminant sources (Bernhard and Field, 2000; Mieszkin et al., 2010), and their decay has been estimated in a limited number of studies (Bae and Wuertz, 2009b; Walters and Field, 2009). However, different markers were found to display different decay properties in these studies, even when directly compared to one another in the same system (Walters and Field, 2009). Cattle specific markers, that can distinguish fecal contamination from cattle and other ruminants,

have also been identified (Shanks et al., 2006), but even less data are available investigating their decay properties (Tambalo et al., 2012).

As the field of MST advances from novel techniques towards widespread use, it is necessary to streamline the methodology used for these studies, and understand the properties associated with markers selected for this standardization. Comparisons among potential markers for use in MST have addressed the sensitivity and specificity of ruminant markers (Boehm et al., 2013; Raith et al., 2013; Reischer et al., 2013). However, while these studies made recommendations for marker use based on these properties, the decay of the indicators was not included when making these recommendations. Prior to use for source tracking studies, the decay properties of these indicators should be understood.

Extending model systems to environmental processes

Ideally, decay studies could be conducted *in situ* for a given water body, so that the exact characteristics of that environment could be recreated for the study. However, aquatic systems of interest involve complex hydrological phenomena, including dispersion, mixing, and settling, making it difficult to study decay in these systems. Thus, simplified reactors are employed to model the environments of interest, and observations made in these artificial settings are assumed to be applicable to the environments of interest.

While the systems selected to study decay have varied, there has been no investigation of how the model system selected to study FIB or marker decay impacts findings. Most decay studies are conducted using a homogenously mixed microcosm consisting of a volume of water seeded with fecal matter from a selected source (Ahmed et al., 2014; Anderson et al., 2005; Green et al., 2011; Jeanneau et al., 2012; Noble et al., 2004; Schulz and Childers, 2011; Sinton et al., 2002; Sokolova et al., 2012; Solecki et al., 2011; Tambalo et al., 2012). An alternative

approach employs dialysis bags attached to a submersible mesocosm (Bae and Wuertz, 2011, 2009b, Korajkic et al., 2014, 2013a) , thus allowing for exchange of small materials with the environment. However, both of these systems constrain microorganisms within the system being studied, introducing a potentially significant difference between the model systems and the environments they are intended to model. To date, no study has explored the biases introduced by conducting decay studies in these closed systems.

Objectives

Although a large volume of research on indicator decay is available in the literature, many inconsistencies and data gaps remain, making it difficult to incorporate FIB or genetic marker decay into environmental applications, such as nonpoint source pollution models. Major limitations to such use include confounding results regarding significant environmental factors, statistical limitations, a lack of data for ruminant sources of fecal contamination, and uncertainty regarding applicability of estimates generated from artificial systems to the environments of interest. The objectives of this research were to use a combination of mathematical modeling and novel experimental designs to address these obstacles.

In the following chapters, I attempt to untangle the confusion in the literature by conducting a meta-analysis of published decay rates. Through this study, I determine trends in the available literature and find areas where more research is needed. I also use a sensitive statistical technique for decay profile comparison to detect differences in decay curves for genetic markers and FIB from dairy cattle. Finally, I develop an adjustment term that can be applied to concentrations observed for indicators in an open system, so that loss due to decay can be identified separately from loss due to removal. This enables the use of an open system to study decay, thus allowing differences in decay profiles generated from open and closed systems

to be compared. The results presented here are not only important and applicable in the field of water quality, but they have widespread implications for microbiology and environmental science in general.

Literature Cited

Ahmed, W., Gyawali, P., Sidhu, J. P. S., Toze, S., 2014. Relative inactivation of faecal indicator bacteria and sewage markers in freshwater and seawater microcosms. Lett. Appl. Microbiol. 348–354. doi:10.1111/lam.12285

Anderson, K.L., Whitlock, J.E., Harwood, V.J., 2005. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Appl. Environ. Microbiol. 71, 3041–3048. doi:10.1128/AEM.71.6.3041-3048.2005

Bae, S., Wuertz, S., 2011. Survival of host-associated *Bacteroidales* cells and their relationship with *Enterococcus* spp., *Campylobacter jejuni*, *Salmonella enterica*, *Serovar Typhimurium*, and Adenovirus in freshwater microcosms as measured by propidium monoazide-Quantitative PCR. Appl. Environ. Microbiol. 78, 922–932. doi:10.1128/AEM.05157-11

Bae, S., Wuertz, S., 2009a. Rapid decay of host-specific fecal *Bacteroidales* cells in seawater as measured by quantitative PCR with propidium monoazide. Water Res. 43, 4850–4859. doi:10.1016/j.watres.2009.06.053

Bae, S., Wuertz, S., 2009b. Discrimination of viable and dead fecal *Bacteroidales* bacteria by quantitative PCR with propidium monoazide. Appl. Environ. Microbiol. 75, 2940–2944. doi:10.1128/AEM.01333-08

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

Bell, A., Layton, A.C., McKay, L., Williams, D., Gentry, R., Sayler, G.S., 2009. Factors influencing the persistence of fecal *Bacteroides* in stream water. J. Environ. Qual. 38, 1224. doi:10.2134/jeq2008.0258

Benham, B.L., Baffaut, C., Zeckoski, R.W., Mankin, K.R., Pachepsky, Y.A., Sadeghi, A.M., Brannan, K.M., Soupir, M.L., Habersack, M.J., 2006. Modeling bacteria fate and transport in watersheds to support TMDLs. Trans. ASAE 49, 987–1002.

Bernhard, A.E., Field, K.G., 2000. A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Appl. Environ. Microbiol. 66, 4571–4574. doi:10.1128/AEM.66.10.4571-4574.2000

Blatchley, E.R., Dumoutier, N., Halaby, T.N., Levi, Y., Laîné, J.M., 2001. Bacterial responses to ultraviolet irradiation. Water Sci. Technol. J. Int. Assoc. Water Pollut. Res. 43, 179–186.

Boehm, A.B., Van De Werfhorst, L.C., Griffith, J.F., Holden, P.A., Jay, J.A., Shanks, O.C., Wang, D., Weisberg, S.B., 2013. Performance of forty-one microbial source tracking methods: a twenty-seven lab evaluation study. Water Res. 47, 6812–6828. doi:10.1016/j.watres.2012.12.046

Boehm, A.B., Yamahara, K.M., Love, D.C., Peterson, B.M., McNeill, K., Nelson, K.L., 2009. Covariation and photoinactivation of traditional and novel indicator organisms and human

viruses at a sewage-impacted marine beach. Environ. Sci. Technol. 43, 8046–8052. doi:10.1021/es9015124

Bougeard, M., Saux, J.-C.L., Teillon, A., Belloir, J., Mennec, C.L., Thome, S., Durand, G., Pommepuy, M., 2011. Combining modeling and monitoring to study fecal contamination in a small rural catchment. J. Water Health 9, 467–482. doi:10.2166/wh.2011.189

Carlucci, A.F., Pramer, D., 1960. An evaluation of factors affecting the survival of *Escherichia coli* in sea water. Appl. Microbiol. 8, 247–250.

Carrillo, M., Estrada, E., Hazen, T.C., 1985. Survival and enumeration of the fecal indicators *Bifidobacterium adolescentis* and *Escherichia coli* in a tropical rain forest watershed. Appl. Environ. Microbiol. 50, 468–476.

Chase, E., Hunting, J., Staley, C., Harwood, V.J., 2012. Microbial source tracking to identify human and ruminant sources of faecal pollution in an ephemeral Florida river. J. Appl. Microbiol. 113, 1396–1406. doi:10.1111/jam.12007

Chern, E.C., Siefring, S., Paar, J., Doolittle, M., Haugland, R.A., 2011. Comparison of quantitative PCR assays for *Escherichia coli* targeting ribosomal RNA and single copy genes. Lett. Appl. Microbiol. 52, 298–306. doi:10.1111/j.1472-765X.2010.03001.x

Cho, K.H., Pachepsky, Y.A., Kim, J.H., Kim, J.-W., Park, M.-H., 2012. The modified SWAT model for predicting fecal coliforms in the Wachusett Reservoir Watershed, USA. Water Res. 46, 4750–4760. doi:10.1016/j.watres.2012.05.057

Cho, K.H., Pachepsky, Y.A., Oliver, D.M., Muirhead, R.W., Park, Y., Quilliam, R.S., Shelton, D.R., 2016. Modeling fate and transport of fecally-derived microorganisms at the watershed scale: State of the science and future opportunities. Water Res. 100, 38–56. doi:10.1016/j.watres.2016.04.064

Claesson, M.J., Jeffery, I.B., Conde, S., Power, S.E., O'Connor, E.M., Cusack, S., Harris, H.M.B., Coakley, M., Lakshminarayanan, B., O'Sullivan, O., Fitzgerald, G.F., Deane, J., O'Connor, M., Harnedy, N., O'Connor, K., O'Mahony, D., van Sinderen, D., Wallace, M., Brennan, L., Stanton, C., Marchesi, J.R., Fitzgerald, A.P., Shanahan, F., Hill, C., Ross, R.P., O'Toole, P.W., 2012. Gut microbiota composition correlates with diet and health in the elderly. Nature 488, 178–184. doi:10.1038/nature11319

Coffey, R., Dorai-Raj, S., O'Flaherty, V., Cormican, M., Cummins, E., 2013. Modeling of pathogen indicator organisms in a small-scale agricultural catchment using SWAT. Hum. Ecol. Risk Assess. Int. J. 19, 232–253. doi:10.1080/10807039.2012.701983

Colford, J.M., Wade, T.J., Schiff, K.C., Wright, C.C., Griffith, J.F., Sandhu, S.K., Burns, S., Sobsey, M., Lovelace, G., Weisberg, S.B., 2007. Water quality indicators and the risk of illness at beaches with nonpoint sources of fecal contamination. Epidemiol. Camb. Mass 18, 27–35. doi:10.1097/01.ede.0000249425.32990.b9

Collins, R., Rutherford, K., 2004. Modelling bacterial water quality in streams draining pastoral land. Water Res. 38, 700–712. doi:10.1016/j.watres.2003.10.045

Craig, D. l., Fallowfield, H. J., Cromar, N. J., 2004. Use of microcosms to determine persistence of *Escherichia coli* in recreational coastal water and sediment and validation with *in situ* measurements. J. Appl. Microbiol. 96, 922–930. doi:10.1111/j.1365-2672.2004.02243.x

Darakas, E., Koumoulidou, T., Lazaridou, D., 2009. Fecal indicator bacteria declines via a dilution of wastewater in seawater. Desalination 248, 1008–1015. doi:10.1016/j.desal.2008.10.017

Davies-Colley, R.J., Bell, R.G., Donnison, A.M., 1994. Sunlight inactivation of Enterococci and fecal coliforms in sewage effluent diluted in seawater. Appl. Environ. Microbiol. 60, 2049–2058.

Dick, L.K., Bernhard, A.E., Brodeur, T.J., Santo Domingo, J.W., Simpson, J.M., Walters, S.P., Field, K.G., 2005. Host distributions of uncultivated fecal *Bacteroidales* bacteria reveal genetic markers for fecal source identification. Appl. Environ. Microbiol. 71, 3184–3191. doi:10.1128/AEM.71.6.3184-3191.2005

Dick, L.K., Stelzer, E.A., Bertke, E.E., Fong, D.L., Stoeckel, D.M., 2010. Relative decay of *Bacteroidales* microbial source tracking markers and cultivated *Escherichia coli* in freshwater microcosms. Appl. Environ. Microbiol. 76, 3255–3262. doi:10.1128/AEM.02636-09

Fong, T.-T., Lipp, E.K., 2005. Enteric viruses of humans and animals in aquatic environments: Health risks, detection, and potential water quality assessment tools. Microbiol. Mol. Biol. Rev. 69, 357–371. doi:10.1128/MMBR.69.2.357-371.2005

Frey, S.K., Topp, E., Edge, T., Fall, C., Gannon, V., Jokinen, C., Marti, R., Neumann, N., Ruecker, N., Wilkes, G., Lapen, D.R., 2013. Using SWAT, *Bacteroidales* microbial source tracking markers, and fecal indicator bacteria to predict waterborne pathogen occurrence in an agricultural watershed. Water Res. 47, 6326–6337. doi:10.1016/j.watres.2013.08.010

Gautam, B., Kasi, M., Lin, W., 2006. Determination of fecal coliform loading and its impact on river water quality for TMDL development. Proc. Water Environ. Fed. 2006, 3851–3874. doi:10.2175/193864706783751609

Geldreich, E.E., 1970. Applying bacteriological parameters to recreational water quality. J. Am. Water Works Assoc. 62, 113–120.

Gerba, C.P., McLeod, J.S., 1976. Effect of sediments on the survival of *Escherichia coli* in marine waters. Appl. Environ. Microbiol. 32, 114–120.

Gordon, K.V., Brownell, M., Wang, S.Y., Lepo, J.E., Mott, J., Nathaniel, R., Kilgen, M., Hellein, K.N., Kennedy, E., Harwood, V.J., 2013. Relationship of human-associated microbial source tracking markers with Enterococci in Gulf of Mexico waters. Water Res. 47, 996–1004. doi:10.1016/j.watres.2012.10.032

Green, H.C., Dick, L.K., Gilpin, B., Samadpour, M., Field, K.G., 2012. Genetic markers for rapid PCR-based identification of gull, Canada goose, duck, and chicken fecal contamination in water. Appl. Environ. Microbiol. 78, 503–510. doi:10.1128/AEM.05734-11

Green, H.C., Shanks, O.C., Sivaganesan, M., Haugland, R.A., Field, K.G., 2011. Differential decay of human faecal *Bacteroides* in marine and freshwater. Environ. Microbiol. 13, 3235–3249. doi:10.1111/j.1462-2920.2011.02549.x

Haller, L., Amedegnato, E., Poté, J., Wildi, W., 2009. Influence of freshwater sediment characteristics on persistence of fecal indicator bacteria. Water. Air. Soil Pollut. 203, 217–227. doi:10.1007/s11270-009-0005-0

Harwood, V.J., Brownell, M., Wang, S., Lepo, J., Ellender, R.D., Ajidahun, A., Hellein, K.N., Kennedy, E., Ye, X., Flood, C., 2009. Validation and field testing of library-independent microbial source tracking methods in the Gulf of Mexico. Water Res. 43, 4812–4819. doi:10.1016/j.watres.2009.06.029

Haugland, R.A., Siefring, S.C., Wymer, L.J., Brenner, K.P., Dufour, A.P., 2005. Comparison of *Enterococcus* measurements in freshwater at two recreational beaches by quantitative polymerase chain reaction and membrane filter culture analysis. Water Res. 39, 559–568. doi:10.1016/j.watres.2004.11.011

Haugland, R.A., Varma, M., Sivaganesan, M., Kelty, C., Peed, L., Shanks, O.C., 2010. Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected *Bacteroidales* species and human fecal waste by qPCR. Syst. Appl. Microbiol. 33, 348–357. doi:10.1016/j.syapm.2010.06.001

Hendricks, C.W., 1972. Enteric bacterial growth rates in river water. Appl. Microbiol. 24, 168–174.

Jeanneau, L., Solecki, O., Wéry, N., Jardé, E., Gourmelon, M., Communal, P.-Y., Jadas-Hécart, A., Caprais, M.-P., Gruau, G., Pourcher, A.-M., 2012. Relative decay of fecal indicator bacteria and human-associated markers: A microcosm study simulating wastewater input into seawater and freshwater. Environ. Sci. Technol. 46, 2375–2382. doi:10.1021/es203019y

Jofre, J., Blanch, A.R., Lucena, F., Muniesa, M., 2014. Bacteriophages infecting *Bacteroides* as a marker for microbial source tracking. Water Res. 55, 1–11. doi:10.1016/j.watres.2014.02.006

Kay, D., Stapleton, C.M., Wyer, M.D., McDonald, A.T., Crowther, J., Paul, N., Jones, K., Francis, C., Watkins, J., Wilkinson, J., Humphrey, N., Lin, B., Yang, L., Falconer, R.A., Gardner, S., 2005. Decay of intestinal enterococci concentrations in high-energy estuarine and coastal waters: towards real-time T90 values for modelling faecal indicators in recreational waters. Water Res. 39, 655–667. doi:10.1016/j.watres.2004.11.014

Korajkic, A., McMinn, B.R., Harwood, V.J., Shanks, O.C., Fout, G.S., Ashbolt, N.J., 2013a. Differential decay of enterococci and *Escherichia coli* originating from two fecal pollution sources. Appl. Environ. Microbiol. 79, 2488–2492. doi:10.1128/AEM.03781-12

Korajkic, A., McMinn, B.R., Shanks, O.C., Sivaganesan, M., Fout, G.S., Ashbolt, N.J., 2014. Biotic interactions and sunlight affect persistence of fecal indicator bacteria and microbial source tracking genetic markers in the upper Mississippi River. Appl. Environ. Microbiol. 80, 3952–3961. doi:10.1128/AEM.00388-14

Korajkic, A., Wanjugi, P., Harwood, V.J., 2013b. Indigenous microbiota and habitat influence *Escherichia coli* survival more than sunlight in simulated aquatic environments. Appl. Environ. Microbiol. 79, 5329–5337. doi:10.1128/AEM.01362-13

Kreader, C.A., 1998. Persistence of PCR-Detectable *Bacteroides distasonis* from human feces in river water. Appl. Environ. Microbiol. 64, 4103–4105.

Kreader, C.A., 1995. Design and evaluation of *Bacteroides* DNA probes for the specific detection of human fecal pollution. Appl. Environ. Microbiol. 61, 1171–1179.

Layton, B. a., Walters, S. p., Lam, L. h., Boehm, A. b., 2010. *Enterococcus* species distribution among human and animal hosts using multiplex PCR. J. Appl. Microbiol. 109, 539–547. doi:10.1111/j.1365-2672.2010.04675.x

Lee, D.-Y., Weir, S.C., Lee, H., Trevors, J.T., 2010. Quantitative identification of fecal water pollution sources by TaqMan real-time PCR assays using *Bacteroidales* 16S rRNA genetic markers. Appl. Microbiol. Biotechnol. 88, 1373–1383. doi:10.1007/s00253-010-2880-0

Ley, V., Higgins, J., Fayer, R., 2002. Bovine enteroviruses as indicators of fecal contamination. Appl. Environ. Microbiol. 68, 3455–3461. doi:10.1128/AEM.68.7.3455-3461.2002

Liang, Z., He, Z., Zhou, X., Powell, C.A., Yang, Y., Roberts, M.G., Stoffella, P.J., 2012. High diversity and differential persistence of fecal *Bacteroidales* population spiked into freshwater microcosm. Water Res. 46, 247–257. doi:10.1016/j.watres.2011.11.004

Lim, C.H., Flint, K.P., 1989. The effects of nutrients on the survival of *Escherichia coli* in lake water. J. Appl. Bacteriol. 66, 559–569.

Ludwig, W., Schleifer, K.-H., 2000. How quantitative is quantitative PCR with respect to cell counts? Syst. Appl. Microbiol. 23, 556–562. doi:10.1016/S0723-2020(00)80030-2

Maraccini, P.A., Ferguson, D.M., Boehm, A.B., 2012. Diurnal variation in *Enterococcus* species composition in polluted ocean water and a potential role for the enterococcal carotenoid in protection against photoinactivation. Appl. Environ. Microbiol. 78, 305–310. doi:10.1128/AEM.06821-11

McQuaig, S., Griffith, J., Harwood, V.J., 2012. Association of fecal indicator bacteria with human viruses and microbial source tracking markers at coastal beaches impacted by nonpoint source pollution. Appl. Environ. Microbiol. 78, 6423–6432. doi:10.1128/AEM.00024-12

Menon, P., Billen, G., Servais, P., 2003. Mortality rates of autochthonous and fecal bacteria in natural aquatic ecosystems. Water Res. 37, 4151–4158. doi:10.1016/S0043-1354(03)00349-X

Mieszkin, S., Yala, J.-F., Joubrel, R., Gourmelon, M., 2010. Phylogenetic analysis of *Bacteroidales* 16S rRNA gene sequences from human and animal effluents and assessment of ruminant faecal pollution by real-time PCR. J. Appl. Microbiol. 108, 974–984. doi:10.1111/j.1365-2672.2009.04499.x

Muela, A., Arana, I., Justo, J., Seco, C., Barcina, I., 1999. Changes in DNA content and cellular death during a starvation-survival process of *Escherichia coli* in river water. Microb. Ecol. 37, 62–69.

Muirhead, R.W., Collins, R.P., Bremer, P.J., 2005. Erosion and subsequent transport state of *Escherichia coli* from cowpats. Appl. Environ. Microbiol. 71, 2875–2879. doi:10.1128/AEM.71.6.2875-2879.2005

Noble, R.T., Lee, I.M., Schiff, K. C., 2004. Inactivation of indicator micro-organisms from various sources of faecal contamination in seawater and freshwater. J. Appl. Microbiol. 96, 464–472. doi:10.1111/j.1365-2672.2004.02155.x

Raith, M.R., Kelty, C.A., Griffith, J.F., Schriewer, A., Wuertz, S., Mieszkin, S., Gourmelon, M., Reischer, G.H., Farnleitner, A.H., Ervin, J.S., Holden, P.A., Ebentier, D.L., Jay, J.A., Wang, D., Boehm, A.B., Aw, T.G., Rose, J.B., Balleste, E., Meijer, W.G., Sivaganesan, M., Shanks, O.C., 2013. Comparison of PCR and quantitative real-time PCR methods for the characterization of ruminant and cattle fecal pollution sources. Water Res., Microbial source tracking 47, 6921–6928. doi:10.1016/j.watres.2013.03.061

Reischer, G., Ebdon, J.E., Bauer, J., Schuster, N., Ahmed, W., Åström, J., Blanch, A., Blöschl, G., Byamukama, D., Coakley, T., Ferguson, C., Goshu, G., Ko, G., De Roda Husman, A.M., Mushi, D., Poma, R., Pradhan, B., Rajal, V., Schade, M., Sommer, R., Taylor, H.D., Toth, E., Vrajmasu, V., Wuertz, S., Mach, R.L., Farnleitner, A.H., 2013. Performance characteristics of qPCR assays targeting human- and ruminant-associated *Bacteroidetes* for microbial source tracking across sixteen countries on six continents. Environ. Sci. Technol. 130611111208000. doi:10.1021/es304367t

Rodriguez, S., 2012. Microbial source tracking: determination of animal sources of enterococci in Oso Creek, Nueces County, Texas.

Savichtcheva, O., Okabe, S., 2006. Alternative indicators of fecal pollution: Relations with pathogens and conventional indicators, current methodologies for direct pathogen monitoring and future application perspectives. Water Res. 40, 2463–2476. doi:10.1016/j.watres.2006.04.040

Schoen, M.E., Ashbolt, N.J., 2010. Assessing pathogen risk to swimmers at non-sewage impacted recreational beaches. Environ. Sci. Technol. 44, 2286–2291. doi:10.1021/es903523q

Schriewer, A., Goodwin, K.D., Sinigalliano, C.D., Cox, A.M., Wanless, D., Bartkowiak, J., Ebentier, D.L., Hanley, K.T., Ervin, J., Deering, L.A., Shanks, O.C., Peed, L.A., Meijer, W.G., Griffith, J.F., SantoDomingo, J., Jay, J.A., Holden, P.A., Wuertz, S., 2013. Performance evaluation of canine-associated *Bacteroidales* assays in a multi-laboratory comparison study. Water Res., Microbial source tracking 47, 6909–6920. doi:10.1016/j.watres.2013.03.062

Schulz, C.J., Childers, G.W., 2011. Fecal *Bacteroidales* diversity and decay in response to variations in temperature and salinity. Appl. Environ. Microbiol. 77, 2563–2572. doi:10.1128/AEM.01473-10

Shanks, O.C., Atikovic, E., Blackwood, A.D., Lu, J., Noble, R.T., Domingo, J.S., Seifring, S., Sivaganesan, M., Haugland, R.A., 2007. Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. Appl. Environ. Microbiol. 74, 745–752. doi:10.1128/AEM.01843-07

Shanks, O.C., Kelty, C.A., Archibeque, S., Jenkins, M., Newton, R.J., McLellan, S.L., Huse, S.M., Sogin, M.L., 2011. Community structures of fecal bacteria in cattle from different animal feeding operations. Appl. Environ. Microbiol. 77, 2992–3001. doi:10.1128/AEM.02988-10

Shanks, O.C., Kelty, C.A., Peed, L., Sivaganesan, M., Mooney, T., Jenkins, M., 2014. Age-related shifts in the density and distribution of genetic marker water quality indicators in cow and calf feces. Appl. Environ. Microbiol. 80, 1588–1594. doi:10.1128/AEM.03581-13

Shanks, O.C., Santo Domingo, J.W., Lamendella, R., Kelty, C.A., Graham, J.E., 2006. Competitive metagenomic DNA hybridization identifies host-specific microbial genetic markers in cow fecal samples. Appl. Environ. Microbiol. 72, 4054–4060. doi:10.1128/AEM.00023-06

Siefring, S., Varma, M., Atikovic, E., Wymer, L., Haugland, R.A., 2008. Improved real-time PCR assays for the detection of fecal indicator bacteria in surface waters with different instrument and reagent systems. J. Water Health 6, 225–237. doi:10.2166/wh.2008.022

Sinton, L., Hall, C., Braithwaite, R., 2007. Sunlight inactivation of *Campylobacter jejuni* and *Salmonella enterica*, compared with *Escherichia coli*, in seawater and river water. J. Water Health 5, 357–365. doi:10.2166/wh.2007.031

Sinton, L.W., Finlay, R.K., Lynch, P.A., 1999. Sunlight inactivation of fecal bacteriophages and bacteria in sewage-polluted seawater. Appl. Environ. Microbiol. 65, 3605–3613.

Sinton, L.W., Hall, C.H., Lynch, P.A., Davies-Colley, R.J., 2002. Sunlight inactivation of fecal indicator bacteria and bacteriophages from waste stabilization pond effluent in fresh and saline waters. Appl. Environ. Microbiol. 68, 1122–1131.

Sokolova, E., Åström, J., Pettersson, T.J.R., Bergstedt, O., Hermansson, M., 2012. Decay of *Bacteroidales* genetic markers in relation to traditional fecal indicators for water quality modeling of drinking water sources. Environ. Sci. Technol. 46, 892–900. doi:10.1021/es2024498

Solecki, O., Jeanneau, L., Jardé, E., Gourmelon, M., Marin, C., Pourcher, A.M., 2011. Persistence of microbial and chemical pig manure markers as compared to faecal indicator bacteria survival in freshwater and seawater microcosms. Water Res. 45, 4623–4633. doi:10.1016/j.watres.2011.06.012

Soller, J.A., Bartrand, T., Ashbolt, N.J., Ravenscroft, J., Wade, T.J., 2010. Estimating the primary etiologic agents in recreational freshwaters impacted by human sources of faecal contamination. Water Res. 44, 4736–4747. doi:10.1016/j.watres.2010.07.064

Tambalo, D.D., Fremaux, B., Boa, T., Yost, C.K., 2012. Persistence of host-associated *Bacteroidales* gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. Water Res. 46, 2891–2904. doi:10.1016/j.watres.2012.02.048

USEPA, 2015. Assesment TMDL tracking and implementation system (ATTAINS), on National Water Quality Inventory Report to Congress.

Wade, T.J., Pai, N., Eisenberg, J.N.S., Colford, J.M., 2003. Do U.S. Environmental Protection Agency water quality guidelines for recreational waters prevent gastrointestinal illness? A systematic review and meta-analysis. Environ. Health Perspect. 111, 1102–1109.

Walters, S.P., Field, K.G., 2009. Survival and persistence of human and ruminant-specific faecal *Bacteroidales* in freshwater microcosms. Environ. Microbiol. 11, 1410–1421. doi:10.1111/j.1462-2920.2009.01868.x

Walters, S.P., Yamahara, K.M., Boehm, A.B., 2009. Persistence of nucleic acid markers of health-relevant organisms in seawater microcosms: Implications for their use in assessing risk in recreational waters. Water Res. 43, 4929–4939. doi:10.1016/j.watres.2009.05.047

Wang, D., Farnleitner, A.H., Field, K.G., Green, H.C., Shanks, O.C., Boehm, A.B., 2013. Enterococcus and Escherichia coli fecal source apportionment with microbial source tracking genetic markers – Is it feasible? Water Res. doi:10.1016/j.watres.2013.02.058

Whitman, R.L., Nevers, M.B., Korinek, G.C., Byappanahalli, M.N., 2004. Solar and temporal effects on *Escherichia coli* concentration at a Lake Michigan swimming beach. Appl. Environ. Microbiol. 70, 4276–4285. doi:10.1128/AEM.70.7.4276-4285.2004

World Health Organization, 1999. Health based monitoring of recreational waters: the feasibility of a new approach (the“ Annapolis Protocol”): outcome of an expert consultation, Annapolis, USA.

Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.-Y., Keilbaugh, S.A., Bewtra, M., Knights, D., Walters, W.A., Knight, R., Sinha, R., Gilroy, E., Gupta, K., Baldassano, R., Nessel, L., Li, H., Bushman, F.D., Lewis, J.D., 2011. Linking long-term dietary patterns with gut microbial enterotypes. Science 334, 105–108. doi:10.1126/science.1208344

Zhao, T., Doyle, M.P., Shere, J., Garber, L., 1995. Prevalence of enterohemorrhagic *Escherichia coli* O157:H7 in a survey of dairy herds. Appl. Environ. Microbiol. 61, 1290–1293.

CHAPTER 2

BAYESIAN META-ANALYSIS TO SYNTHESIZE DECAY RATE CONSTANT ESTIMATES FOR COMMON FECAL INDICATOR BACTERIA

Lauren E. Brooks and Katharine G. Field

Water Research 104: 262-271

Abstract

For decades, fecal indicator bacteria have been used as proxies to quantitatively estimate fecal loading into water bodies. Widely used cultured indicators (e.g. *E. coli* and *Enterococcus* spp.) and more recently developed genetic markers are well studied, but their decay in the environment is still poorly understood. We used Hierarchical Bayesian Linear Modeling to conduct a series of meta-analyses using published decay rate constant estimates, to synthesize findings into pooled estimates and identify gaps in the data preventing reliable estimates. In addition to the meta-analysis assuming all estimates come from the same population, meta-regressions including covariates believed to contribute to decay were fit and used to provided synthesized estimates for specific combinations of significant variables. Additionally, statements regarding the significance of variables across studies were made using the 95% confidence interval for meta-regression coefficients. These models were used to construct a mean decay rate constant estimate as well as credible intervals for the mean and the distribution of all likely data points. While synthesized estimates for each targeted indicator bacteria were developed, the amount of data available varied widely for each target, as did the predictive power of the models as determined by testing with additional data not included in the modeling. Temperature was found to be significant for all selected indicators, while light was found to be significant only for culture-based indicators. Results from the models must be interpreted with caution, as they are based only on the data available, which may not be representative of decay in other scenarios.

1. Introduction

Differences between the persistence of fecal indicator bacteria (FIB) used to detect the presence of fecal contamination and enteric pathogens in water could limit the ability to accurately predict public exposure and health risks. As a result, the public could be exposed to pathogens at higher levels than predicted using these measures, or, conversely, unwarranted precautions such as closures to the public could be implemented. Having an accurate range of plausible values for the decay of FIB in the environment is vital to the success of their use as proxies for assessing contamination.

In addition to cultured *E. coli* and *Enterococcus spp.*, genetic markers detected by quantitative Polymerase Chain Reaction (qPCR) are used with increasing frequency. Genetic markers provide advantages over culture-based enumeration, including a shorter period between sample collection and quantification (Dick and Field, 2004; Haugland et al., 2005). Additionally, host specific markers have been identified that are able to distinguish host species of origin for microbial source tracking (Bernhard and Field, 2000). As the use of genetic markers for microbial source tracking increases, it has become important to understand decay of these markers as related to traditional FIB and pathogens as well as to each other for source allocation (Wang et al., 2013).

Previous studies have investigated the decay of FIB in controlled environments (Anderson et al., 2005; Bae and Wuertz, 2009, 2015; Bell et al., 2009; Dick et al., 2010; Green et al., 2011). Additionally, specific environmental conditions, such as sunlight, salinity, temperature and predation, have been studied to determine their effect on decay of FIB and genetic markers (e.g. Bell et al., 2009; Korajkic et al., 2014; Okabe and Shimazu, 2007; Schulz and Childers, 2011; Walters and Field, 2009). While these studies have provided valuable information, major

disagreements among both decay rate constant estimates (Figure 2.1) and the significance of environmental conditions make it difficult to forecast how FIB will persist in the environment.

Meta-analysis is a statistical approach for synthesizing prior studies estimating the same parameter (Sutton and Abrams, 2001). Meta-analyses can include defined study characteristics (fixed effects), unexplained variance (random effects), or a combination of the two (mixed effects). Fixed effects meta-analyses assume that studies to be synthesized not only estimate the same parameter, but that all studies are exchangeable as estimates of that parameter. Conversely, random effects models assume no exchangeability between studies. By using a mixed effects model, it is possible to assign variance in the data to predictor variables, leaving remaining variation accounted for by the undefined random effects.

The Bayesian approach to meta-analysis allows for, and can explicitly model, parameter uncertainty. Under the Bayesian paradigm, both the data and the parameters in the model are treated as unknowns with their own distributions. Using Bayes' Theorem, the likelihood function, which defines the plausibility of the data values given the model parameters, is combined with prior estimates and used to construct the posterior credible interval. Unlike Frequentist confidence intervals, direct probability statements about the posterior distribution can be made, allowing for easier interpretation of the posterior credible interval (Thompson and Higgins, 2002).

This study used Bayesian hierarchical linear models to analyze and synthesize existing decay rate constant estimates for common FIB. We selected FIB targets of significance for water quality monitoring or microbial source tracking for which sufficient data were available for synthesis via meta-analysis. Fixed effects meta-analysis models for the general decay rate constant, excluding predictor variables, provided a synthesized estimate of decay rate constants

for the selected indicators of fecal contamination. Mixed effects meta-regression models, including predictor variables provided in the description of each published study, were used to determine which variables included in the model were significant. Synthesized estimates for decay rate constants from combinations of significant variables were also generated. Our objectives were to provide synthesized decay rate constant estimates, improve our understanding of decay rates by determining what variables significantly alter decay rate estimates, and identify gaps in the current data that limit applications of FIB and molecular markers.

2. Methods

2.1 Literature Search and Data Selection

Two classic indicators, cultivable *E. coli* and *Enterococcus spp.*, were selected for this analysis, as they have been used for decades and have been the focus of many studies (e.g. Kay et al., 2005; Noble et al., 2004; Sinton et al., 2002). More recently, *Bacteroides* associated markers have become the focus of a number of studies, as both indicators of general contamination, and to distinguish sources of contamination (Ahmed et al., 2014; Green et al., 2011; Jeanneau et al., 2012; Tambalo et al., 2012; Walters and Field, 2009). To ensure enough data points were available, *Bacteroides* associated markers from the same general hosts (e.g. "*Bacteroides* ruminant- associated markers" or "*Bacteroides* human-associated markers") were grouped together for the meta-analyses, although they do not necessarily target the same phylogenetic groups or match the same coverage within these clades (see Table S2 for complete list of primers included).

Data from the literature were compiled for the selected FIB or genetic markers using databases available through the Web of Science citation indexing service. The literature search

was conducted using a combination of the search terms “Fecal Indicator Bacteria” or “Microbial Source Tracking” combined with either “persistence”, “decay”, or “inactivation”, and “water”. This initial screening returned 411 results (Supplemental Table 2.1). To assess inclusion for the meta-analysis, titles and abstracts for all papers were reviewed and analyzed to determine if they included original data relevant to the decay of FIB or microbial source tracking markers in water bodies. Following this initial screening, full papers were read and screened for inclusion based on the following criteria:

Experimental Design Selected papers were screened based on several key features of the experimental design. As the purpose of this meta-analysis was to synthesize results for the decay rate constants of FIB as indicators of fecal contamination, we included studies that used fecal material, including sewage influent and effluent as well as raw fecal material, as a spike source and excluded studies that seeded waters using laboratory strains or isolated environmental strains. Additionally, studies that tested persistence under artificial pressures, such as chlorine, were excluded. Finally, the persistence data reported in a study must have been from a natural water body or from a microcosm that was constructed using natural waters. While microcosms that included sediment were included, data collected from the sediments were not included in this analysis.

Data Format Papers that graphically represented data but did not report any quantitative analysis (e.g. decay rate constants) were excluded from this meta-analysis. While these papers qualitatively provide valuable information, it could not be used for synthesis performed here. Additionally, to be included in the models, first order decay rate estimates (k) must have been presented in the form $k = \ln(C_0/C_t)$ in units of day^{-1} with corresponding estimates of variance (e.g. standard error, 95% confidence interval) for each data point. Studies presenting first-order

decay rates in other units (e.g. $hour^{-1}$) were transformed into compatible format. Papers that did not present data in the form of the first order decay model or did not provide variance estimates, but did provide a value for T90 or T99, the time necessary for 90% or 99%, respectively, of the initial measurement to decay, were used to assess the ability of the model to predict data not used to generate the model.

Minor data conversions such as unit conversions were performed both for decay rate estimates and for metadata. Major changes to data, however, such as calculating a first order decay rate from a T90 value, were not performed in order to respect the initial decisions of the authors not to represent the data using a first order decay model.

Metadata Another criteria for inclusion in this study was the reporting of metadata. Data were compiled for a number of possible covariates (Supplemental Table 2.2), but few studies included all possible categories of metadata. For mixed effects models, predictor variables presented in the original papers were selected to explain variance. Variables that have been suggested to impact decay of FIB in the environment (light, salinity, temperature) were included in the models to assess significance across studies. In addition to environmental conditions that have been tested, variables for experimental design effects (sampling duration, size of experiment) were included to measure the significance of any unintended variables as a result of experimental design. Categorical indicator variables for each of these terms were constructed and used as terms in the meta-regression model.

For quantitative variables, such as temperature or length of experiment, categories were constructed by identifying breaks in the data points using the ClassIntervals package in R (Bivand et al., 2015). For temperature, breaks were identified at 16°C and 30°C, providing categorical variables for low temperature < 16°C, mid temperature between 16 and 30°C, and

high temperature > 30°C. Duration of experiment was similarly categorized (breaks at 7 days and 15 days) as was size (breaks at 10 L and 50 L). Indicator variables were constructed for each category and used in the mixed effects model.

2.2. Model definition

All models were constructed as Normal-Normal Hierarchical models, meaning that an assumption was made that all estimates came from a normally distributed population. The Central Limit Theorem states that estimates of a mean will be normally distributed, even if the values themselves do not come from a normal distribution. To check for normality, data for each individual target were assessed using histograms and QQ-plots (not shown). Outliers ($k > 3$ day^{-1}) were detected for *E. coli* and *Enterococcus spp.* but were not removed from analyses.

2.2.1. Fixed Effects Meta-analysis

Models for each target were constructed in which the observed decay rate estimates were the first level parameters, defined as coming from a second level distribution with mean μ and variance τ (Equation 1). Weakly informative normal prior densities with mean = 0.5 day^{-1} and high variance (variance = 10 day^{-1}) values were assigned to μ. A non-informative prior gamma density was assigned for τ estimates in all models, and values were estimated using the posterior density generated from the output of the model (Equation 1). All models were coded using the r2jags package in RStudio (Su and Yajima, 2014) and output for each model was visualized using the mcmcplots package in RStudio (Goldin, 2012).

Equation 1: Meta-analysis (no-covariates)

$$y_i \sim Norm(\mu, \tau)$$

Prior Densities:

$$\mu \sim \text{Norm}(0.5 , 10)$$

$$\tau \sim Gamma(0.001, 0.001)$$

2.2.2. Meta-Regression

Models for meta-regression were defined similarly to the meta-analysis, with the exception that data from a specific combination of variables were assigned to come from unique distributions, with mean θ_i and variance τ_i (Equation 2). A unique theta value was given for each combination of predictor variables (denoted by i), representing the higher-level distribution of the observed data. Non-informative priors were used for all parameters estimated in the model to provide unbiased estimates of θ. Variables to be included in the final model were determined based on significance of the coefficients (β values). If the 95% confidence interval for coefficients included zero, variables were not considered significant, and were left out of further analyses to model decay rates under specific conditions.

Equation 2: Meta-regression (covariates included)

$$y_i \sim Norm(\theta_i, \tau)$$

$$\theta_i = \beta_0 + \beta_j \times x \qquad \text{where } x \text{ is a predictor variable}$$

Prior Densities:

$$\theta_i \sim \text{Norm}(0.5, 10)$$

$$\tau \sim Gamma(0.001, 0.001)$$

$$\beta_j \sim Norm(0, 1000)$$

2.3. Model Assessment

Convergence was assessed using the Gelman-Rubin Statistic (Gelman and Rubin, 1992) produced by R2jags for all parameters in all models. A value of one was used to indicate

complete convergence. Additionally, plots produced with the mcmcplots package were viewed for each parameter to visually assess convergence and autocorrelation. To ensure convergence and minimize auto-correlation, a burn-in value (the number of iterations before data were collected) of 1000, a thinning value of 30 (1 of every 30 values was used in the model), and a total of 100,000 iterations were used for all models. Deviance Information Criterion (DIC) values produced by R2jags, a measure of deviance in the model, were assessed to compare models. A lower DIC value indicates lower deviance, and was used to select the best fitting models.

The fit of each model was assessed by posterior predictive checking using the data points included in the models (Meng, 1994). Additionally, the predictive intervals generated from the model were constructed by creating simulations for the output from the models, and tested with additional T90 and T99 values that had not been included in the original analysis. The I^2 statistic (Higgins et al., 2003), presented as a percentage of variation that is not explained in the model, was used to assess heterogeneity within estimates for a given parameter.

3. Results

3.1. Overall Model Performance

Models for all targets were successful in reaching convergence with minimal autocorrelation. DIC values were lower for all mixed effects models when compared to the corresponding fixed effect model for each target, suggesting mixed effects models including covariates fit the data better even when accounting for a penalty based on inclusion of additional parameters. Posterior checking showed that for most models, all data used to construct the model fell within the posterior distribution. Models constructed for *Enterococcus spp.* and one of the combinations

for *Bacteroides* Human-Associated Markers failed to encompass all constructive data points, each with one point lying outside of the credible interval.

Fixed effects models resulted in output for posterior intervals for the mean that were nearly identical to a simple averaging of the data. The posterior predictive distribution for all data was wider than ranges estimated for the mean in all cases (Table 2.2). The model definition for the mixed effects model used a shared variance term among all combinations of significant variables for a target FIB, resulting in different distributions for estimates of the mean decay rate of specific combinations as compared to simple frequentist averaging (data not shown). These differences were not consistent across targets, with some models providing wider confidence intervals for the mean than frequentist estimates, while others resulted in narrower ranges of credible estimates.

3.2. Model Performance by Target

3.2.1. E. coli

Because *E. coli ha*s been used as an indicator for decades, there are many studies focusing on decay in the environment. Of these studies, 21 data points from 6 papers reported first order decay rates with estimates of variance and could therefore be included in the model. Two data points were notably higher than other values and were not included in the final models (see Discussion), resulting in 19 data points being used to construct the model. Additionally, 21 data points from 7 papers presenting either T90 or T99 values were used to assess the model fit.

Estimates of distributions for the mean and credible distribution centered on the mean value of $\mu = 0.74$ day^{-1}, but the credible interval had a much wider range (Figure 2.2). Posterior checking suggested the model fit the data well, with 100% inclusion of points used to construct

the model. However, using data points that were not included in the model, only 43% of the 21 data points fell within the posterior credible distribution.

Incorporation of covariates for the mixed effects model suggested that temperature was the best predictor variable for decay rate. Additionally, light was included as a variable in the final model, as inclusion resulted in a lower DIC value, although it was not statistically significant if the outliers were not included in the model. Increased temperature and the inclusion of light both resulted in an increased rate of decay, although temperature had a larger effect on the decay rate than light. Inclusion of predictor variables explained some of the heterogeneity in the data points, as measured by lower I^2 values, but large amounts of variance remained unexplained. Posterior checking of the model using T90 and T99 values not included in the model had mixed results, with a range of 20%- 67% inclusion of points used to assess fit (Table 2.2).

3.2.2. Enterococcus spp.

Like *E. coli*, *Enterococcus spp.* has been used as an indicator for decades, resulting in many studies addressing decay after leaving the host. Sixteen data points from 6 papers were used to construct the model (Table 2.1) and an additional 28 data points from 7 papers were used to assess the predictive power of the model.

The fixed effects meta-analysis resulted in wide credible intervals for both the mean (μ = 0.84 day^{-1}) and for the distribution (-0.5 to 2.18). Large amounts of variation in the data resulted in a high variance term, leading to negative decay rate constant values in the credible interval. Only 94% of the data points used to construct the model fell within this wide posterior interval, and only 54% of T90/T99 data fell within this wide interval, with all points above the estimated T90/T99 ranges, suggesting the model overestimated decay rate.

Using covariates to explain variance indicated that both temperatures and exposure to light were significant predictors of decay rate for *Enterococcus spp*. Explanation of variance as indicated by the I^2 statistic was slightly improved with the inclusion of these predictor variables, but heterogeneity remained high. Posterior predictive checking showed that the model did a good job of predicting estimates for both temperature categories in the dark, with 100% of data points used to test the model falling within the credible interval, while no data points in the light category for either temperature category fell within the posterior distribution. Additional data points were not included in the posterior checking as they were generated in the high temperature category, which was not assessed in the model.

3.2.3. General Bacteroides *Associated Markers*

As a newer indicator, qPCR primer sets targeting *Bacteroides* Associated Markers had fewer data available for inclusion in the model. Only 11 data points from 2 studies were available to construct the model, while 19 data points from 5 studies were used for assessing the model fit. The data that were available were limited in scope (Table 2.1); only one data point included light, and no studies conducted in saline water fit the requirements to be used to construct the model. Additionally, each of the studies used to construct and test the model relied on different primer sets, making it impossible to test the effects of combining these markers into one target.

Possibly as a result of these limits to the available data, the fixed effects model provided relatively narrow confidence intervals for the mean (0.71 day^{-1}) and distribution (0.19 to 1.24), relative to those constructed for *E. coli* and *Enterococcus spp.* (Figure 2.2). All data points used to construct the model fell within the posterior distribution, and the model output distribution included 74% of T90/T99 data points used to test the model.

The mixed effects model showed that temperature category was a significant predictor of decay rate for General *Bacteroides* Associated Markers. The model output showed a significant difference between data points between the two studies, but the source of that difference (e.g., primer targets, experimental design) was impossible to identify given limitations in the data. As the source of this variation was not able to be determined, temperature was used as the only significant variable for the mixed effects model even though this resulted in a higher DIC for the model. Using temperature alone, the model output included three data points available for checking the model in the low temperature category, but could only predict 33% of data points in the mid temperature range.

3.2.4. Bacteroides *Human-Associated Markers*

As was the case for general *Bacteroides* associated markers, *Bacteroides* human-associated markers have been the focus of fewer studies than the traditional FIB. A total of 12 data points from 5 studies were available for construction of the model, using primer sets BacH, BacHum-UCD, and HF183. Additionally, 16 data points from 6 studies were used to check the fit of the model.

The fixed effects model provided a tight interval for the mean (0.95 day^{-1}), but a wide distribution for the credible interval (-0.08, 1.99), including negative values for k. In spite of this wide interval, only 56% of the data available for checking fell into the T90/T99 boundaries defined by the model output.

Inclusion of predictor variables again found that temperature was the most significant variable for predicting decay rates, although all but 2 of the data points used to construct the model fell into the low temperature category. The two low temperature estimates were very close, resulting in a very low amount of heterogeneity ($I^2 = 0$). However, with the inclusion of

temperature as a predictor variable, fewer data points used to check model fit fell inside the boundaries predicted by the corresponding model, with no data points from the low temperature category falling within the credible interval and only 62% of the mid temperature category falling into the posterior interval.

3.2.5. Ruminant Associated Bacteroides *Markers*

While there were enough data from Ruminant Associated *Bacteroides* markers to construct the model, only 7 data points were available to assess the model fit. Data consisted of 12 data points from 3 different papers, using four different primer sets targeting different markers. The variables included in the model were limited by the fact that all studies were conducted in fresh water, making it impossible to include salinity as a potential significant variable.

The posterior interval for the mean (0.76) was narrow while the credible distribution was wide (-0.14, 1.65), again including negative values as part of the distribution credible interval. Inclusion of temperature, the only significant predictor variable, failed to either narrow the credible intervals, or improve the remaining heterogeneity. Additionally, while 5 of the 7 data points available for checking the fit of the model fell within the bounds of the fixed effects model output, the mixed effects models did a poor job of predicting the additional data points with only 66% fitting from the low temperature category and 50% fitting from the mid temperature category.

4. Discussion

4.1. Model limitations

While meta-analysis can be a useful tool to synthesize estimates of a parameter, interpretations should be made with caution and an awareness of the limitations. First, this study

was hampered by the reliance on first order decay rates. First order decay rates were selected for this study because the majority of papers report decay rate estimates in the first order format. First order decay rates are readily interpreted, as there is only one parameter, yet they are not always the best fitting model for describing the data (Bae and Wuertz, 2009, 2011; Green et al., 2011). When non-linear data are forced into the first order decay model, there is potential not only to present a poorly fitting model, but also to lose valuable information. For instance, many studies report a shoulder or lag period prior to decay that may be related to the health of the microbial community (Green et al., 2011). For this reason, we opted not to calculate first order decay rates based on a provided T90/T99 value in papers that observed non-linear decay curves, and instead used only the data calculated and presented by the original authors. By selecting only first order decay rates in this model, it is possible that data are biased towards estimates that fit linear models best, possibly explaining the poor fit of data not used to construct the models.

In addition to the limited data available for construction of the models, in order to synthesize papers to measure for significant covariates metadata provided by the original authors was necessary. While all papers included provided information with respect to temperature, other potential covariates were not provided in all or were provided without quantitative information. For instance, due to reporting methods of the original papers, light was defined here as a binary variable (i.e. light or dark). However, as light intensity has a wide range within the "natural light" category, this may not be an accurate reflection of the true experimental design. However, few studies reported data for light intensity, making it necessary to treat light as a categorical variable. Another inherent limitation in the construction of the models presented here is the grouping of similar targets as though they were exchangeable. This merging was done primarily as a way to collect enough data for synthesis, but may introduce additional variance into the data

that would not exist for a single marker. For culture based methods, different enumeration techniques were used (e.g. Colilert and plate counts), which may introduce discrepancies in the data, but generally target overlapping or closely related organisms. However, for the genetic markers, discrepancies between the coverage are more prevalent. While all of the markers selected for this study focus on a broad class (i.e. "*Bacteroides* Human-Associated Markers") there are known differences among the coverage of taxa that amplify with each primer set (Dick et al., 2005; Kildare et al., 2007; Layton et al., 2006). While sequences targeted by these primer sets fall within the phylum *Bacteroidetes*, discrepancy exists between the coverage of sequences within this taxonomic grouping, and possible biological differences between targets are poorly understood.

In addition to the limitations in collecting data used to construct the model, there are limitations within the model definition. The variance term for combinations in the meta-regression analysis was defined as the same for all combinations, leading to wide credible intervals even when the data appeared to be tight (e.g. *Bacteroides* Human-Associated Markers – Low Temperature). As defined, the models assume that all combinations are actually from the same population, rather than different populations with their own variance terms. While this makes the credible interval wider, it could actually be more appropriate to assume the same variance for all groups, even if the data available were tightly bound.

4.2. Posterior Distributions and Model Checking

4.2.1. Posterior Distributions

Output from the models provided mean decay rate constant estimates and credible ranges for both the mean values as well as the distribution. While the ranges of posterior intervals for the mean decay rate constants were smaller, credible intervals for the distribution resulted in wide

ranges for all targets in both fixed effects and mixed effects models. The broad credible intervals reflect the high amounts of variation in estimates from the literature, and were often only slightly improved by the incorporation of predictor variables. Even using these broad distribution ranges, it is clear that the credible ranges determined by the models encompass different values, and thus the targets likely experience differences in decay. This becomes more evident as the credible ranges for decay rate constants are transformed into T90 and T99 credible ranges (Figure 2.3).

While having a wide credible range may not be useful for predicting future values, narrower ranges for credible intervals of the mean could be useful for predicting average values. Such averages have potential for use in source tracking and allocation. The ratio method of source apportionment was effective only when contamination was fresh, while variation in decay over time led to difficulties in correctly assigning relative source amounts (Wang et al., 2013). It is possible that using average values, such as those presented here, may be useful for source apportionment, although future work is needed to test this idea.

4.2.2. Posterior Checking

Although model output for credible intervals provided wide ranges, in many cases the posterior distribution was not able to predict values checked by T90 and T99 estimates from outside studies. One possible explanation of this poor predictive power was the use of data points that came from data fit to different models, suggesting that even by using T90/T99, which can be compared across different models, the linear decay model did a poor job of predicting other models. However, the poor predictive power of the models can also be explored to draw additional conclusions that cannot be ascertained mathematically using the models.

Two papers (Sinton et al., 2007, 2002) contained data points that fell outside of the credible ranges, likely due to experimental design not accounted for in the model. T90 estimates

presented in these papers for both *Enterococcus spp.* and *E. coli* in natural light fell below the credible range for fixed effects models as well as the corresponding mixed effects models. Samples for both papers were collected within a period of hours following the addition of the fecal spike, in contrast to days or weeks long experiments for most of the other papers. While duration of the experiment was one of the variables tested, none of the studies used to construct the model consisted of such a short duration, possibly explaining why the model output failed to encompass these data (Figure 2.3).

In addition to the data collected from experiments using a short sampling scheme, there were two other papers that contained points that were not predicted by the data. One paper (Noble et al., 2004) reported data points that mostly fit well within the model, with the exception of data from one experiment. These data points were identified as outliers above, and had decay rate constant estimates considerably higher than those reported elsewhere. For *E. coli*, only the data points that were tested under "High Solar Irradiation" fell outside of the credible ranges, while for *Enterococcus spp.*, data points collected in this experiment fell outside of the credible ranges regardless of light treatment (data points collected under both low and high solar irradiation).

While there are possible explanations for the lack of fit for the papers mentioned above, additional data points presented in the paper by Dick et al. (2010) also consistently fell outside of the ranges for the corresponding targets. Data points presented in this study from experiments where the microbial community was removed from the receiving water fell into credible ranges produced from the models, while all other data points from non-sterile water estimated significantly lower T99 estimates (higher decay) than estimated by the model output. Unlike the other estimates that fell outside the credible ranges, no additional explanatory variables were

reported that could help explain this discrepancy, suggesting an unidentified difference between this study and others incorporated in the model.

For the genetic markers, limited data were available to assess the fit of the models, and for General *Bacteroides* Associated Markers and *Bacteroides* Human-Associated Markers, most of the data used to check the model fit failed to fall within the credible ranges of the model. However, as there was little other available data to check the posterior fit of the model, it is difficult to assess whether these data points were truly different or if the model just did a poor job of predicting credible distributions. Future testing and additional studies for these genetic markers are necessary to truly assess significant variables impacting decay of these targets in the environment.

4.3. Variables

One of the main objectives of this paper was to determine which predictor variables have a significant role in predicting the rate of decay. Mixed effects models relied on covariate information presented in the papers, and thus were limited to the variables that were reported in all or most papers. Potentially relevant environmental parameters such as DO, nutrients, and TSS were not reported in the majority of papers, and were therefore not included in the analysis. Additionally, for some of the variables tested, only one study had different values for several variables; as a result, the source of any explained variance could not be identified. For example, in the *E. coli* model, data points from studies reporting inclusion of sediment, small volume, and short experiment were all the same data points, making detection of significance more difficult, as these data points only represent that particular combination. For many of the models, there were not enough data available to test the significance of some variables, leaving unanswered questions that need to be addressed in future studies.

Some variables that have been shown to influence decay were left out of this model due to restrictions for the study inclusion. For example, the microbial community has been shown to play a major role in the removal of FIB (Korajkic et al., 2014a). Most studies measuring this effect have compared the differences when the environmental microbial community is removed from the system, thereby showing that is the microbial community is significant, but not providing information with respect to the mechanisms contributing to the significance of the microbial community. As a result, while the microbial community of the receiving water is likely to play a large role in the removal of FIB, it cannot be incorporated into modelling efforts as anything other than a binary variable (microbial community present or absent) without a better understanding of how the microbial community contributes to decay.

Temperature was a significant predictor of decay rate constants for all models tested. Simple Linear Regressions using temperature as the sole predictor variables were conducted to assess the predictive power of temperature as a stand-alone variable (Figure 2.4). An ANCOVA to assess the difference between the slopes of these regression lines suggested there was no difference in slope ($p = 0.455$) but that the intercept of the regression line was significant ($p = 0.028$). This difference in intercepts but not slope suggests that the FIB and markers have a similar response to changes in temperature, but the rate at which they decay is different from one another across all temperatures. The significance of temperature on decay of FIB in the environment has been reported previously (Bae and Wuertz, 2015; Noble et al., 2004). Bae and Wuertz (2015) provided a way to adjust decay rate constants estimated based on the significance of temperature, using an Arhennius correction to extrapolate decay rate constant estimates for comparison across temperatures. The findings presented here use a simple way to assess the

impact of temperature on the decay rate constant estimates, but incorporate data from multiple studies to assess the broader impact of temperature.

While temperature was found to be a good predictor across targets, the impacts of light on decay were less clear. For both culture based targets, light was included in the best fitting model as a predictor variable, while for genetic markers, it was not significant nor did it improve model fit. It is important to note that as many studies did not report the amount of solar radiation observed under light treatments, light was treated as a binary variable rather than a continuous scale, possibly introducing variation as some light treatments may have been stronger than others. Additionally, due to the lack of data available for the genetic markers, it is difficult to say whether light has no impact on decay or if there simply have not been enough data collected to detect significant effects.

Even with the inclusion of significant variables, most of the models did a poor job of explaining variation, as evidenced by the high I^2 values in the output of model models. This high unexplained variance suggests that the models are not including all, if any, of the most significant variables affecting decay in the environment. Possible additional sources of variation include strain specific differences of the FIB themselves (Maraccini et al., 2012; Noble et al., 2004). Additionally, it is possible that within even the same sources of spiking material, there could be a large amount of variation in the types and concentrations of FIB in the feces, leading to differences in decay in the environment. As the microbial community of the receiving water has been demonstrated to significantly impact decay of FIB (Korajkic et al., 2014b), it is possible that differences in microbial communities that degrade contaminants are also contributing to observed differences.

5. Conclusions

The data and models presented here synthesize data collected in previous studies to summarize what is known about decay of FIB in the environment, while highlighting the many data gaps that prevent accurate estimates for decay of FIB and the associated risks associated with fecal pathogens in the environment. Here, we focused on 5 FIB and genetic markers that have relevance for monitoring efforts and potential source allocation. However, given enough data, similar analyses for pathogens of concern could also be conducted.

Large amounts of unexplained variation found in the data suggest that in spite of numerous studies, identification of additional predictors for decay is necessary to generate more precise credible intervals for decay rate estimates. Our findings suggest that more data is needed before synthesized estimates for decay rate constants can be generated with confidence. Specifically, additional research is needed with respect to the host specific genetic markers, as even the two most heavily studied source organisms, human and ruminant, did not have sufficient data to provide useful estimates.

While more data is needed, it is crucial to make sure that future studies provide data that can be used by other researchers. For instance, providing the raw data for decay studies as supplemental information would be helpful for researchers attempting to compare their data quantitatively to past studies. It is also recommended that researchers provide all available metadata so that effects of environmental variables could be better understood. Finally, as with all scientific studies, it is crucial that researchers choose experimental designs that are sufficient to answer lingering questions regarding FIB and genetic marker persistence. Specifically, the number of replicates chosen for a given type of analysis should be investigated with power analyses when conducting a study comparing the effects of environmental variables on decay rates.

This study is not intended to end the discussion of FIB decay, but rather to summarize what is known to date and draw attention to unknowns to be addressed. The data gathered for this analysis should be used as a starting point, upon which additional data could be added and evaluated as the field continues to grow. Researchers are cautioned about using these data as a synthesized estimate, as many unknowns remain about the assumption that all decay rate constant estimates are from a single population. Although temperature and, in some cases, light were the only covariates found to be significant, other covariates that were not tested, or not identified, are likely to be present and should be included in future analyses.

Finally, nearly all of these studies have been conducted in artificial environments rather than in an open-system, suggesting possible limitations in the applicability of these estimates in the environment. Before decay rate constant estimates such as those synthesized in this study can be used with confidence in the environment, the impacts of conducting studies in a closed system should be examined and where possible, minimized using experimental designs that better simulate the environment.

The major findings of this study are as follows:

- Sufficient data are only available to synthesize first-order decay rate constants, and are not available to model individual genetic markers separately.
- Large amounts of variation exist in the data, leading to wide credible intervals for estimates of decay rate constants.
- Significant predictor variables included temperature for all targets, and light for culture based *E. coli* and *Enterococcus spp.*

- Inclusion of significant variables tested in this study failed to explain heterogeneity, and sources of large amounts of variation in the data remain unidentified.
- Future studies are necessary to not only understand decay in artificial environments, but also to assess the applicability of these rates to environments of interest.

Acknowledgements

This work was partially funded by Grant #2012-51130-20330 from USDA-NIFA. Lauren Brooks was supported by an Eckelman Foundation Fellowship.

References

Ahmed, W., Gyawali, P., Sidhu, J.P.S., and Toze, S. (2014). Relative inactivation of faecal indicator bacteria and sewage markers in freshwater and seawater microcosms. Lett. Appl. Microbiol. *59*, 348–354.

Anderson, K.L., Whitlock, J.E., and Harwood, V.J. (2005). Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Appl. Environ. Microbiol. *71*, 3041–3048.

Bae, S., and Wuertz, S. (2009). Discrimination of viable and dead fecal *Bacteroidales* bacteria by quantitative PCR with propidium monoazide. Appl. Environ. Microbiol. *75*, 2940–2944.

Bae, S., and Wuertz, S. (2011). Survival of host-associated *Bacteroidales* cells and their relationship with *Enterococcus spp.*, *Campylobacter jejuni*, *Salmonella enterica*, *Serovar Typhimurium*, and Adenovirus in freshwater microcosms as measured by propidium monoazide-Quantitative PCR. Appl. Environ. Microbiol. *78*, 922–932.

Bae, S., and Wuertz, S. (2015). Decay of host-associated *Bacteroidales* cells and DNA in continuous-flow freshwater and seawater microcosms of identical experimental design and temperature as measured by PMA-qPCR and qPCR. Water Res. *70*, 205–213.

Bell, A., Layton, A.C., McKay, L., Williams, D., Gentry, R., and Sayler, G.S. (2009). Factors influencing the persistence of fecal *Bacteroides* in stream water. J. Environ. Qual. *38*, 1224.

Bernhard, A.E., and Field, K.G. (2000). A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Appl. Environ. Microbiol. *66*, 4571–4574.

Bivand, R., Ono, H., Dunlap, R., and Stigler, M. (2015). classInt: Choose Univariate Class Intervals.

Dick, L.K., and Field, K.G. (2004). Rapid estimation of numbers of fecal *Bacteroidetes* by use of a quantitative PCR assay for 16S rRNA genes. Appl. Environ. Microbiol. *70*, 5695–5697.

Dick, L.K., Bernhard, A.E., Brodeur, T.J., Santo Domingo, J.W., Simpson, J.M., Walters, S.P., and Field, K.G. (2005). Host distributions of uncultivated fecal *Bacteroidales* bacteria reveal genetic markers for fecal source identification. Appl. Environ. Microbiol. *71*, 3184–3191.

Dick, L.K., Stelzer, E.A., Bertke, E.E., Fong, D.L., and Stoeckel, D.M. (2010). Relative decay of *Bacteroidales* microbial source tracking markers and cultivated *Escherichia coli* in freshwater microcosms. Appl. Environ. Microbiol. *76*, 3255–3262.

Gelman, A., and Rubin, D.B. (1992). Inference from Iterative Simulation Using Multiple Sequences. Stat. Sci. *7*, 457–472.

Goldin, S.M.C. with contributions from I. (2012). mcmcplots: Create Plots from MCMC Output.

Green, H.C., Shanks, O.C., Sivaganesan, M., Haugland, R.A., and Field, K.G. (2011). Differential decay of human faecal *Bacteroides* in marine and freshwater. Environ. Microbiol. *13*, 3235–3249.

Haugland, R.A., Siefring, S.C., Wymer, L.J., Brenner, K.P., and Dufour, A.P. (2005). Comparison of *Enterococcus* measurements in freshwater at two recreational beaches by quantitative polymerase chain reaction and membrane filter culture analysis. Water Res. *39*, 559–568.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., and Altman, D.G. (2003). Measuring inconsistency in meta-analyses. BMJ *327*, 557–560.

Jeanneau, L., Solecki, O., Wéry, N., Jardé, E., Gourmelon, M., Communal, P.-Y., Jadas-Hécart, A., Caprais, M.-P., Gruau, G., and Pourcher, A.-M. (2012). Relative decay of fecal indicator bacteria and human-associated markers: A microcosm study simulating wastewater input into seawater and freshwater. Environ. Sci. Technol. *46*, 2375–2382.

Kay, D., Stapleton, C.M., Wyer, M.D., McDonald, A.T., Crowther, J., Paul, N., Jones, K., Francis, C., Watkins, J., Wilkinson, J., et al. (2005). Decay of intestinal enterococci concentrations in high-energy estuarine and coastal waters: towards real-time T90 values for modelling faecal indicators in recreational waters. Water Res. *39*, 655–667.

Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B., and Wuertz, S. (2007). 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal Bacteroidales: A Bayesian approach. Water Res. *41*, 3701–3715.

Korajkic, A., McMinn, B.R., Shanks, O.C., Sivaganesan, M., Fout, G.S., and Ashbolt, N.J. (2014a). Biotic Interactions and Sunlight Affect Persistence of Fecal Indicator Bacteria and Microbial Source Tracking Genetic Markers in the Upper Mississippi River. Appl. Environ. Microbiol. AEM.00388-14.

Korajkic, A., McMinn, B.R., Shanks, O.C., Sivaganesan, M., Fout, G.S., and Ashbolt, N.J. (2014b). Biotic interactions and sunlight affect persistence of fecal indicator bacteria and microbial source tracking genetic markers in the upper Mississippi River. Appl. Environ. Microbiol. *80*, 3952–3961.

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R., and Sayler, G. (2006). Development of *Bacteroides* 16S rRNA gene TaqMan-based Real-Time PCR assays for estimation of total, human, and bovine fecal pollution in water. Appl. Environ. Microbiol. *72*, 4214–4224.

Maraccini, P.A., Ferguson, D.M., and Boehm, A.B. (2012). Diurnal variation in *Enterococcus* species composition in polluted ocean water and a potential role for the enterococcal carotenoid in protection against photoinactivation. Appl. Environ. Microbiol. *78*, 305–310.

Meng, X.-L. (1994). Posterior Predictive p-Values. Ann. Stat. *22*, 1142–1160.

Noble, R.T., Lee, I.M., and Schiff, K. c. (2004). Inactivation of indicator micro-organisms from various sources of faecal contamination in seawater and freshwater. J. Appl. Microbiol. *96*, 464–472.

Okabe, S., and Shimazu, Y. (2007). Persistence of host-specific *Bacteroides–Prevotella* 16S rRNA genetic markers in environmental waters: effects of temperature and salinity. Appl. Microbiol. Biotechnol. *76*, 935–944.

Schulz, C.J., and Childers, G.W. (2011). Fecal *Bacteroidales* diversity and decay in response to variations in temperature and salinity. Appl. Environ. Microbiol. *77*, 2563–2572.

Seurinck, S., Defoirdt, T., Verstraete, W., and Siciliano, S.D. (2005). Detection and quantification of the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. Environ. Microbiol. *7*, 249–259.

Sinton, L., Hall, C., and Braithwaite, R. (2007). Sunlight inactivation of *Campylobacter jejuni* and *Salmonella enterica*, compared with *Escherichia coli*, in seawater and river water. J. Water Health *5*, 357–365.

Sinton, L.W., Hall, C.H., Lynch, P.A., and Davies-Colley, R.J. (2002). Sunlight inactivation of fecal indicator bacteria and bacteriophages from waste stabilization pond effluent in fresh and saline waters. Appl. Environ. Microbiol. *68*, 1122–1131.

Su, Y.-S., and Yajima, M. (2014). R2jags: A Package for Running jags from R.

Sutton, A.J., and Abrams, K.R. (2001). Bayesian methods in meta-analysis and evidence synthesis. Stat. Methods Med. Res. *10*, 277–303.

Tambalo, D.D., Fremaux, B., Boa, T., and Yost, C.K. (2012). Persistence of host-associated *Bacteroidales* gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. Water Res. *46*, 2891–2904.

Thompson, S.G., and Higgins, J.P.T. (2002). How should meta-regression analyses be undertaken and interpreted? Stat. Med. *21*, 1559–1573.

Walters, S.P., and Field, K.G. (2009). Survival and persistence of human and ruminant-specific faecal *Bacteroidales* in freshwater microcosms. Environ. Microbiol. *11*, 1410–1421.

Wang, D., Farnleitner, A.H., Field, K.G., Green, H.C., Shanks, O.C., and Boehm, A.B. (2013). *Enterococcus* and *Escherichia coli* fecal source apportionment with microbial source tracking genetic markers – Is it feasible? Water Res. *47*, 6849–6861.

Figures and Tables:

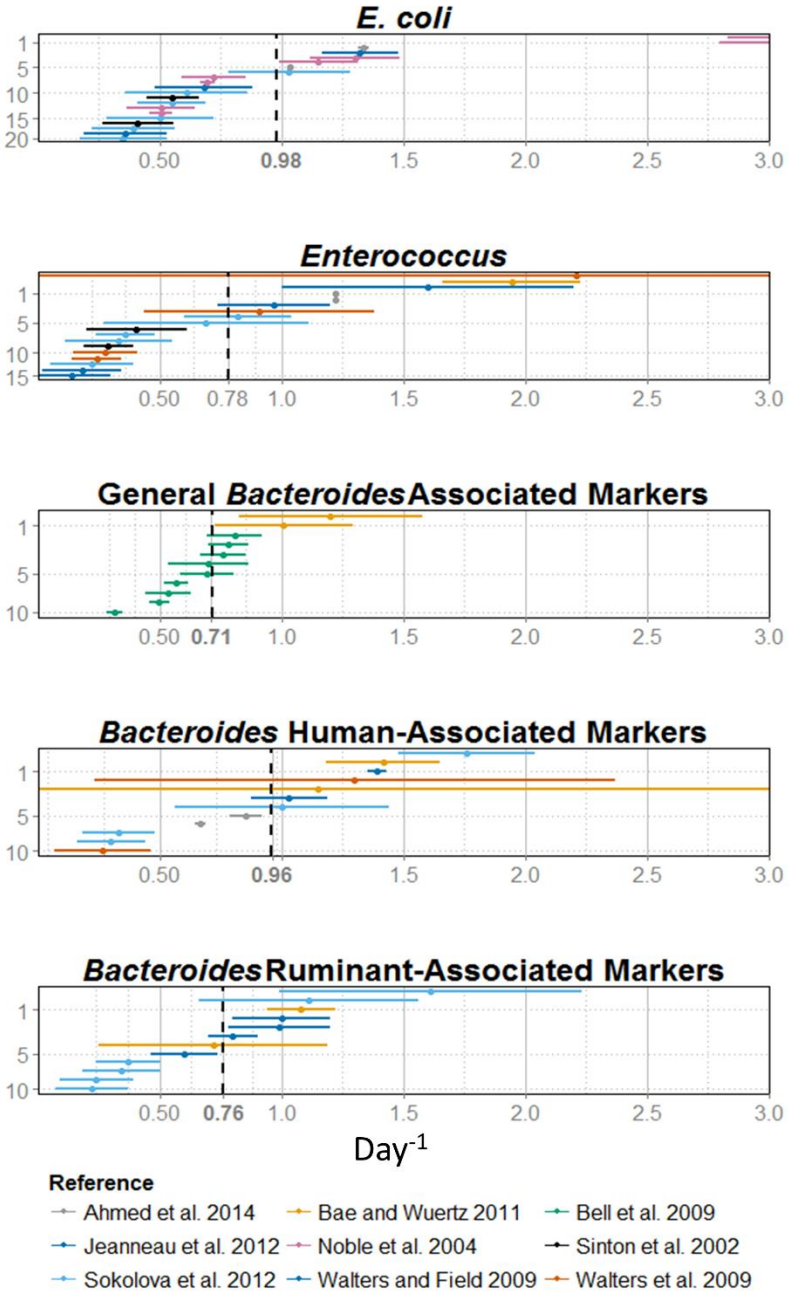

Figure 2.1 First order decay rate estimates and standard errors from published papers for selected FIB reveal inconsistency in the literature among decay rate estimates.

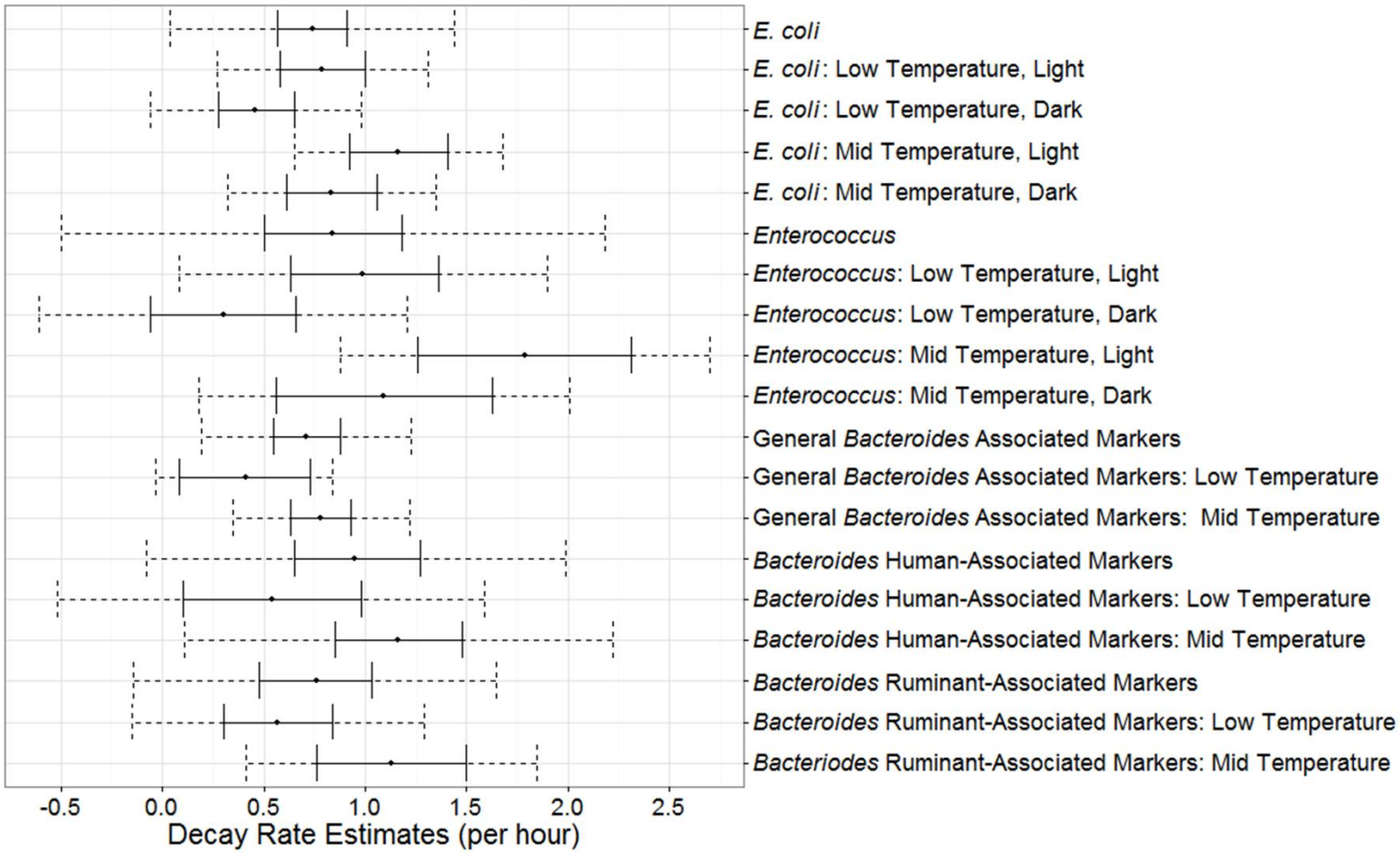

Figure 2.2 95% Credible Intervals for the mean (solid line) and distribution (dashed line) generated from each model output.

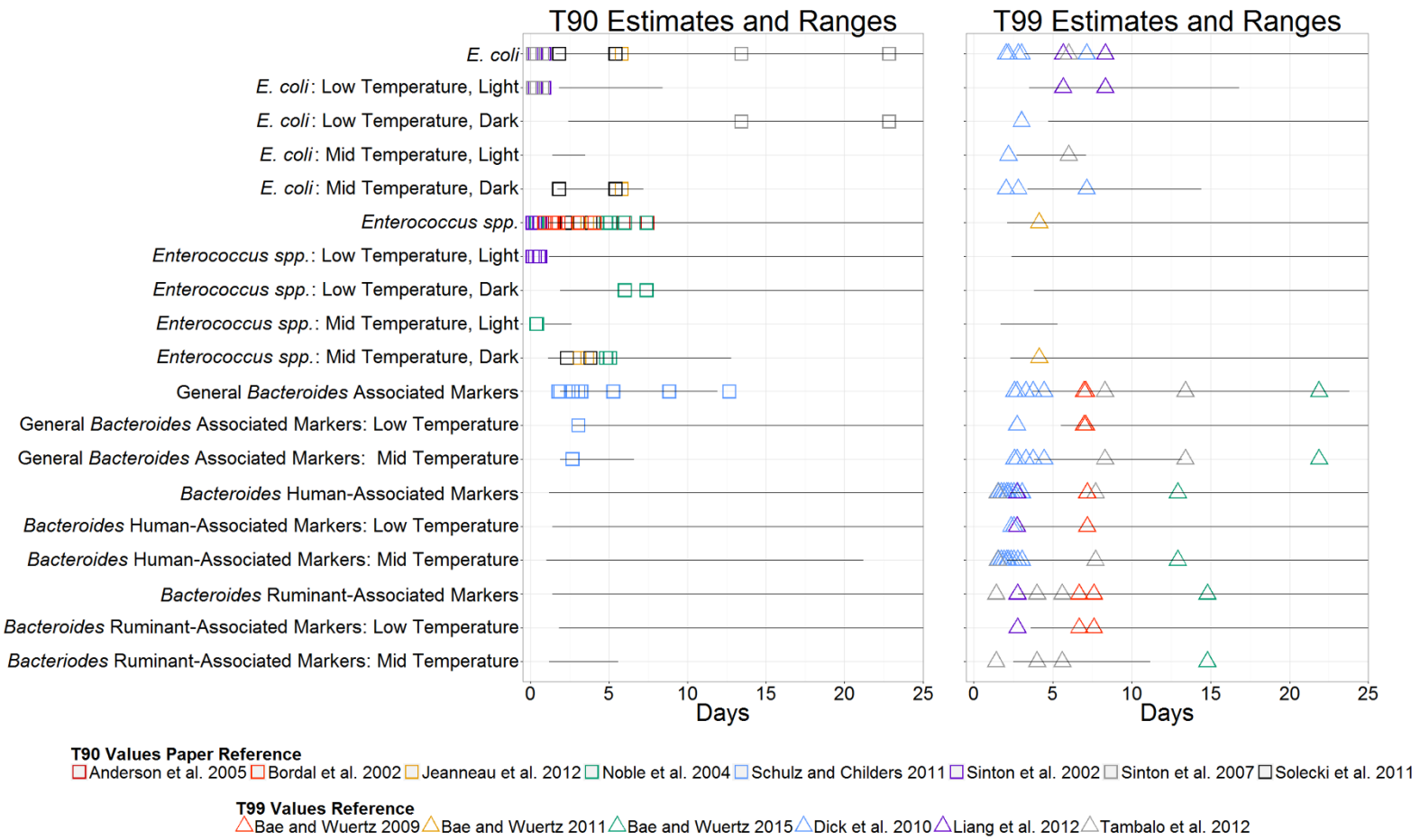

Figure 2.3 Data points used for posterior predictive checking and credible T90/T99 ranges generated by each model. Lines represent the ranges predicted from the output of the models.

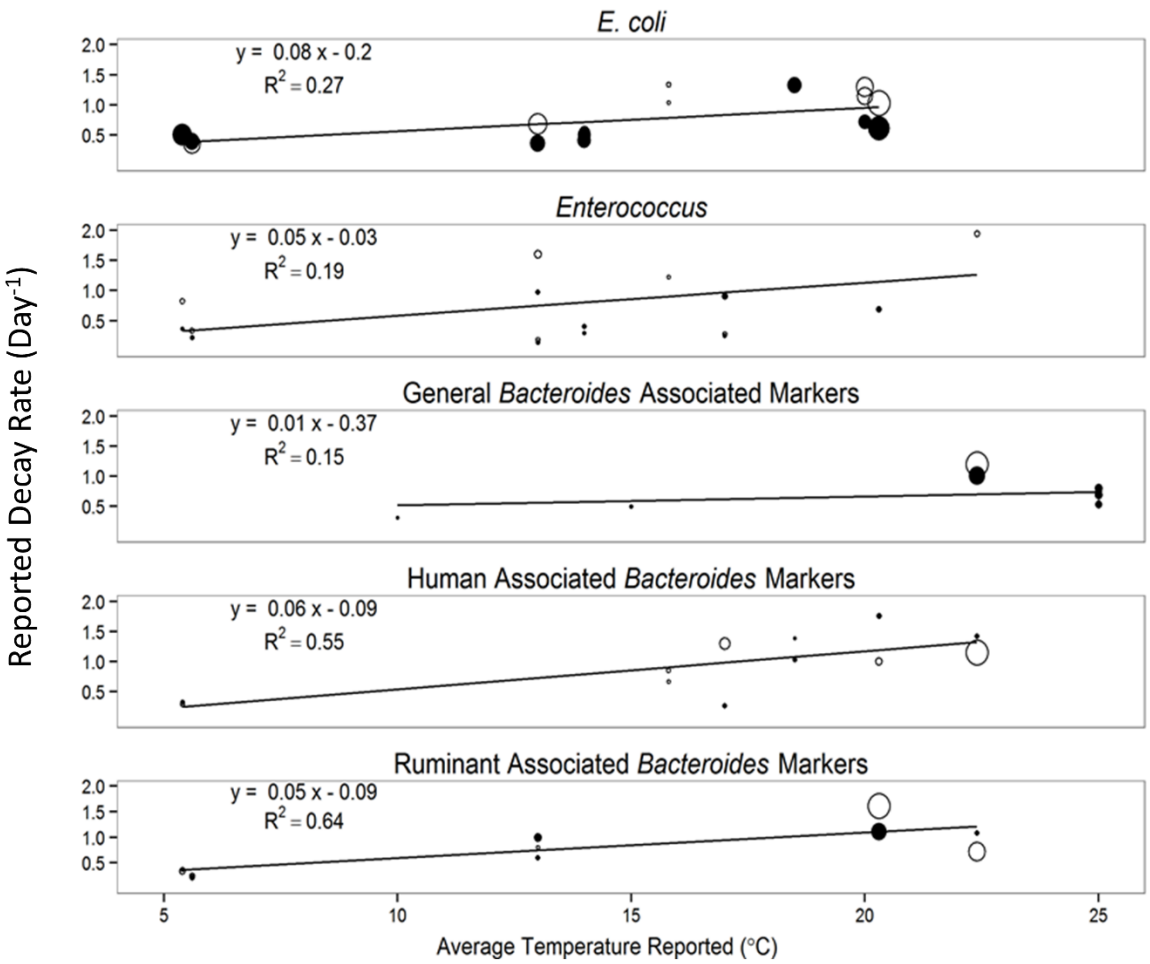

Figure 2.4 Simple linear regression showing effect of temperature on decay rate estimates with a distinction for data points generated under light treatments (○) and dark treatments (●). Circle size corresponds to standard error estimates.

Table 2.1 References for data used to construct and test models. Variables were defined as either binary ([b]), or categorical ([c])

Target	Model Data Paper References	Variables Tested	Checking Data Paper References
Culturable *E. coli*	Ahmed et al. 2014 Jeanneau et al. 2012 Noble et al. 2004 Sinton et al. 2002 Sokolova et al. 2012 Walters and Field 2009	Light[b] Water Type[c] Temperature[c] Fecal Spike[c] Sediment Inclusion[b] Culture Media[c] Experiment Size[c] Experiment Length[c]	Dick et al. 2010 Jeanneau et al. 2012 Liang et al. 2012 Sinton et al. 2002 Sinton et al. 2007 Solecki et al. 2011 Tambalo et al. 2012
Culturable *Enterococcus*	Ahmed et al. 2014 Bae and Wurtz 2011 Sinton et al. 2002 Sokolova et al. 2012 Walters and Field 2009 Walters et al. 2009	Light[b] Water Type[c] Temperature[c] Fecal Spike[c] Culture Media[c] Experiment Size[c] Experiment Length[c]	Anderson et al. 2005 Bae and Wurtz 2011 Bordalo et al. 2002 Jeanneau et al. 2012 Noble et al. 2004 Sinton et al. 2002 Solecki et al. 2011
General *Bacteroides* Associated Markers	Bae and Wurtz 2011 Bell et al. 2009	Light[b] Temperature[c] Experiment Length[c] Primer set[c]	Bae and Wuertz, 2009 Bae and Wurtz 2015 Dick et al. 2010 Schulz and Childers 2011 Tambalo et al. 2012
Bacteroides Human-Associated Markers	Ahmed et al. 2014 Bae and Wurtz 2011 Jeanneau et al. 2012 Sokolova et al. 2012 Walters et al. 2009	Light[b] Water Type[c] Temperature [c] (Low and Mid only) Fecal Spike[c] Experiment Size[c] Experiment Length[c]	Bae and Wuertz, 2009 Bae and Wurtz 2015 Dick et al. 2010 Liang et al. 2012 Sokolova et al. 2012 Tambalo et al. 2012
Bacteroides Ruminant-Associated Markers	Bae and Wurtz 2011 Sokolova et al. 2012 Walters and Field 2009	Light[b] Temperature[c] Experiment Length[c]	Bae and Wuertz, 2009 Bae and Wurtz 2015 Liang et al. 2012 Tambalo et al. 2012

Table 2.2 Output from meta-analyses

Target	N	Mean (95% Credible Interval)	Variance (95% Credible Interval)	DIC	I^2	Distribution 95% Credible Interval	% of Data in Credible Interval	T90 95% Credible Interval (T99)	% of T90/99 Data in Credible Interval (proportion)
E. coli	19	0.74 (0.57, 0.91)	0.13 (0.07, 0.25)	16.87	100	0.04, 1.44	100	1.6, 55.7 (3.2, 129.2)	43 (9/21)
Mid Temperature, Dark	4	0.83 (0.61, 1.06)	0.07 (0.03, 0.15)	7.58	98	0.32, 1.35	100	1.7, 7.2 (3.4, 14.4)	57 (4/7)
Low Temperature, Dark	7	0.46 (0.28, 0.65)	0.07 (0.03, 0.15)	7.58	70	-.06, 0.98	100	2.4, NA (4.7, NA)	67 (2/3)
Low Temperature, Light	5	0.79 (0.58, 1)	0.07 (0.03, 0.15)	7.58	99	0.27, 1.31	80	1.8, 8.4 (3.5, 16.8)	20 (2/10)
Mid Temperature, Light	3	1.16 (0.92, 1.41)	0.07 (0.03, 0.15)	7.58	98	0.65, 1.68	100	1.4, 3.5 (2.7, 7.1)	50 (1/2)
Enterococcus	16	0.84 (0.5, 1.18)	0.48 (0.23, 0.99)	34.77	100	-0.5, 2.18	94	1.1, NA (2.1, NA)	54 (15/28)
Low Temperature, Dark	6	0.3 (-0.06, 0.66)	0.23 (0.1, 0.5)	25.6	93	-0.61, 1.21	100	1.9, NA (3.8, NA)	100 (2/2)
Mid Temperature, Dark	2	1.09 (0.56, 1.63)	0.23 (0.1, 0.5)	25.6	96	0.18, 2.01	100	1.1, 12.8 (2.3, 25.6)	100 (7/7)
Low Temperature, Light	6	0.99 (0.63, 1.36)	0.23 (0.1, 0.5)	25.6	98	0.08, 1.9	100	1.2, 28.3 (2.4, 56.6)	0 (0/4)
Mid Temperature, Light	2	1.79 (1.26, 2.31)	0.23 (0.1, 0.5)	25.6	99	0.88, 2.7	100	0.9, 2.6 (1.7, 5.3)	0 (0/4)
General *Bacteroides* Associated Markers	11	0.71 (0.55, 0.88)	0.08 (0.03, 0.19)	3.76	99	0.19, 1.24	100	1.9, 12 (3.7, 24)	74 (14/19)
Mid Temperature	8	0.78 (0.63, 0.94)	0.05 (0.02, 0.14)	1.33	99	0.35, 1.22	100	1.9, 6.6 (3.8, 13.1)	33 (3/9)
Low Temperature	2	0.41 (0.08, 0.73)	0.05 (0.02, 0.14)	1.33	65	-0.03, 0.84	100	2.8, NA (5.5, NA)	100 (3/3)
***Bacteroides* Human-Associated Markers**	12	0.95 (0.65, 1.27)	0.29 (0.12, 0.7)	20.33	99	-0.08, 1.99	100	1.2, NA (2.3, NA)	56 (9/16)

Low Temperature	4	0.54 (0.1, 0.98)	1.16 (0.85, 1.48)	16.88	95	-0.52, 1.59	100	14, NA (2.9, NA)	0 (0/3)
Mid Temperature	8	1.16 (0.85, 1.48)			98	0.11, 2.22	100	1, 21.2 (2.1, 42.3)	62 (8/13)
***Bacteroides* Ruminant-Associated Markers**	12	0.76 (0.48, 1.03)	0.22 (0.09, 0.51)	16.73	95	-0.14, 1.65	100	1.4, NA (2.8, NA)	74 (5/7)
Low Temperature	8	0.57 (0.3, 0.84)	0.14 (0.06, 0.36)	13.26	89	-0.15, 1.29	100	1.8, NA (3.6, NA)	66 (2/3)
Mid Temperature	4	1.13 (0.76, 1.5)			96	0.41, 1.85	100	1.2, 5.6 (2.5, 11.2)	50 (2/4)

CHAPTER 3

GLOBAL MODEL FITTING TO COMPARE SURVIVAL CURVES FOR FECAL INDICATOR BACTERIA AND RUMINANT-ASSOCIATED GENETIC MARKERS

Lauren E. Brooks and Katharine G. Field

In preparation for the Journal of Applied Microbiology

Abstract

The environmental decay of molecular markers for fecal contamination impacts their utility for microbial source tracking, and must be understood before recommendations for genetic markers can be made. We conducted a decay study to describe the decay profiles of cultivable *E. coli*, a general *Bacteroidales* marker (GenBac3), ruminant associated markers (CF128, Rum2Bac), and cattle associated markers (CowM2, CowM3). We selected best fitting models for the decay profile of each marker and tested for differences between decay profiles of different indicators using Global Model Fitting techniques to directly test for differences between curves. Initial concentrations varied across targets, leading to variability in the time to non-detection, despite similar survival profiles for some markers. After normalizing for initial concentration, we found statistically significant differences between the decay curves of *E. coli* and all genetic markers except for CowM3. The decay curves for CF128 compared to GenBac3 and Rum2Bac also differed, but no differences were observed between Rum2Bac and GenBac3 decay profiles. The differences across decay profiles and time to non-detection suggest caution is necessary when interpreting microbial source tracking results using these markers, as differential decay may result in different findings depending on the marker selected. However, as the Rum2Bac and GenBac3 marker had similar decay profiles, these two markers could potentially be used reliably over time for source allocation using the ratio method.

Introduction

Identification of fecal contamination from ruminant sources can affect the interpretation of water quality monitoring results for both public risk assessment and source mitigation. As ruminant sources of fecal contamination have been shown to contribute to elevated incidents of gastrointestinal illnesses (Soller et al., 2010), elevated fecal indicator counts from these sources could be of greater risk to the public. Additionally, as agricultural contributions of nonpoint sources of contaminants are the leading cause of stream impairments (USEPA, 2010), it is important to have methods that are capable of detecting fecal contamination from agricultural ruminants, such as cattle and sheep.

Host specific genetic markers have been identified that can distinguish sources of fecal contamination. However, the characteristics of selected markers, such as decay in the environment, impact interpretation of results. As host specific markers are typically used to identify sources that contribute to impairment in a given water body, variations in the ability to detect marker presence in the period of time following loading to the water body can lead to different results, and thus different actions taken depending on the marker selected for use. Efforts are underway to standardize the use of genetic markers, including making recommendations for markers to be used for microbial source tracking (MST) (Ahmed et al., 2016; Boehm et al., 2013; Schriewer et al., 2013; Shanks et al., 2010; Staley et al., 2012; Ahmed et al., 2015).

Based on their assay sensitivity and specificity, primers targeting ruminant associated *Bacteroidales* (Rum2Bac) or targeting bacterial genes involved in host interactions (CowM2 and CowM3) have been recommended for MST studies involving ruminant fecal contamination (Raith et al., 2013; Shanks et al., 2010). However, information on decay is lacking for these three

recommended markers, while numerous studies have addressed decay for other fecal indicators (Bae and Wuertz, 2011; Green et al., 2011; Jeanneau et al., 2012; Sinton et al., 2007, 2002; Sokolova et al., 2012; Tambalo et al., 2012; Walters and Field, 2009), including a different marker targeting ruminant sources of contamination, CF128 (Bernhard and Field, 2000; Walters and Field, 2009). While CF128, like Rum2Bac, targets microbes from the order *Bacteroidales*, differences between organisms detected by the two primer sets are not well defined.

This paper reports findings of a decay study to assess persistence of ruminant markers in freshwater mesocosms spiked with a slurry of cattle feces. Additionally, we measured decay of GenBac3, a general marker for members of the group *Bacteroidales* (Dick and Field, 2004; Siefring et al., 2008), and cultivable *E. coli*, a long-used general indicator of microbial contamination (Geldreich, 1970). Decay profiles for the markers were described with mathematical models individually as well as in comparison to one another and cultivable *E. coli*. We used global model fitting (GMF) to compare decay profiles as whole curves, rather than assessing individual time points or parameters. Additionally, the concentration ratios of different host specific markers compared to the GenBac3 assay over time were compared to assess the feasibility of the ratio method for source allocation (Wang et al., 2013) with these selected markers.

Methods

Mesocosm construction

Decay was studied in a simulated freshwater system at the John D. Fryer Aquatic Animal Health Laboratory (AAHL, Corvallis OR), using three replicate 100 L mesocosms. Indoor tanks exposed to ambient sunlight were filled with unfiltered river water pumped from the Willamette

River. Water quality parameters were recorded using a YSI ProPlus Multiparameter meter (Yellow Springs International, Yellow Springs, OH) throughout the experiment (mean pH = 7.74, mean DO = 9.88 mg/l, mean temperature = 11°C). Constant mixing was achieved using a Koralia Circulation and Wave Pump (Hydror, Sacramento, CA) and an air-stone was added to the tanks to ensure aerobic conditions.

To simulate contamination typical of runoff from a dairy farm, 200 ml feces were collected from each of 10 separate patties at the OSU Dairy Cattle Farm. The fecal samples were transported on ice to the laboratory and immediately mixed together with 3.8 L river water to create a fecal slurry (5% V:V dilution). The slurry was homogenized and 500 ml were dispensed into individual tanks and allowed to mix prior to sampling

Sample Collection, Enumeration of FIB and Genetic Markers

Water samples from each tank were collected in sterile 500 ml polypropylene bottles and brought to the laboratory on ice for immediate processing. Samples were collected daily from days 0-7 for *E. coli* analysis, and filters for DNA extraction were collected on days 0, 1, 3, 5, and 7. To enumerate *E. coli*, 100 ml of sample was combined with a Colilert packet in a disposable 120 ml tear-off vessel without sodium thiosulfate and mixed until the reagents were dissolved. Mixed samples were then poured into a Quanti-Tray/2000 (Idexx Technologies, Westbrook, ME) and heat sealed prior to incubating for 24 hours at 37°C. All observations were below the upper range of quantitation (2419.2 CFU/100 ml) for this methodology, and samples were collected until all tanks fell below the lower range (1 CFU/100 ml).

To collect bacteria for the enumeration of genetic markers, 50 ml of collected samples were filtered onto 0.4 µm polycarbonate filters (Pall Corporation, Fort Washington, NY) and frozen at -80°C for less than one year until DNA extractions could be performed. DNA

extractions were performed following the manufacturer's protocols using the Gene-Rite DNA extraction kit with modifications and processing controls implemented as recommended for analysis of human specific markers (Shanks et al., 2016), including the addition of a Sample Processing Control (Haugland et al., 2005). Extracted DNA was stored at -20°C until used for qPCR amplification

qPCR assays

To establish a standard curve for absolute quantitation, synthetic plasmids (Integrated DNA Technologies, Coralville, IA) containing sequences corresponding to the targeted regions for each primer set were diluted from a range of 10^8 to 10^1. Calibration curves were constructed for single runs using serially diluted synthesized standards containing the target. Two separate standards were generated, one of which contained matching regions for GenBac3 (Dick and Field, 2004; Siefring et al., 2008), CF128 (Bernhard and Field, 2000; Seurinck et al., 2005), and Rum2Bac (Mieszkin et al., 2010), and the other contained CowM2 and CowM3 (Shanks et al., 2007). All runs achieved acceptable linearity as noted by the R^2 value being >0.9, and all runs maintained an efficiency between 90 and 100%, with the exception of CF128, which had an efficiency of 82% (Table S3.1).

To enumerate marker concentration, 2 µl of extracted DNA samples were amplified using qPCR. Each 25 µl reaction consisted of 10 µl Quanta PerfeCTa Toughmix, 500 nmol l^{-1} each primer, and 250 nmol l^{-1} of probe (Table 3.1). For CF128, 0.1× SYBR green I dye (Life Technologies, Gaithersburg, MD) was used as a general probe for double stranded DNA, while for other assays TaqMan probes were designed based on published sequences (Table 3.1). All reactions were performed in triplicate in MicroAmp optical 96-well plates with optical adhesive film (Thermo Scientific, Wilmington, DE), using the ABI 7500 Fast Real-Time PCR system with

cycling parameters consisting of a 2 minute start at 94°C followed by 40 cycles of 15s at 94°C and 32s at 60°C. Cycle threshold was determined by the software provided with the ABI 7500 Fast Real-Time PCR system and used to calculate the concentration using equations fit from the standard curves (Table S3.1). Results were exported to Microsoft Excel for analysis.

Quality Controls

To ensure quality of data for analysis, several quality control steps were implemented at various stages of sample processing. To measure potential contamination during filtration as well as to create extraction blanks for comparisons using a sample processing control, 3 filters were used to collect 100 ml deionized water instead of a sample. To detect contamination and provide inhibition free samples to establish a metric for amplification inhibition, 9 no template controls (NTCs) were added to each plate. Extraction blanks and NTCs were found to be free of contamination for all samples.

To detect variation in extraction efficiency, a sample processing control spiked into the extraction buffer for all samples was measured and compared to extraction blanks as described previously (Haugland et al., 2005). The cycle thresholds (C_t) of the extraction blanks were compared to the C_t for each samples for the Sketa22 assay to measure for inhibition from the sample matrix. The extraction blanks had a mean C_t of 32.5 (SD = 1.13), creating an acceptable range of 29.1 to 35.9 cycles for samples. All samples fell well within the bounds generated by the blanks.

In addition to the sample processing controls, an internal amplification control (Haugland et al., 2010) was tested by conducting a multiplex assay for the GenBac3 assay, measuring both the GenBac3 marker and the IAC marker (UCP1). The mean from the no template controls (mean = 27.4, SD = 1.2) was used to construct acceptable bounds for the samples, with a range

from 23.7 to 31.1. All samples fell within this bound as well, suggesting there were no problems with inhibition in the samples.

Decay Profiles and Statistical Calculations

To compare markers, decay profiles were constructed for all markers and indicators. All statistical analyses and curve fitting were conducted using R (R Core Team, 2015) through the Rstudio interface (2015). Raw concentration data were normalized by dividing by the initial concentration for each replicate tank to generate the Survival Ratio ($S_t = C_t/C_0$) for each replicate at each time point.

Visual inspection of the data indicated the presence of a shoulder period, so in addition to the commonly used Chick Watson model (Eq. 1), data were fit using two models with shoulders, the Delayed Chick Watson (DCW, Eq. 2) model, as used previously to model indicator decay (Green et al., 2011), and the Geeraerd Shoulder (GS, Eq.3) model (Geeraerd et al., 2000). Concentrations from each replicate tank were treated individually, resulting in three data points for each time point.

Eq. 1: $S_t = e^{-kt}$, where k = decay rate (hour^{-1}), t = time (hours)

Eq. 2: If t <= lag: $S_t = 1$

If t > lag: $S_t = e^{-k(t-lag)}$, where lag = duration of no decay period (hours)

Eq. 3: $S_t = e^{-kt}\left(\frac{e^{ks}}{1+(e^{ks}-1)e^{-kt}}\right)$, where s = shoulder period (hours)

Models were coded and fit in R using the r2ADMB package (Bolker et al., 2015). This package allows interface through R with Automatic-Differentiation Model Builder (Fournier et al., 2012). Best fitting model selection was made based on Akaike Information Criterion (AIC) values (Table 3.2), which account for not only the sum of squared errors, but also penalize for the incorporation of additional model parameters (Akaike, 1992). AIC values for different model

fits were directly compared, and the probability of choosing the correct model was calculated for each dataset:

Eq. 4 $Probability\ Model\ 1\ is\ correct = \frac{e^{-0.5(AIC_2 - AIC_1)}}{1+e^{-0.5(AIC_{2-}AIC_1)}}$

Once the best fitting model for individual markers or FIB was determined, statistical comparisons to detect changes in survival profiles were made using GMF. Prior to conducting this experiment, we conducted a power analysis (data not shown) to determine sensitivity of this statistical method versus other methods of fitting non-log linear data (e.g. repeated measures ANOVA). In our application of GMF, separate datasets from different markers are combined and fit using models (Table S3.2) as described in detail by Motulsky and Christopoulos (Motulsky and Christopoulos, 2004).

Results

Significant differences ($p < 0.05$) were detected among all initial concentrations (concentrations of markers in seeded tanks at time zero) except for GenBac3 compared with Rum2Bac ($p = 0.054$); differences between these two markers were nearly significant. The GenBac3 marker had the highest initial concentration (mean $\log_{10} C_0$/100 ml= 6.96), followed by the two markers that target ruminant-specific *Bacteroidales* ribosomal genes, Rum2Bac (mean $\log_{10} C_0$/100 ml = 6.61) and CF128 (mean $\log_{10} C_0$/100 ml = 6.04). Less abundant were the markers targeted by the CowM2 (mean $\log_{10} C_0$/100 ml = 4.14) and CowM3 (mean $\log_{10} C_0$/100 ml = 3.79) assays. Cultured *E. coli* had the lowest initial concentration. Although for cultured *E. coli* concentration is measured as CFU, as opposed to marker copy number for the markers, and is thus not strictly comparable, its concentration is useful because of the information it provides about relative ability to detect the indicators (Table 3.3).

After normalizing for the initial concentration using the survival ratio for each tank, decay profiles revealed a period of little or no decay for all datasets (Figure 3.1). However, the models selected to best represent this feature varied for the indicators. Decay profiles for the CowM2 assay and *E. coli* were best described by the inclusion of a lag period with no decay, as defined in the DCW model, while the decay profiles for all *Bacteroidales* associated markers were fit best by the GS decay model. High residual errors were detected on day 1 for CF128, Rum2Bac, and GenBac3, suggesting the best fitting model underestimated concentrations on day 1 (Figure 3.2). Models fit to the *E. coli* dataset did the poorest job of fitting the data, despite having a higher number of sample points.

Interpretation of the lag period from the DCW model is relatively straightforward, as it assumes no decay takes place before that point. For CowM3, a lag period was visible in the decay profile, but was not depicted by the best fitting model for the CowM3 dataset (Table 3.3). Probabilities of selecting the correct model for the dataset revealed the models chosen for each dataset significantly improved fit over the alternative model that incorporates a lag or shoulder period (Table 3.2). Different lag periods observed for CowM2 and cultivable *E. coli* supported the detection of significant differences in the overall curve shape by GMF (Table 3.4).

The shoulder term in the GS model is less intuitive, as it does not correspond precisely to a notable shift in the data. Shoulder estimates for the three *Bacteroidales* associated markers were similar, while the decay rate estimates varied, with Rum2Bac and GenBac3 having higher estimated decay. These differences were supported by GMF, which found moderately significant ($p<0.1$) differences between decay profiles for CF128 compared with Rum2Bac or GenBac3, while no differences ($p>0.1$) were found for Rum2Bac and GenBac3.

Making comparisons across parameter estimates from different decay models is challenging, but the overall curves shape can be compared using GMF. Although the best fitting model differed for CowM3, no significant differences were detected for CowM3 compared to any of the decay profiles for the other indicators Additionally, differences between CowM2, GenBac3, and Rum2Bac were not significant ($p>0.1$), while differences were detected between decay profiles for CowM2 and CF128.

Discussion

The differences observed in the initial concentrations of the indicators highlight the fact that even though these markers may all be used to indicate the presence of fecal contamination, they target different organisms, and will not necessarily have the same properties. The higher concentrations of the markers targeting organisms from the order *Bacteroidales* is supported by other reports finding high concentrations of *Bacteroidales* in the guts of animals (Fiksdal et al., 1985; Kreader, 1998), including ruminants (Shanks et al., 2011), while the markers targeting functional genes typically have lower concentrations in the gut of animals (Raith et al., 2013). This difference in abundance affects the sensitivity of these markers for use in source tracking (Raith et al., 2013; Shanks et al., 2010), as the ability to detect markers over time depends on both the initial concentration and the decay in the environment.

While decay profiles were standardized by initial fecal slurry seeding concentration prior to fitting the data to the curves, the differences among initial concentrations for each indicator still had a pronounced effect on the shape of the curves by limiting the number of data points collected for lower abundance markers. For the CowM2 and CowM3 markers, only two data points past the initial measurement were reported, as concentrations from subsequent days were

below the range of quantitation (Table S3.1). A past study measuring CowM2 concentrations over time was limited by the same issue, with data only collected through 24 hours before it was below the limit of detection (Tambalo et al., 2012). While CowM2 and CowM3 markers are highly specific for cattle (Raith et al., 2013), they may only be suitable for microbial source tracking studies of recent contamination events

We made statistical comparisons using GMF, a sensitive statistical technique that compares the fit of models to the data when the data from two separate datasets are grouped as a shared dataset to those fits generated by treating the data separately. Other techniques that have been employed for comparison of decay profiles exhibiting non-linear patterns rely on comparing estimates for individual model parameters or time points (Ahmed et al., 2014; Anderson et al., 2005; Bell et al., 2009; Jeanneau et al., 2012; Sokolova et al., 2012; Solecki et al., 2011) or comparing individual concentration estimates at sampling points using repeated measures ANOVA (Bae and Wuertz, 2011; Liang et al., 2012; Solecki et al., 2011). While those techniques provide valuable data, they fail to test whether or not there is a difference between the overall shapes of the curves.

While the GMF approach is a more sensitive method, it requires fitting the data to mathematical models for decay, and any lack of fit for the models chosen can decrease the power of the GMF approach. Thus, it is essential to select well-fitting mathematical models for any dataset being compared with GMF. The different model shapes that were selected to represent our data may be a reflection of truly different biological phenomena (i.e. the different targets are truly experiencing a difference in persistence), or they may be artifacts of the sampling scheme and the particular time-points selected.

Both models chosen in this study estimate two terms, the log linear decay rate (k) and the lag or shoulder period (lag or s), although the interpretation of these parameters differs between the two models. The DCW model has been used previously to describe data in fecal indicator decay studies (Bae and Wuertz, 2011; Green et al., 2011), and has been proposed to result from a buildup of predators necessary for decay to take place (Green et al., 2011). The GS model has predominantly been used in food science (Swinnen et al., 2004), and was designed to reflect a decline in a "critical component" necessary for the survival of the indicator (Geeraerd et al., 2000). Our selection of best fitting model was strongly supported by the high probabilities calculated for selecting the correct model for a given dataset (Table 3.2). However, this was based only on the mathematical description of these data, and cannot be used to support claims regarding the mechanisms or interpretation of these models.

It is likely that the lack of a shoulder or lag period detected for the CowM3 dataset resulted from the sampling scheme rather than truly indicating that the decay rate was uniform at all times, as implied by the CW model. The DCW model described the data well, but the standard error estimates were wide ranging, resulting in the log linear model being selected, although there is a clear discrepancy in the predicted and observed values (Figure 3.2).

Differences observed between *E. coli* and all other markers highlight the need to use caution when interpreting microbial source tracking results, even when doing a presence/absence screening for host specific markers. Results from our comparisons using GMF revealed statistical differences in the data when combining *E. coli* and all markers with the exception of CowM3. This may be a result of the poor fit of the CowM3 model to the data, rather than a suggestion that *E. coli* and CowM3 truly have the same shape. Past studies comparing the decay of *E. coli* and general *Bacteroidales* markers have found no differences (Dick et al., 2010;

Tambalo et al., 2012), in contrast to our findings here. However, some studies have reported differences between *E. coli* and ruminant associated markers (Tambalo et al., 2012), while others reported no difference (Sokolova et al., 2012). While no data were available comparing GenBac3 to ruminant markers, past studies reported that different ruminant markers decay at different rates (Walters and Field, 2009).

Previous work has assessed feasibility of source allocation using ratios between host specific and more general indicators, such as *E. coli* (Wang et al., 2013) or, potentially GenBac3. Under this method, the concentration of a host specific marker is divided by the concentration of the general indicator or marker, thus establishing a ratio. However, in order for the ratio method to work, the markers quantified in the ratio must have similar decay patterns across time (Field and Samadpour, 2007). Of the ruminant and cattle markers tested in this study, only Rum2Bac had a curve that was not different than that of GenBac3 ($p > 0.1$).

When the ratios of host specific markers versus GenBac3, as determined from the best fitting models, are plotted over time, it is clear that Rum2Bac has the most consistent ratio over time, although neither the GenBac3 nor Rum2Bac models decay in a linear fashion (Figure 3.3). However, the initial concentrations of Rum2Bac and GenBac3 were similar, suggesting that at least in cattle feces, the two assays are detecting similar targets, explaining their similar decay. If additional sources were added, such as human feces, GenBac3 would presumably detect more diverse targets, and thus its decay might not behave similarly to the decay of GenBac3 from ruminant sources. This has been shown for other indicators (Korajkic et al., 2013). Thus, more work is necessary to assess the potential applicability of these markers in source allocation by the ratio method.

Our data were collected using samples from a single herd of dairy cattle. Past studies have found variability in concentrations of FIB based on feeding practices (Shanks et al., 2011) and animal age (Shanks et al., 2014). Thus, the shapes described here may not be applicable with fecal matter from other types of cattle. Additionally, this study tested decay in one setting, using mesocosms filled with river water. Results obtained here may not be applicable in other water sources. Finally, as with all experiments conducted in artificial settings, it is important to understand what effects are induced by the experimental design. While the mesocosms used in this study were large and filled with unaltered river water, it is possible that the closed system induced artifacts that are reflected in the measurements, and may not make decay estimates applicable in natural settings.

References

Ahmed, W., Gyawali, P., Sidhu, J. p. s., and Toze, S. (2014). Relative inactivation of faecal indicator bacteria and sewage markers in freshwater and seawater microcosms. Lett. Appl. Microbiol. 348–354.

Ahmed, W., Harwood, V.J., Nguyen, K., Young, S., Hamilton, K., and Toze, S. (2016). Utility of *Helicobacter spp.* associated GFD markers for detecting avian fecal pollution in natural waters of two continents. Water Res. *88*, 613–622.

Ahmed, W., Triplett, C., Gomi, R., Gyawali, P., Hodgers, L., and Toze, S. (2015). Assessment of genetic markers for tracking the sources of human wastewater associated *Escherichia coli* in environmental waters. Environ. Sci. Technol. *49*, 9341–9346.

Akaike, H. (1992). Information theory and an extension of the maximum likelihood principle. In Breakthroughs in Statistics, S. Kotz, and N.L. Johnson, eds. (Springer New York), pp. 610–624.

Anderson, K.L., Whitlock, J.E., and Harwood, V.J. (2005). Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Appl. Environ. Microbiol. *71*, 3041–3048.

Bae, S., and Wuertz, S. (2011). Survival of host-associated *Bacteroidales* cells and their relationship with *Enterococcus spp.*, *Campylobacter jejuni*, *Salmonella enterica*, *Serovar Typhimurium*, and Adenovirus in freshwater microcosms as measured by propidium monoazide-Quantitative PCR. Appl. Environ. Microbiol. *78*, 922–932.

Bell, A., Layton, A.C., McKay, L., Williams, D., Gentry, R., and Sayler, G.S. (2009). Factors influencing the persistence of fecal *Bacteroides* in stream water. J. Environ. Qual. *38*, 1224.

Bernhard, A.E., and Field, K.G. (2000). A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Appl. Environ. Microbiol. *66*, 4571–4574.

Boehm, A.B., Van De Werfhorst, L.C., Griffith, J.F., Holden, P.A., Jay, J.A., Shanks, O.C., Wang, D., and Weisberg, S.B. (2013). Performance of forty-one microbial source tracking methods: a twenty-seven lab evaluation study. Water Res. *47*, 6812–6828.

Bolker, B., Skaug, H., and Laake, J. (2015). R2admb: "ADMB" to R Interface Functions.

Dick, L.K., and Field, K.G. (2004). Rapid estimation of numbers of fecal *Bacteroidetes* by use of a quantitative PCR assay for 16S rRNA genes. Appl. Environ. Microbiol. *70*, 5695–5697.

Dick, L.K., Stelzer, E.A., Bertke, E.E., Fong, D.L., and Stoeckel, D.M. (2010). Relative decay of *Bacteroidales* microbial source tracking markers and cultivated *Escherichia coli* in freshwater microcosms. Appl. Environ. Microbiol. *76*, 3255–3262.

Field, K.G., and Samadpour, M. (2007). Fecal source tracking, the indicator paradigm, and managing water quality. Water Res. *41*, 3517–3538.

Fiksdal, L., Maki, J.S., LaCroix, S.J., and Staley, J.T. (1985). Survival and detection of *Bacteroides spp.*, prospective indicator bacteria. Appl. Environ. Microbiol. *49*, 148–150.

Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., and Sibert, J. (2012). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Methods Softw. *27*, 233–249.

Geeraerd, A.H., Herremans, C.H., and Van Impe, J.F. (2000). Structural model requirements to describe microbial inactivation during a mild heat treatment. Int. J. Food Microbiol. *59*, 185–209.

Geldreich, E.E. (1970). Applying bacteriological parameters to recreational water quality. J. Am. Water Works Assoc. *62*, 113–120.

Green, H.C., Shanks, O.C., Sivaganesan, M., Haugland, R.A., and Field, K.G. (2011). Differential decay of human faecal *Bacteroides* in marine and freshwater. Environ. Microbiol. *13*, 3235–3249.

Haugland, R.A., Siefring, S.C., Wymer, L.J., Brenner, K.P., and Dufour, A.P. (2005). Comparison of *Enterococcus* measurements in freshwater at two recreational beaches by quantitative polymerase chain reaction and membrane filter culture analysis. Water Res. *39*, 559–568.

Haugland, R.A., Varma, M., Sivaganesan, M., Kelty, C., Peed, L., and Shanks, O.C. (2010). Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected *Bacteroidales* species and human fecal waste by qPCR. Syst. Appl. Microbiol. *33*, 348–357.

Jeanneau, L., Solecki, O., Wéry, N., Jardé, E., Gourmelon, M., Communal, P.-Y., Jadas-Hécart, A., Caprais, M.-P., Gruau, G., and Pourcher, A.-M. (2012). Relative decay of fecal indicator bacteria and human-associated markers: A microcosm study simulating wastewater input into seawater and freshwater. Environ. Sci. Technol. *46*, 2375–2382.

Korajkic, A., McMinn, B.R., Harwood, V.J., Shanks, O.C., Fout, G.S., and Ashbolt, N.J. (2013). Differential decay of enterococci and *Escherichia coli* originating from two fecal pollution sources. Appl. Environ. Microbiol. *79*, 2488–2492.

Kreader, C.A. (1998). Persistence of PCR-Detectable *Bacteroides distasonis* from human feces in river water. Appl. Environ. Microbiol. *64*, 4103–4105.

Liang, Z., He, Z., Zhou, X., Powell, C.A., Yang, Y., Roberts, M.G., and Stoffella, P.J. (2012). High diversity and differential persistence of fecal *Bacteroidales* population spiked into freshwater microcosm. Water Res. *46*, 247–257.

Mieszkin, S., Yala, J.-F., Joubrel, R., and Gourmelon, M. (2010). Phylogenetic analysis of *Bacteroidales* 16S rRNA gene sequences from human and animal effluents and assessment of ruminant faecal pollution by real-time PCR. J. Appl. Microbiol. *108*, 974–984.

Motulsky, H., and Christopoulos, A. (2004). Fitting models to biological data using linear and nonlinear regression: A practical guide to curve fitting (Oxford University Press, USA).

R Core Team (2015). R: A language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing).

Raith, M.R., Kelty, C.A., Griffith, J.F., Schriewer, A., Wuertz, S., Mieszkin, S., Gourmelon, M., Reischer, G.H., Farnleitner, A.H., Ervin, J.S., et al. (2013). Comparison of PCR and quantitative real-time PCR methods for the characterization of ruminant and cattle fecal pollution sources. Water Res. *47*, 6921–6928.

Schriewer, A., Goodwin, K.D., Sinigalliano, C.D., Cox, A.M., Wanless, D., Bartkowiak, J., Ebentier, D.L., Hanley, K.T., Ervin, J., Deering, L.A., et al. (2013). Performance evaluation of canine-associated *Bacteroidales* assays in a multi-laboratory comparison study. Water Res. *47*, 6909–6920.

Seurinck, S., Defoirdt, T., Verstraete, W., and Siciliano, S.D. (2005). Detection and quantification of the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. Environ. Microbiol. *7*, 249–259.

Shanks, O.C., Atikovic, E., Blackwood, A.D., Lu, J., Noble, R.T., Domingo, J.S., Seifring, S., Sivaganesan, M., and Haugland, R.A. (2007). Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. Appl. Environ. Microbiol. *74*, 745–752.

Shanks, O.C., White, K., Kelty, C.A., Hayes, S., Sivaganesan, M., Jenkins, M., Varma, M., and Haugland, R.A. (2010). Performance assessment PCR-based assays targeting *Bacteroidales* genetic markers of bovine fecal pollution. Appl. Environ. Microbiol. *76*, 1359–1366.

Shanks, O.C., Kelty, C.A., Archibeque, S., Jenkins, M., Newton, R.J., McLellan, S.L., Huse, S.M., and Sogin, M.L. (2011). Community structures of fecal bacteria in cattle from different animal feeding operations. Appl. Environ. Microbiol. *77*, 2992–3001.

Shanks, O.C., Sivaganesan, M., Peed, L., Kelty, C.A., Blackwood, A.D., Greene, M.R., Noble, R.T., Bushon, R.N., Stelzer, E.A., Kinzelman, J., et al. (2012). Interlaboratory comparison of Real-Time PCR protocols for quantification of general fecal indicator bacteria. Environ. Sci. Technol. *46*, 945–953.

Shanks, O.C., Kelty, C.A., Peed, L., Sivaganesan, M., Mooney, T., and Jenkins, M. (2014). Age-related shifts in the density and distribution of genetic marker water quality indicators in cow and calf feces. Appl. Environ. Microbiol. *80*, 1588–1594.

Shanks, O.C., Kelty, C.A., Oshiro, R., Haugland, R.A., Madi, T., Brooks, L., Field, K.G., and Sivaganesan, M. (2016). Data acceptance criteria for standardized human-associated fecal source identification quantitative Real-Time PCR methods. Appl. Environ. Microbiol. AEM.03661-15.

Siefring, S., Varma, M., Atikovic, E., Wymer, L., and Haugland, R.A. (2008). Improved real-time PCR assays for the detection of fecal indicator bacteria in surface waters with different instrument and reagent systems. J. Water Health *6*, 225–237.

Sinton, L., Hall, C., and Braithwaite, R. (2007). Sunlight inactivation of *Campylobacter jejuni* and *Salmonella enterica*, compared with *Escherichia coli*, in seawater and river water. J. Water Health *5*, 357–365.

Sinton, L.W., Hall, C.H., Lynch, P.A., and Davies-Colley, R.J. (2002). Sunlight inactivation of fecal indicator bacteria and bacteriophages from waste stabilization pond effluent in fresh and saline waters. Appl. Environ. Microbiol. *68*, 1122–1131.

Sokolova, E., Åström, J., Pettersson, T.J.R., Bergstedt, O., and Hermansson, M. (2012). Decay of *Bacteroidales* genetic markers in relation to traditional fecal indicators for water quality modeling of drinking water sources. Environ. Sci. Technol. *46*, 892–900.

Solecki, O., Jeanneau, L., Jardé, E., Gourmelon, M., Marin, C., and Pourcher, A.M. (2011). Persistence of microbial and chemical pig manure markers as compared to faecal indicator bacteria survival in freshwater and seawater microcosms. Water Res. *45*, 4623–4633.

Soller, J.A., Schoen, M.E., Bartrand, T., Ravenscroft, J.E., and Ashbolt, N.J. (2010). Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. Water Res. *44*, 4674–4691.

Staley, C., Gordon, K.V., Schoen, M.E., and Harwood, V.J. (2012). Performance of two quantitative PCR methods for microbial source tracking of human sewage and implications for microbial risk assessment in recreational waters. Appl. Environ. Microbiol. *78*, 7317–7326.

Swinnen, I.A.M., Bernaerts, K., Dens, E.J.J., Geeraerd, A.H., and Van Impe, J.F. (2004). Predictive modelling of the microbial lag phase: a review. Int. J. Food Microbiol. *94*, 137–159.

Tambalo, D.D., Fremaux, B., Boa, T., and Yost, C.K. (2012). Persistence of host-associated *Bacteroidales* gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. Water Res. *46*, 2891–2904.

USEPA, 2015. Assesment TMDL tracking and implementation system (ATTAINS), on National Water Quality Inventory Report to Congress.

Walters, S.P., and Field, K.G. (2009). Survival and persistence of human and ruminant-specific faecal *Bacteroidales* in freshwater microcosms. Environ. Microbiol. *11*, 1410–1421.

Wang, D., Farnleitner, A.H., Field, K.G., Green, H.C., Shanks, O.C., and Boehm, A.B. (2013). Enterococcus and Escherichia coli fecal source apportionment with microbial source tracking genetic markers – Is it feasible? Water Res.

Rstudio Team (2015). RStudio: Integrated Development for R. (Boston, MA: RStudio, Inc.).

Figures and Tables

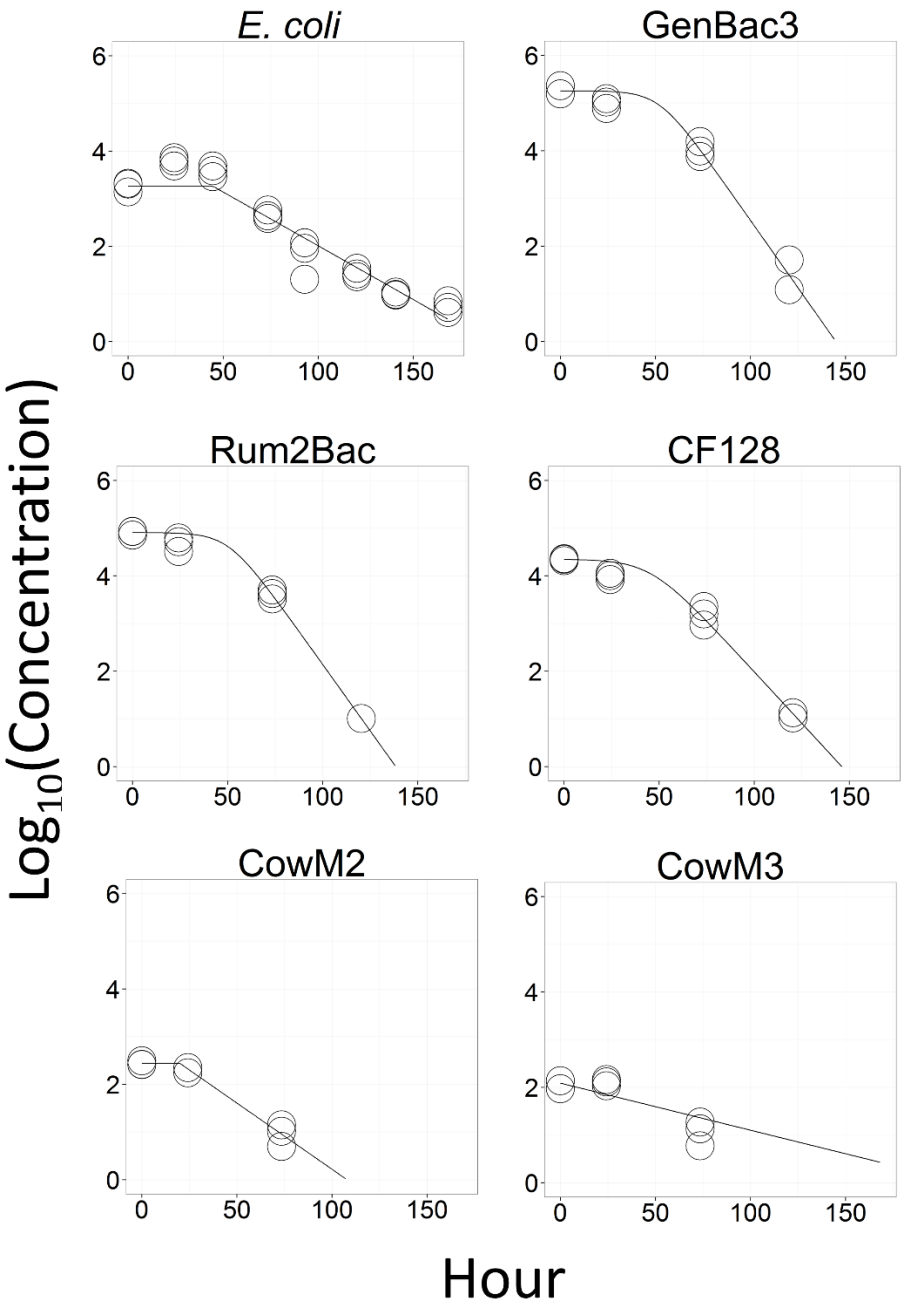

Figure 3.1 Mathematical models fit to measured concentrations for different indicators. Concentration values represent CFU/100 ml for *E. coli* or copies/100 ml for markers.

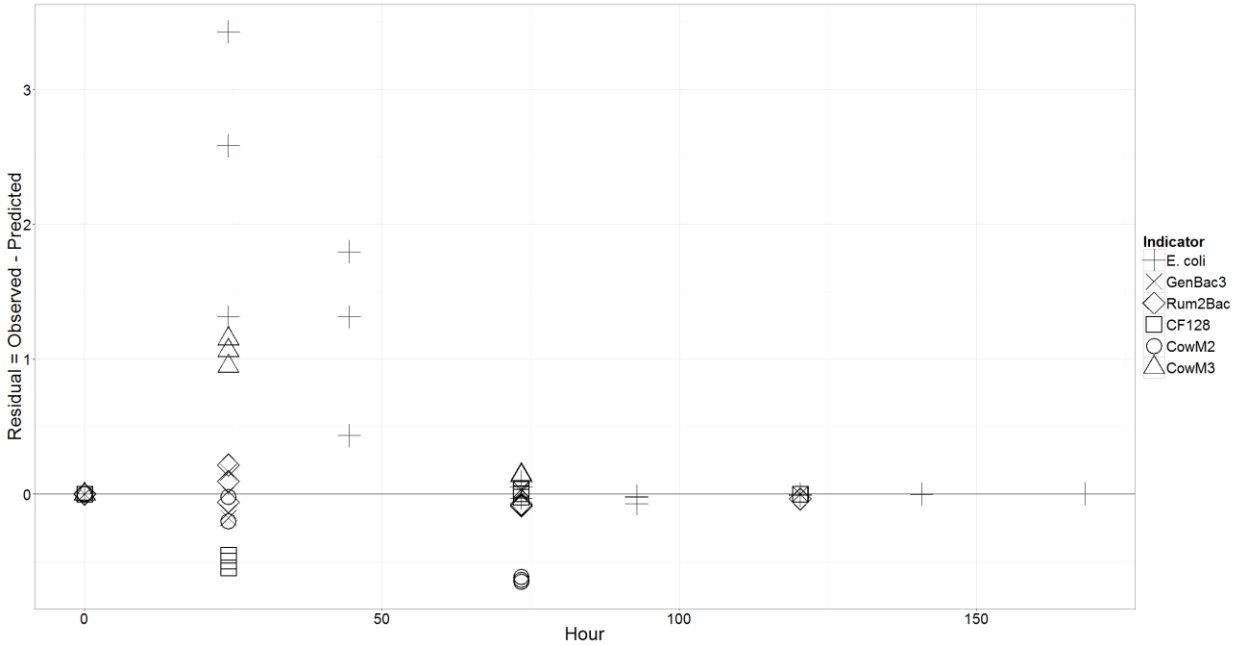

Figure 3. 2 Residual errors plotted for best fitting models reveal discrepancy between observed and predicted.

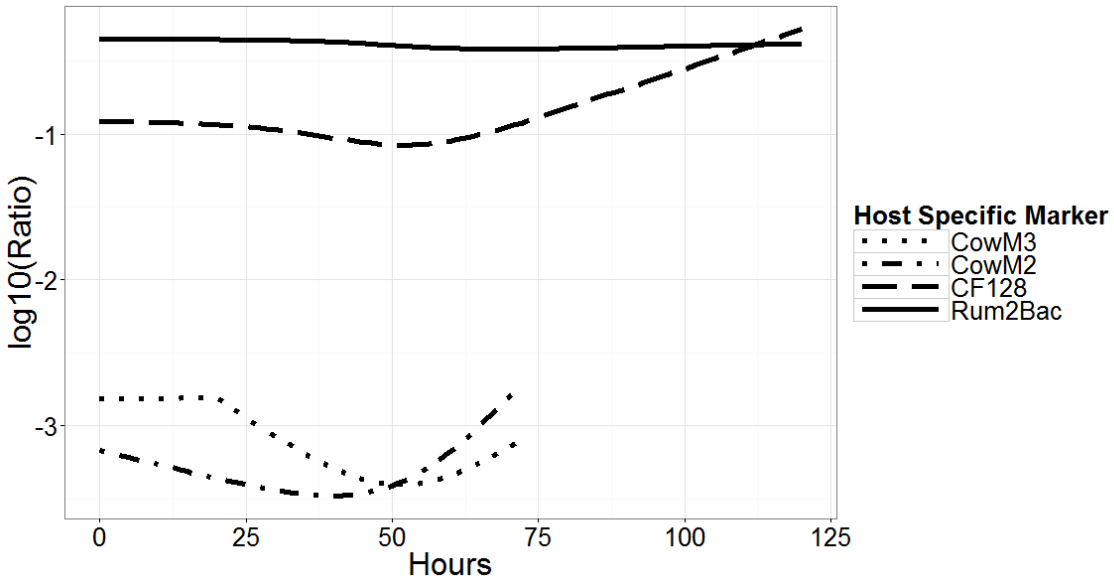

Figure 3.3 Ratios of the concentration of ruminant or cattle associated markers to the GenBac3 marker (predicted by model fits) were plotted over time to assess the feasibility of source allocation.

Table 3.1 Primers and probes used in this study

Target	Primers and Probes	Reference
GenBac3 (General *Bacteroidales*)	Forward: GGGGTTCTGAGAGGAAGGT Reverse:CCGTCATCCTTCACGCTACT Probe:[FAM]CAATATTCCTCACTGCTGCCTCCCGTA[TAMRA]	(Dick and Field, 2004; Siefring et al., 2008)
CF128 (Ruminant Associated *Bacteroidales*)	Forward:CCAACYTTCCCGWTACTC Reverse: ACCCCGCCTACTATCTAATG	(Bernhard and Field, 2000; Seurinck et al., 2005)
Rum2Bac (Ruminant Associated *Bacteroidales*)	Forward: ACAGCCCGCGATTGATACTGGTAA Reverse: CAATCGGAGTTCTTCGTGAT Probe: [FAM]-ATGAGGTGGATGGAATTCGTGGTGT-[BHQ-1]	(Mieszkin et al., 2010)
CowM2 (Bovine Associated HDIG domain protein)	Forward: CGGCCAAATACTCCTGATCGT Reverse: GCTTGTTGCGTTCCTTGAGATAAT Probe: [FAM]-AGGCACCTATGTCCTTTACCTCATCAACTACAGACA-[TAMRA]	(Shanks et al., 2007)
CowM3 (Bovine Associated sialic acid-specific 9-0 acetylesterase secretory protein homolog)	Forward: CCTCTAATGGAAAATGGATGGTATCT Reverse: CCATACTTCGCCTGCTAATACCTT Probe: [FAM]-TTATGCATTGAGCATCGAGGCC-[TAMRA]	(Shanks et al., 2007)

Table 3. 2 Statistics of Delayed Chick Watson (DCW) and Geeraerd Shoulder (GS) model fits for individual datasets.

Dataset	Best Fitting Model	Degrees of Freedom	Sum of Squared Errors	Akaike Information Criterion	Probably DCW Best Model	Probably GS Best Model
E. coli	DCW	21	12.8596	-10.4037	100	<1
CF128	GS	8	1.97298	-13.4019	<1	100
GenBac	GS	8	1.70465	-15.0099	<1	100
Rum2Bac	GS	7	0.903411	-18.3273	<1	100
CowM2	DCW	5	0.55874	-14.8921	100	<1
CowM3	CW	6	1.27413	-12.0308	100	<1

Table 3.3 Summary statistics for individual curves fit with the Delayed Chick Watson (DCW) and Geeraerd Shoulder (GS) models. Standard error estimates are shown in parentheses.

Target	Best Fitting Model	Mean $\log_{10}(C_0)$ / 100 ml	Lag or Shoulder (Hours)	Decay Rate ($Hour^{-1}$)	Last Time Point (Hours)
Cultivable *E. coli*	DCW	3.26 (2.60)	44.5 (0.094)	0.0520 (0.0022)	168
GenBac3	GS	6.96 (6.34)	52.0 (5.4)	0.130 (0.015)	120
CF128	GS	6.04 (4.73)	45.4 (9.1)	0.0994 (0.017)	120
Rum2Bac	GS	6.61 (5.60)	50.4 (6.0)	0.129 (0.019)	120
CowM2	DCW	4.14 (3.17)	19.9 (8.5)	0.0639 (0.0131)	72
CowM3	CW	3.79 (3.10)	NA	0.0227 (0.0063)	72

Table 3.4 Global Model Fitting results for comparisons of different markers or indicator bacteria. Bolded values represent significance.

Dataset	*E. coli*	GenBac3	CF128	Rum2Bac	CowM2
E. coli					
GenBac3	**P <0.001**				
CF128	**P <0.001**	**P < 0.1**			
Rum2Bac	**P <0.001**	P > 0.1	**P < 0.1**		
CowM2	**P <0.001**	P > 0.1	**P < 0.05**	P > 0.1	
CowM3	P > 0.1	P > 0.1	P > 0.1	P > 0.1	P > 0.1

Table S3.1 Standard curve calibration equations and range of quantitation for general and host-specific qPCR assays

Assay	Standard curve equation	Efficiency	R^2	Range of quantitation	Range of quantitation (/100 ml)
GenBac3	Y = -3.535x+36.95	91.806	.991	10^8-10^2	10^9-10^3
CF128	Y = -3.834x +37.358	82.234	.999	10^7-10^2	10^8-10^3
Rum2Bac	Y = -3.455x+39.055	94.305	.996	10^8-10^1	10^9-10^2
CowM2	Y = -3.499x+38.193	93.099	.998	10^8-10^1	10^9-10^2
CowM3	Y = -3.464x + 35.212	94.391	.993	10^6-10^2	10^7-10^3

Table S3.2 Statistics of best fitting models for curve comparisons

Target 1	Target 2	Separate Sum of Squared Errors	Separate Degrees of Freedom	Shared Sum of Squared Errors	Shared Degrees of Freedom	Sum of Squares Relative Difference	Degrees of Freedom Relative Difference	F Ratio	P value
E. coli	GenBac3	14.56425	29	66.827	32	358.8427	10.34483	34.68813	3.57E-10
E. coli	Rum2Bac	13.76301	28	42.5255	32	208.984	14.28571	14.62888	6.75E-07
E. coli	CF128	14.83258	29	58.6835	33	295.6392	13.7931	21.43384	8.63E-09
E. coli	CowM2	13.41834	26	22.7508	29	69.55003	11.53846	6.02767	0.002544
E. coli	CowM3	14.13373	27	15.7249	30	11.25796	11.11111	1.013217	0.40057
GenBac3	Rum2Bac	2.608061	15	2.64431	18	1.389883	20	0.069494	0.975479
GenBac3	CF128	3.67763	16	5.72312	19	55.61979	18.75	2.966389	0.058053
GenBac3	CowM2	2.26339	13	2.82787	16	24.93958	23.07692	1.080715	0.385386
GenBac3	CowM3	2.97878	14	3.06592	17	2.925359	21.42857	0.136517	0.936857
Rum2Bac	CF128	2.876391	15	4.42915	18	53.98289	20	2.699145	0.076422
Rum2Bac	CowM2	1.462151	12	1.80727	15	23.60351	25	0.944141	0.444075
Rum2Bac	CowM3	2.177541	13	2.47435	16	13.63047	23.07692	0.590654	0.630006
CF128	CowM2	2.53172	13	5.59767	16	121.1015	23.07692	5.24773	0.010324
CF128	CowM3	3.24711	14	3.25469	17	0.233438	21.42857	0.010894	0.998379
CowM2	CowM3	1.83287	11	3.32864	14	81.60808	27.27273	2.992296	0.066777

CHAPTER 4

SIMULATIONS TO TEST A CONTINUOUS-FLOW STIRRED TANK REACTOR AS AN OPEN SYSTEM TO STUDY FECAL INDICATOR BACTERIA DECAY

Lauren E. Brooks and Katharine G. Field

In preparation for Environmental Science and Technology Letters

Abstract

Model systems used to investigate the decay of fecal indicator bacteria differ in many respects from the environments they are intended to simulate. One major difference is the reactor type used to conduct decay studies, which typically relies on closed, recirculating conditions. Open systems are usually avoided in decay studies, as it is challenging to distinguish between removal from the system and loss due to decay. Here, we present an adjustment to observed concentration data generated from a Continuous-flow Stirred Tank Reactor that could be used to distinguish removal and decay. We tested the performance of this adjustment across a range of decay shapes and flow rates to assess the feasibility of this type of system for use in decay studies. Our simulations revealed that the proposed adjustment provides a reliable way to compare closed and open systems.

Introduction

Fecal indicator bacteria (FIB) and molecular markers are used to signify the presence of fecal matter in recreational waters. To be used as reliable indicators of fecal contamination, the loss of measurability, referred to as decay, of indicators must be understood. While the systems selected to study decay have varied, there has been no investigation of how the model system selected to study FIB or marker decay impacts findings. Most decay studies are conducted using a homogenously mixed microcosm consisting of a volume of water seeded with fecal matter from a selected source (Ahmed et al., 2014; Anderson et al., 2005; Green et al., 2011; Jeanneau et al., 2012; Noble et al., 2004; Schulz and Childers, 2011; Sinton et al., 2002; Sokolova et al., 2012; Solecki et al., 2011; Tambalo et al., 2012). An alternative approach employs dialysis bags attached to a submersible mesocosm (Bae and Wuertz, 2011, 2009b, Korajkic et al., 2014, 2013a) , thus allowing for exchange of small materials with the environment. However, both of these systems constrain microorganisms within the system being studied, introducing a potentially significant difference between the model systems and the environments they are intended to model. To date, no study has explored the biases introduced by conducting decay studies in these closed systems.

The Continuous-flow Stirred Tank Reactor (CSTR) is a mixed reactor with homogenous water matrix composition throughout the reactor volume. Additionally, in a CSTR, a controlled volume is continuously added to the reactor, which is held at constant volume, and thus an equivalent volume is continuously removed. This setup is commonly used as a model for well mixed ponds or lakes, but can also be used as a preliminary step for studying more complex systems, as it eliminates the need to account for processes that occur in the environment, such as dispersion. By using an idealized reactor, certain characteristics of interest, such as fate of a

contaminant, can be studied, while other properties, such as dispersion, can be controlled and thus explicitly accounted for by modeling efforts.

In an open system, changes in concentration are not only caused by decay, but also by removal from the system as a result of flow. Open systems are usually avoided in decay studies, as it is challenging to distinguish between removal from the system and loss due to decay. To use such a system to study persistence of FIB, adjustments to the measured concentration must be made to account for loss due to removal, allowing for modeling of loss due to decay. Here, we establish an adjustment that can be incorporated into previously developed decay models, and calculate an adjusted concentration for use in decay calculations and models. Our objective was to use simulations to evaluate the possibility of using a CSTR as an alternative experimental design for the study of FIB, in which the system is open with respect to all water constituents.

Methods

Theoretical development of the concentration adjustment

To develop the adjustment that could distinguish loss as a result of decay from loss due to removal, we started with mass balance calculations associated with CSTRs:

Eq. 1: Accumulation (time rate of change) = ± transport ± sources/sinks ± kinetic changes

Assuming proper mixing within the tank, and negligible inputs relative to the high concentration of FIB added, the equation can be simplified and rewritten specifically for a decay study:

Eq 2: $-\Delta C/\Delta t = C_{t-1} - C_t$ = Total Loss = Decay + Removal

Defining total loss as the change in the observed concentration during time *t*, this will include removal of FIB from the system as well as loss due to a loss of measurability. This can be written

dynamically as shown in Equation 2, or as is more common for decay studies, it can be written statically based on the concentration at time 0 (C_0)

Eq 3: $C_t = C_0$*Decay*Removal

Additionally, making the assumption that homogenous mixing has been achieved, and as a result, each bacterium has an equally likely chance of being removed from the system, removal can be modeled using a first order model with the rate of removal being equal to the flow rate (Q) divided by the volume (V) of the reactor (Ramaswami et al., 2005). If we let the rate of removal be expressed as $k_f = Q/V$, then the static equation can be rewritten as:

Eq 4: $C_t = C_0$*Decay*exp(-k_f*time)

Finally, we rearrange this equation by dividing both sides by the removal term, resulting in:

Eq 5: $C_{adj} = C_t/\exp(-k_f*time) = C_0$*Decay

This is equivalent to the typical models used to describe data collected in decay studies with C_{adj} in place of C_t. The decay term can then be modeled by an appropriate decay model using the adjusted concentration data.

Simulations to verify concept

We conducted a series of simulations to test the ability of decay models to estimate decay parameters from an open system using adjusted concentrations. All simulations and analyses were conducted in R (R Core Team, 2015) using the Rstudio interface (2015). Data were generated with no replicates and negligible variability, and were intended to represent idealized data that could be observed from a decay study.

Prior to conducting the open system simulations, we tested the ability of the models to fit the data generated by combinations of the decay parameters in a simulated closed system (i.e. no simulated flow). Decay profiles were simulated using parameters based on the range of values

for common decay shapes, including those with a shoulder and a tailing feature (Table 4.1). Decay models simulated included the Chick Watson (CW) model, the Delayed Chick Watson model (DCW), and the Chick Watson with Tail model (CWT). Combinations of decay parameters resulted in 5 simulations from the CW model, 25 from the DCW model, and 15 from the CWT model.

Data were generated using the equation for each decay model (Table 4.1), assuming a sampling frequency of every 24 hours for 14 days, or until the simulated concentration was less than 1. Datasets consisting of three or fewer data points were removed from analysis. For each simulated data set, decay models were fit using the R2ADMB package (Bolker et al., 2015), both with the flow adjustment and without.

Models fit to each simulated dataset were screened for successful fits based on whether or not estimates were provided for each parameter in the model. Additionally, fits were screened for imprecise estimates, as determined by wide ranging standard errors for the estimates. Models that provided estimates with standard errors greater than the model parameter were not considered to be successfully fit. For some combinations of model parameters, simulated model features (i.e. lag period or tailing) were not evident in the data, and the simulated model failed to fit the dataset. Combinations of parameters that were not successfully fit by the correct model were excluded from further analysis.

Combinations of decay parameters that were successfully fit by the simulated model for a no flow system were combined with simulated flow rates to generate concentrations observed from an open system. Flow rates were simulated ranging from 0 to 10 L per hour with the simulated volume set to 100 L. Final counts of simulations were 99 from the CW model, 233 from the DCW model, and 101 from the CWT model.

Simulated concentrations were fit to both the observed concentration with no adjustment and the adjusted concentrations and assessed for accuracy of parameter estimates. Additionally, to assess whether inaccurate estimates were the result of the adjustment or the shortened experiment resulting from removal in addition to decay, end point fits were calculated using the same model parameters under a no flow scenario, but datasets were shortened to the same duration as observed for the simulated datasets that included a flow rate.

The simulations were assessed based on the ability of the best fitting model to accurately estimate the simulated decay parameters as well as to correctly describe the simulated data shape, as determined using Akaike Information Criterion (AIC) values (Akaike, 1992) to determine the best fitting model for each simulated dataset. Simulations for each shape were summarized by the percent of simulations that correctly described the data shape, and the correlation coefficient (R^2) between the simulated or end point values and estimated values.

Results and Discussion

We simulated data for common decay models in a CSTR under a range of flow rates to assess the possibility of using an open system to study fecal indicator decay. Because of contaminant removal by flow-through, the contaminant of interest becomes undetectable in a CSTR earlier than in a closed system with an identical decay profile, shortening the duration of experiments. Our simulations demonstrated that the effects of this shortened duration had different impacts on our ability to detect and estimate decay parameters, varying across decay shapes.

For data simulated from the CW decay model, when no adjustment was made to the concentration there was a poor correlation between the estimated decay rate and the simulated

decay rate (R^2 = 0.3883). However, when the adjustment was made, regardless of the duration of the experiment there was a perfect correlation between the simulated and estimated values. This similarity suggests that the adjustment provided a reliable method for estimating the decay rate, and that the duration does not matter for linear decay data. The DCW model was selected to represent decay profiles that include a lag or shoulder period in the data. Decay curves with this shape have a period of little to no change in the beginning, followed by exponential decay. Simulations conducted based on a DCW model of decay showed a poor correlation between the estimated decay parameters (decay rate and lag duration) when no correction was made (Table 4.1). Unlike the CW based simulations, those that included a lag period were still not perfectly correlated once the adjustment term was included in the models (R^2 = .8505 for decay rate, 0.5018 for lag duration).

As the decay shape is not identical throughout time for this model, estimates of model parameters appeared inaccurate when duration of the experiment was not considered. However, if instead of comparing the estimated values to the simulated values, the data are compared to the endpoint value, in which the data were generated based on no flow, but were cut off so that the simulation only lasted the same duration as those with flow, perfect correlation (R^2 = 1) was observed. Thus, when fitting decay rate models of the DCW shape, the estimates are only valid for the time frame modeled, and cannot be compared past this period.

The ability to detect the lag feature in the data depended on the simulated flow rate. Dataset simulated for the DCW model high flow rates were not fit by the correct model (Figure 4.1). Instead, only 48% of the simulations correctly identified the lag feature, but those estimates were accurate based on the endpoint values. This is because the shortened experimental duration results in inadequate data points to justify the inclusion of the lag parameter in the model.

Data simulated from the Chick Watson with Tail (CWT) model also experience differential decay rates throughout the data. In this shape, however, the period of minimal decay occurs following the exponential decay period, rather than before. The tailing feature was not detected by 58% of the simulations from the CWT model, the majority of which were at high flow rates (Figure 4.1). However, the datasets that were fit successfully by the CWT model accurately estimated the simulated values, regardless of duration. Thus, the tailing feature, if reached during the shortened experimental duration of the open system, was accurately estimated.

While numerous studies have been conducted to measure FIB decay, characterization of potential biases introduced by a closed system to study FIB decay has never been studied. Before any conclusions drawn from decay studies can be extended to the environment, it is essential to understand the effects of the closed system on decay rate estimates. Here, we have provided an adjustment to the typical decay models that allows for the use of an open system to study FIB decay, thus enabling comparisons to be made between open and closed systems. The correction we defined is based on a known flow rate, and still required a simplified artificial setting in the form of a CSTR, but is an improvement over the classic methods of measuring FIB decay, as it allows for an exchange of the microbial community between the contaminated and clean waters. This adjustment can be applied to any cultivated bacterial indicator or genetic marker used to estimate decay rates for any target of interest.

References

Ahmed, W., Gyawali, P., Sidhu, J.., and Toze, S. (2014). Relative inactivation of faecal indicator bacteria and sewage markers in freshwater and seawater microcosms. Lett. Appl. Microbiol. 348–354.

Akaike, H. (1992). Information theory and an extension of the maximum likelihood principle. In Breakthroughs in Statistics, S. Kotz, and N.L. Johnson, eds. (Springer New York), pp. 610–624.

Anderson, K.L., Whitlock, J.E., and Harwood, V.J. (2005). Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Appl. Environ. Microbiol. *71*, 3041–3048.

Anu Ramaswami, Jana B. Milford, and Mitchell J. Small (2005). Integrated Environmental Modeing: Pollutant Transport, Fate, and Risk in the Environment (John Wiley & Sons, Inc).

Bae, S., and Wuertz, S. (2009). Discrimination of viable and dead fecal *Bacteroidales* bacteria by quantitative PCR with propidium monoazide. Appl. Environ. Microbiol. *75*, 2940–2944.

Bae, S., and Wuertz, S. (2011). Survival of host-associated *Bacteroidales* cells and their relationship with *Enterococcus spp.*, *Campylobacter jejuni*, *Salmonella enterica*, *Serovar Typhimurium*, and Adenovirus in freshwater microcosms as measured by propidium monoazide-Quantitative PCR. Appl. Environ. Microbiol. *78*, 922–932.

Bolker, B., Skaug, H., and Laake, J. (2015). R2admb: "ADMB" to R Interface Functions.

Field, K.G., and Samadpour, M. (2007). Fecal source tracking, the indicator paradigm, and managing water quality. Water Res. *41*, 3517–3538.

Green, H.C., Shanks, O.C., Sivaganesan, M., Haugland, R.A., and Field, K.G. (2011). Differential decay of human faecal *Bacteroides* in marine and freshwater. Environ. Microbiol. *13*, 3235–3249.

Jeanneau, L., Solecki, O., Wéry, N., Jardé, E., Gourmelon, M., Communal, P.-Y., Jadas-Hécart, A., Caprais, M.-P., Gruau, G., and Pourcher, A.-M. (2012). Relative decay of fecal indicator bacteria and human-associated markers: A microcosm study simulating wastewater input into seawater and freshwater. Environ. Sci. Technol. *46*, 2375–2382.

Korajkic, A., McMinn, B.R., Harwood, V.J., Shanks, O.C., Fout, G.S., and Ashbolt, N.J. (2013). Differential decay of enterococci and *Escherichia coli* originating from two fecal pollution sources. Appl. Environ. Microbiol. *79*, 2488–2492.

Korajkic, A., McMinn, B.R., Shanks, O.C., Sivaganesan, M., Fout, G.S., and Ashbolt, N.J. (2014). Biotic interactions and sunlight affect persistence of fecal indicator bacteria and microbial source tracking genetic markers in the upper Mississippi River. Appl. Environ. Microbiol. *80*, 3952–3961.

Noble, R.T., Lee, I.M., and Schiff, K. c. (2004). Inactivation of indicator micro-organisms from various sources of faecal contamination in seawater and freshwater. J. Appl. Microbiol. *96*, 464–472.

R Core Team (2015). R: A language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing).

Rstudio Team (2015). RStudio: Integrated Development for R. (Boston, MA: RStudio, Inc.).

Schulz, C.J., and Childers, G.W. (2011). Fecal *Bacteroidales* diversity and decay in response to variations in temperature and salinity. Appl. Environ. Microbiol. *77*, 2563–2572.

Sinton, L.W., Hall, C.H., Lynch, P.A., and Davies-Colley, R.J. (2002). Sunlight inactivation of fecal indicator bacteria and bacteriophages from waste stabilization pond effluent in fresh and saline waters. Appl. Environ. Microbiol. *68*, 1122–1131.

Sokolova, E., Åström, J., Pettersson, T.J.R., Bergstedt, O., and Hermansson, M. (2012). Decay of *Bacteroidales* genetic markers in relation to traditional fecal indicators for water quality modeling of drinking water sources. Environ. Sci. Technol. *46*, 892–900.

Solecki, O., Jeanneau, L., Jardé, E., Gourmelon, M., Marin, C., and Pourcher, A.M. (2011). Persistence of microbial and chemical pig manure markers as compared to faecal indicator bacteria survival in freshwater and seawater microcosms. Water Res. *45*, 4623–4633.

Tambalo, D.D., Fremaux, B., Boa, T., and Yost, C.K. (2012). Persistence of host-associated *Bacteroidales* gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. Water Res. *46*, 2891–2904.

Table 4.1 Models and parameter values simulated to test the utility of the adjustment. Performance was assessed using the R^2 correlation coefficient for each comparison (simulated values and raw data, simulated values and adjusted data, and end point estimates and adjusted data) and the percent of simulations that were matched to the correct model.

Simulated Model	**Static Decay Model**	**Simulated Values**	**Simulated parameters / Estimated (no adjustment) R^2**	**Simulated parameter/ Estimated (adjustment) R^2**	**Estimated (end point)/ Estimated (adjustment) R^2**	**Corrected % Shape Match**
Chick Watson	$C_t = C_0 e^{-kt}$	k (hr^{-1}) = 0.001, 0.005, 0.01, 0.05, 0.1	k: 0.3883	k: 1	k: 1	100%
Delayed Chick Watson	If t <= lag: $C_t = C_0$ If t > lag: $C_t = C_0 e^{-k(t-lag)}$	k (hr^{-1}) = 0.001, 0.005, 0.01, 0.05, 0.1 lag (hr) = 25, 75, 100, 250, 500	k: 0.5101 Lag: 0.102	k: 0.774 Lag: 0.404	k: 1 Lag: 1	48%
Chick Watson with Tail	$C_t = (C_0 - C_{res})e^{-kt} + C_{res}$	k (hr^{-1}) = 0.001, 0.005, 0.01, 0.05, 0.1 C_{res} (% of C_0) = 1, 10, 25	k: 0.1029 Tail: 0.123	k: 1 Tail: 1	k: 1 Tail: 1	58%

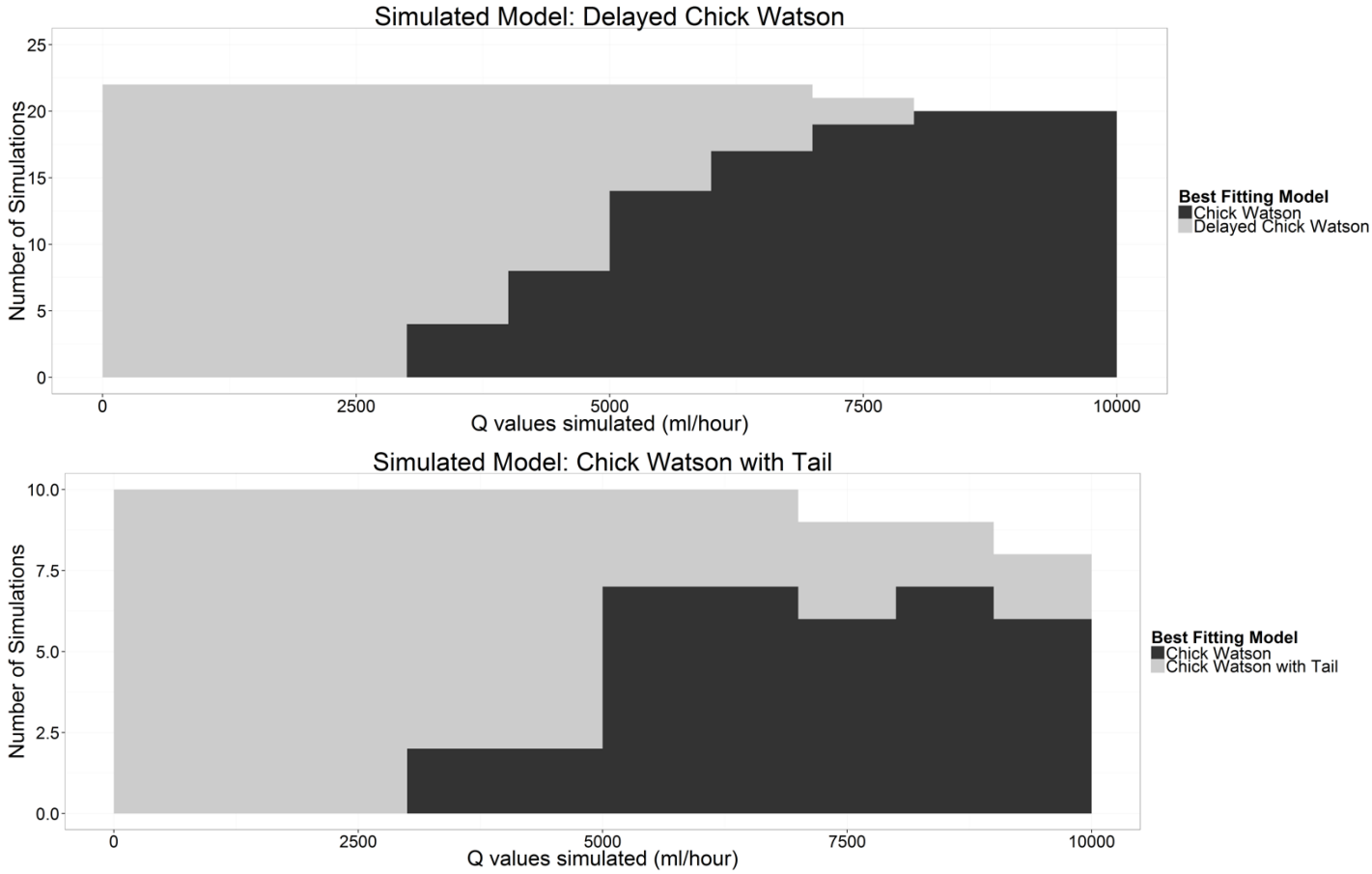

Figure 4.1 Changes in the best fitting model for non-linear simulated data with simulated flow rate (Q)

CHAPTER 5

COMPARISONS OF FECAL INDICATOR DECAY PROFILES GENERATED FROM CLOSED, PARTIALLY CLOSED, AND OPEN SYSTEMS

Lauren E. Brooks and Katharine G. Field

In Preparation for Environmental Science and Technology

Abstract

Estimates of fecal indicator bacteria decay are primarily generated in artificial systems, making it difficult to extend findings to environments of interest. While numerous studies have investigated decay and the causes of decay for fecal indicator bacteria, the potential biases introduced through the use of closed systems, such as microcosms, or partially closed systems, using dialysis bags, have not been studied. To investigate effects of these systems, we conducted simultaneous decay studies for several fecal indicators using closed, partially closed, and analogous Continuous-flow Stirred Tank Reactor systems. Significantly different decay profiles for indicators measured in each of the different systems were observed. These differences could have serious ramifications for water quality applications.

Introduction

Fecal indicators (FI) are used to monitor water quality. Traditionally, culture based techniques to grow and count FI bacteria, such as *E. coli*, are the preferred methods for enumeration, but recently genetic markers that target specific bacteria (Bernhard and Field, 2000; Chern et al., 2009, 2011; Haugland et al., 2005; Ludwig and Schleifer, 2000; Mieszkin et al., 2010; Seurinck et al., 2005; Shanks et al., 2007, 2009)Siefring et al., 2008) have received increased attention. However, uncertainty regarding indicator persistence in the environment limits their applications (Ahmed et al., 2016).

The results of past studies investigating the decay of FI have identified a number of environmental variables that impact their decay, such as the origin of fecal material (Anderson et al., 2005; Korajkic et al., 2013a; Liang et al., 2012), water matrix (Darakas et al., 2009; Green et al., 2011; Jeanneau et al., 2012; Korajkic et al., 2013b; Solecki et al., 2011), temperature (Craig et al., 2004; Kreader, 1998; Noble et al., 2004; Sokolova et al., 2012), light (Bae and Wuertz, 2009; Davies-Colley et al., 1994; Dick et al., 2010; Green et al., 2011; Korajkic et al., 2013b; Walters and Field, 2009), and microbial community (Korajkic et al., 2013b). These studies vary with respect to the shape of the decay profile and estimates of decay rate (Ahmed et al., 2010; Dick and Field, 2004; Liang et al., 2012; Noble et al., 2004; Sinton et al., 2002; Sokolova et al., 2012; Solecki et al., 2011; Tambalo et al., 2012; Walters and Field, 2009). While variability across studies is to be expected, it is likely that some inconsistency is due to the experimental decay system selection, not intended environmental stressors.

FI decay is typically studied using a closed system, in which a microcosm spiked with fecal matter is sampled over time (Ahmed et al., 2014; Anderson et al., 2005; Green et al., 2011; Jeanneau et al., 2012; Noble et al., 2004; Schulz and Childers, 2011; Sinton et al., 2002;

Sokolova et al., 2012; Solecki et al., 2011; Tambalo et al., 2012). Alternatively, to avoid the effects of a closed system, some studies have used dialysis bags attached to a submersible support frame (Bae and Wuertz, 2009, 2011, Korajkic et al., 2013a, 2014). This system is only partially closed, as it allows for some exchange between the spiked water and the surrounding water, but still restricts exchange of the microbial community, including predators, which have been shown to have significant effects on the decay of FIB (Dick et al., 2010; Korajkic et al., 2014; Rhodes and Kator, 1988; Wanjugi and Harwood, 2013).

In an open system, mass can be lost or gained from the environment, allowing water constituents to be freely exchanged between the reactor and the surrounding environment. One example of an open system is a Continuous-flow Stirred Tank Reactor (CSTR). Implementing this system for studying FI decay may help to alleviate some potential biases in closed and partially closed systems, such as continual exposure of the indigenous water microbial community to the seeded water. However, a direct comparison of decay in closed, partially closed, and open systems is not possible, as in an open system, changes in FI concentration are the result of both decay, and the removal that results from simply leaving the system. To isolate the influence of decay only, so that the different systems could be compared, we accounted for removal through mathematical adjustment.

In this study, we report the development and implementation of a novel open system reactor strategy to measure FI decomposition in ambient freshwater. Parallel decay experiments with closed and partially closed systems were conducted to identify potential similarities and differences related to the system selected to study decay. We conducted uncertainty analyses specifically using the observations for each FI to demonstrate the ranges of estimates that can be accurately detected given our experimental design. Observed differences in decay trends between

systems could have serious ramifications for water quality applications that rely on accurate indicator decay information.

Methods

Construction of different water systems

Open system replicate tanks consisted of 125 L plastic sink basins with standpipes connected to a drain to accommodate a total volume of 100 L prior to overflow. To continuously add water to the open system, we utilized pre-existing infrastructure at the Aquatic Animal Health Laboratory (AAHL; Corvallis, OR) to pump water from the Willamette River to a 9500 L holding tank. Water was pumped from the holding tank through a 2 in (5.08 cm) diameter pipe located above the tanks, to which tubing couplers were added for insertion of ½ in (1.27 cm) diameter tubing (Figure 1). Flow reducers (VWR International, Arlington Heights, IL) were added to each line to create a connection onto the 1.27 cm tubing, to which was added ¼ in (0.635 cm) inner diameter Nalgene 180 Clear Plastic Metric Tubing (Nalge Nunc International, Rochester, NY).

To regulate flow, pressure dependent button drippers were added onto the ¼ in (0.635 cm) tubes, allowing flow rates to be adjusted via the inclusion of additional tubes or through increasing or decreasing the pressure in the pipe. Two flow rates were selected; the low flow setting contained only one inflow per tank, consisting of one DIG 1gph button dripper (Model #W221B, Home Depot, Atlanta, GA). The high flow setting was made by adding a 120° y-connector (VWR International, Arlington Heights, IL) to the tube, onto which two 2gph button drippers (Model #W222B, Home Depot, Atlanta, GA) were added. A high level of pressure was maintained throughout the experiment by adding an additional discharge area that was placed

higher than the barbs used to transport water to the tanks (Figure 1), thus ensuring the pressure gradient was constant across time.

Three replicate tanks were assigned to each of the following treatments: low flow, control; low flow with spike; high flow, control; and high flow with spike. Flow rate measurements were recorded daily by measuring discharge from the standpipe for one hour for each tank. The average hourly flow rate for the high flow setup was 5331 ml/hour (SD = 122 ml/hour, range from 5178 to 5424), while the low flow setup had an average hourly flow rate of 1308 (SD = 37 ml/hour, range from 1248 to 1352).

Closed system replicates were constructed identically to the open systems, but with no water continuously added to the system. A partially-closed system was constructing using dialysis bags with a molecular weight cutoff of 14000 Molecular Weight Standards (Aldon Corp SE, Avon, NY). For each bag, 45 cm of 47.7 x 45 mm dialysis tubing was measured and soaked in DI water prior to receiving the sample. One end of the dialysis tubing was closed off by tying two knots on one end, and tightly wrapping a rubber band around the tubing between the two knots. Dialysis bags were stored in DI water until use.

Inoculation of water systems with fecal slurry

All systems were filled with river water from the adjacent Willamette River on day 0 of this experiment, approximately 2 hours before the addition of the slurry. We combined 1 L of fresh cow feces, collected from the Oregon State University Dairy Cattle Farm and transported on ice to the AAHL, with 4 L of unfiltered river water, to create a 20% V:V dilution of feces to river water in an 18.9 L bucket. The slurry was mixed by shaking, and 500 ml of the slurry was immediately added to each spiked water system tank. Control tanks with no fecal slurry were

used to monitor the input levels from the water. All tanks were continuously mixed with a Koralia Circulation and Wave Pump (Hydror, Sacramento, CA).

To fill the dialysis bags, 200 ml were subsampled from each of the recirculating tanks and added to the prepared dialysis bags. The tubing was then closed on the other end as described above, and added to additional tanks that were constructed identically to the high-flow open system tanks, but were not seeded with any fecal slurry. Three tanks containing 15 dialysis bags each were treated as replicates.

Sampling and E. coli *Enumeration*

Water samples were collected from recirculating and open-system tanks from the surface of the water near the location of the standpipe. Dialysis bags were sampled using destructive sampling in which the dialysis tubing was cut open and contents poured into a sample bottle. All samples were collected in sterile 500 ml polypropylene bottles and transported on ice for immediate processing. Samples were processed within six hours of collection.

Samples for *E. coli* enumeration were collected twice daily on days 0 through 3, then daily from days 4 until day 7 for the closed and partially closed system, or until the concentration matched the un-spiked controls for the open systems. The closed and partially closed systems were also sampled every other day from days 7 through 14.

To enumerate *E. coli*, 100 ml of sample was combined with a Colilert (Idexx Technologies, Westbrook, ME) packet in a disposable 120 ml tear-off vessel without sodium thiosulfate and mixed until the reagents were dissolved. When concentrations were expected to be above the range of quantification, serial dilutions were made from the raw samples prior to adding the Colilert packet. Mixed samples were then poured into a Quanti-Tray/2000 and heat sealed prior to incubating for 24 hours at 37°C. The number of large and small fluorescent wells

in the Quanti-Tray/2000 were counted and converted to *E. coli* MPN using the IDEXX Quanti-Tray/2000 MPN Table.

Marker Enumeration

Samples were processed for genetic marker analysis at time 0 followed by every 48 hours for 9 days, and again on day 12. To concentrate bacteria, 50 ml of each water sample was collected on a 0.4 µm disposable polycarbonate filter (Pall Corporation, Fort Washington, NY) and frozen at -80°C until DNA was extracted. DNA extractions were performed following the manufacturer's protocols using the Gene-Rite DNA extraction kit with modifications and processing controls implemented as recommended previously (Haugland et al., 2005). Extracted DNA was stored at -20°C for approximately one month until used for qPCR amplification.

We quantified genetic markers from two indicators of general fecal contamination (Table 5.1): the EC23S857 marker for *E. coli* (Chern et al., 2011), and the *Bacteroidales* associated marker (GenBac3; (Dick and Field, 2004; Siefring et al., 2008) combined with an Internal Amplification Control (Haugland et al., 2010). Additionally, we quantified the presence of the Rum2Bac marker (Mieszkin et al., 2010), as an indicator of ruminant fecal contamination (Raith et al., 2013). Details for qPCR runs and quality assurance and quality control measures are provided in supplemental materials (Appendix 1).

Additional measurements

Temperature, dissolved oxygen, and pH were monitored throughout the experiment using an YSI ProPlus Multiparameter meter (Yellow Springs International, Yellow Springs, OH) and revealed no significant differences ($p>0.05$) in water parameters between systems (mean temperature = 11°C, mean dissolved oxygen = 9.994 mg/L). All tanks were exposed to ambient

light within the AAHL, which is covered, but lined with windows, resulting in some light exposure during the experiment.

Control tanks were analyzed for both the high flow and low flow open system to monitor for influxes of FI from sources other than the fecal slurry, such as elevated FI concentrations in the river water. FI levels in the control tanks were negligible relative to the spike levels throughout the experiment (Supplemental Materials).

Uncertainty analysis

We used simulations based on the observed concentrations and variability from our study to determine the range of parameters that could be accurately estimated given our study design. Simulations are described in depth in supplemental materials (Appendix 2). Briefly, using initial concentrations, sampling times, and variability observed specifically for the indicators used for this study, we generated confidence intervals for each estimate from the simulations, assuming a t-distribution with the degrees of freedom observed for each simulation. Estimates were classified as inaccurate when the simulated value fell outside the confidence interval generated from the model parameter estimate mean and standard error.

Open system concentration adjustment and statistical comparison

To generate comparable decay profiles between the open and closed systems, we adjusted the observed concentration data for each target from the open systems. To adjust observed concentration data (C_t), we divide the concentration observed by the concentration removed, as calculated using assumptions for a CSTR, resulting in:

Eq 1: $C_{Adjusted} = C_{Observed}/\exp(-k_f*time)$

where k_f is calculated as Flow Rate/Volume using average flow rates for each tank (Figure 5.3).

Following this adjustment, we made statistical comparisons between the adjusted decay profiles. Decay models were fit to the survival ratio ($St = C/C_0$) for each system using individual data points generated. Decay models fit to the data included the Chick Watson model (CW), Delayed Chick Watson (DCW), and Delayed Chick Watson with Tail (DCWT; Equations S1-S3). Best fitting models were assigned based on the model that had the lowest Akaike Information Criterion (AIC) value (Table S5.2), which incorporates a penalty for the inclusion of additional parameters to be estimated, thus avoiding using unnecessarily complex models. Goodness of fit was assessed visually using residual errors for each best fitting model (data not shown).

Statistical comparisons were made between the decay profiles for each FI across all systems using Global Model Fitting, as recommended for comparing nonlinear curves (Motulsky and Christopoulos, 2004). Briefly, data generated from different treatments were combined into one dataset, and the null hypothesis that the two datasets should be combined was compared to the alternative hypothesis that there was enough difference in the datasets to support separation into two different models. For all comparisons, data were cut short based on the duration of the shortest dataset used in the comparison. Using the sum of squared errors and the degrees of freedom for data fit separately or as one shared model, we generated the F-ratio and used it to determine the P-value for each comparison. To determine which features differed between decay profiles generated for different systems, we used parameter means and standard errors to compute a t-statistic, which was then used to generate P-values for each comparison.

Results

Uncertainty analysis

An analysis of uncertainty for our estimates revealed that the accuracy of model parameter estimates and the ability of the models to fit the data varied depending on the simulated value, even when no flow was simulated (Figure 5.2). Simulations with high decay rates did not contain enough data points to fit the models, resulting in no model estimates for many simulations. Simulations based on the sampling and initial concentrations observed for cultivable *E. coli* had the fewest errors, likely due to the increased frequency of sampling, but the DCW model still fit the data poorly under the high flow scenario. No simulations for the EC23S857 contained enough data to be fit by models for the high flow system. Errors were generally lower for simulations conducted using the initial concentration of the two *Bacteroidales* associated markers (GenBac3 and Rum2Bac), but were still detected and were more prevalent for the open systems. However, high flow simulations resulted in more inaccurate estimates or an inability to fit the data, as the shortened datasets resulted in an inability to fit the DCW model to the data.

FI decay in different water systems

Differences in duration of the experiment were the result of different starting concentrations (C_0) (Table 5.1), system, and the limits of detection for each FI (Table S5.1). Data were unavailable at later time points from the open systems, and in the case of the EC23S857 marker, not enough data were available for analyses from the high flow system.

Decay profiles generally consisted of a combination of a lag period, followed by exponential decay, followed in some profiles by a tailing feature (Figure 5.4). For datasets generated from the high flow systems, no lag periods were detected, while lag periods were detected for all other indicators in all other systems, with the exception of the EC23S857 marker in the closed system (Table 5.1). No tailing features were detected from any indicator in either of

the open systems, but tailing features were detected for most FI in the closed and partially closed systems.

Statistical comparisons for all designs except for the high flow open system indicated highly significant differences ($p < 0.00001$) in decay profiles for cultivable *E. coli*. Statistical comparisons made between model parameters using the t-test revealed that there was no significant difference in the decay rates of the closed, partially closed, and low flow systems ($p>0.05$)(Figure 5.5). Duration of the lag period was significantly different between the closed and partially closed systems ($p = < 0.01$), but not significant between the low flow system and either the closed or partially closed systems ($p > .05$). For both the closed and partially closed systems, a tailing effect was detected and found to be significantly different ($p = < 0.00001$), with the dialysis treatment having a much higher residual concentration (10% of C_0 for the partially closed system, <1% for the closed system) at the end of the experiment on day 14.

Significant differences were detected in the decay profiles of EC23S857 between all systems (Table 5.2). The only feature detected for this marker in the closed system was the exponential decay. No significant differences were detected in the estimated decay rates between the closed and partially closed ($p = 0.30$) or low flow ($p = 0.08$) systems, while a difference was observed between the partially closed and low flow system ($p = 0.029$; Figure 5.5). Lag periods were detected for both the partially closed and low flow systems, but did not differ significantly from one another ($p = 0.28$). The decay profile for EC23S857 from the partially closed dataset was the only profile that exhibited a tail, with nearly 10% of the initial concentration remaining at the end of the experiment (Table 5.1).

A comparison of decay profiles generated for the GenBac3 marker revealed no differences between the high flow system and the closed systems, but a difference was detected

between low flow and all other systems. For the low flow system, the decay rate estimate was significantly higher than estimates from the closed and partially closed systems, while it was significantly lower than the high flow system (Figure 5.5; Tables S5.3 and S5.4). A lag period was detected for the low flow system, even when shortened to the duration of the high flow experiment. Additionally, the observed lag period was significantly longer for the low flow system than the closed system ($p = 0.012$), while it did not differ from the partially closed system ($p = 0.92$). Comparisons between the closed and partially closed systems revealed no significant differences between the decay rate or lag duration, but did indicate a significant difference in the residual concentration ($p = 0.031$).

For the Rum2Bac marker, differences were detected for all comparisons made across systems. No lag period was observed for the high flow system, while lag periods were observed for each of the other systems even when cut off to the duration of the high flow system (Table S5.4). No significant difference was observed between the lag period for the low flow system and either the closed or partially closed systems ($p>0.05$). However, there was a difference between both the lag estimates ($p = 0.0074$) for the closed and partially closed systems. The closed and partially closed systems, however, did not experience a significant difference in decay rates ($p = 0.49$) or residual concentrations ($p = 0.094$).

Discussion

Paired experiments measuring the decay of select cultivated and genetic FI demonstrated that different experiment systems used to study decay can lead to significantly different results. Causes of FIB decay are often attributed to environmental variables like temperature and light (Craig et al., 2004; Dick et al., 2010; Noble et al., 2004; Walters and Field, 2009; Whitman et al.,

2004). In this experiment, the only variable tested was system type, and all other known covariates were equivalent for all experiments. Thus, differences between the data can be attributed to the effects of the closed, partially closed, or open systems.

The number of inaccurate estimates parameter combinations from our uncertainty analysis revealed the difficulties associated with making estimates of model parameters, regardless of the system choice. The accuracy of estimates varied with the value of the model parameter, as well as indicator specific settings such as starting concentration (C_0). Typically, when the Chick Watson model was successfully fit to the data, it did a good job of estimating the simulated value, but many of the simulations failed to be fit due to either high variability or a poor fit, even when trying to fit the same model that had been simulated. Experimental design features such as sampling frequency also impacted the accuracy of model parameter estimates, highlighting the need to assess uncertainty when conducting decay studies or comparing reported estimates from different studies. We proceeded with estimates generated from best fitting models, but used caution in comparisons that rely on precise estimates for our parameters.

At the beginning of the experiment, prior to exponential decay, the survival ratios from the FI were similar across systems, suggesting the adjustment term was a reliable way to correct for removal from the system. Thus, we feel confident that the adjustment was successful and the discrepancies between the curves are the result of biological differences rather than a failure of the adjustment term. However, one disadvantage to implementing the open system experimental design is the shortening of the experiment, as due to marker removal, marker concentrations drop below the lower limit of quantitation earlier than in closed systems, regardless of the shape of decay profiles. Thus, estimates from the open systems can only be generated for the beginning of the decay profile, and cannot be extended past the last data point.

The log-linear relationship between concentration and time is often modeled for decay studies (Ahmed et al., 2014; Bell et al., 2009; Jeanneau et al., 2012; Noble et al., 2004; Schulz and Childers, 2011; Sinton et al., 2007, 2002; Sokolova et al., 2012; Walters and Field, 2009). While additional features were detected for most decay profiles, some portion of nearly all decay profiles were found to undergo exponential decay that was modeled using a log-linear relationship for either a portion of the experiment or a portion of the initial concentration.

The partially closed system allowed for the removal of small materials, while constraining the microbial community, and experienced no difference in decay rate for any marker relative to the closed system. This suggests that the cause or causes of decay were not being removed from the system in the partially closed system, and were thus larger than the dialysis bag molecular weight cutoff.

As decay rates observed for the high flow systems were low, and for *E. coli*, not significantly different from zero, it appeared that the indicators in the high flow systems did not enter the exponential decay phase, and were instead only observed in a period of little to no decay. Additionally, lag periods tended to be longer for low flow and partially closed systems relative to the closed system, suggesting that the causes of exponential decay took longer to take effect as the system became more open. Both these observations suggest that diluting cells prevents or delays their decay; observed decay was thus density dependent. The presence of a lag period before the exponential decay phase has been observed previously in experimental designs using the partially closed system deployed into freshwater (Bae and Wuertz, 2011; Korajkic et al., 2014). While possible explanations for the lag period have been proposed to be a buildup of predators (Green et al., 2011) or the breakdown of a critical component key to the survival of the indicator (Geeraerd et al., 2000), no cause has been demonstrated to directly cause the lag period.

Tails have previously been detected using a closed system (Jeanneau et al., 2012; Solecki et al., 2011) and have been hypothesized to be the result of a resistant sub population that fails to succumb to the pressures that lead to exponential decay (Cerf, 1977). However, it is also conceivable that the tail could result from a decreased potency in the cause(s) of decay over time, leading to decreased decay again at the later hours of the experiment. While no tailing effects were detected in the open systems, likely as a result of removal even if a tail were present, significantly different tailing effects were observed for the closed and partially closed systems. Higher residual concentrations were found for the partially closed systems, suggesting that the exchange of small materials provided an environment in which indicators were not being removed from the system.

Our results suggest that the less closed the system, the longer the persistence of the indicators tested in this study. This result implies that since the environment is an open system, we may be overestimating how recent a contamination event is, if we use decay rates derived from closed systems experiments. This study focused on fecal indicators, and did not measure the decay of any associated fecal pathogens. However, it is possible that the same extended persistence could occur for pathogens. Future research is warranted to characterize the influence of open system conditions on public health relevant pathogens.

Although the open system approach described here clearly demonstrates how experimental design can influence decay findings, it is important to recognize limitations in this novel approach. Uncertainty analysis indicated a shortened duration of measurable concentrations and a loss of indicator measurement sensitivity, likely due to the additional variability in an open system. Despite these limitations, using an open system to study decay has several advantages over traditional setups. The exchange of all water constituents has the

possibility to influence estimates of decay, as was reported in this study. Here we used flow rates at both ends of the feasible ranges, resulting in different decay profiles for each flow rate. The flow rates set in this experiment were determined based on the results of the simulations, rather than with any environmental considerations.

The results from this study provide a good starting point to further investigate differences between estimates of decay profiles generated from an artificial setting and what is likely to occur in the environments of interest. Our findings demonstrate that the system used to study decay can greatly influence results, but do not fully model the environments of interest, as a CSTR differs greatly from natural water bodies. In future studies, flow rates could be simulated based on observed hydrological phenomena, and used to more accurately estimate decay, thus better simulating the environment. Additionally, this type of modelling could be adjusted from the perfect mixing assumption to incorporate other hydrological phenomena, such as dispersion or settling, thus allowing decay studies to be conducted in the environments of interest.

References

Ahmed, W., Yusuf, R., Hasan, I., Goonetilleke, A., and Gardner, T. (2010). Quantitative PCR assay of sewage-associated *Bacteroides* markers to assess sewage pollution in an urban lake in Dhaka, Bangladesh. Can. J. Microbiol. *56*, 838–845.

Ahmed, W., Gyawali, P., Sidhu, J. p. s., and Toze, S. (2014). Relative inactivation of faecal indicator bacteria and sewage markers in freshwater and seawater microcosms. Lett. Appl. Microbiol. 348–354.

Ahmed, W., Hughes, B., and Harwood, V.J. (2016). Current status of marker genes of *Bacteroides* and related taxa for identifying sewage pollution in environmental waters. Water *8*, 231.

Anderson, K.L., Whitlock, J.E., and Harwood, V.J. (2005). Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Appl. Environ. Microbiol. *71*, 3041–3048.

Bae, S., and Wuertz, S. (2009). Discrimination of viable and dead fecal *Bacteroidales* bacteria by quantitative PCR with propidium monoazide. Appl. Environ. Microbiol. *75*, 2940–2944.

Bae, S., and Wuertz, S. (2011). Survival of host-associated *Bacteroidales* cells and their relationship with *Enterococcus spp.*, *Campylobacter jejuni*, *Salmonella enterica*, *Serovar Typhimurium*, and Adenovirus in freshwater microcosms as measured by propidium monoazide-Quantitative PCR. Appl. Environ. Microbiol. *78*, 922–932.

Bell, A., Layton, A.C., McKay, L., Williams, D., Gentry, R., and Sayler, G.S. (2009). Factors influencing the persistence of fecal *Bacteroides* in stream water. J. Environ. Qual. *38*, 1224.

Bernhard, A.E., and Field, K.G. (2000). A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Appl. Environ. Microbiol. *66*, 4571–4574.

Cerf, O. (1977). A REVIEW Tailing of Survival Curves of Bacterial Spores. J. Appl. Bacteriol. *42*, 1–19.

Chern, E.C., Brenner, K.P., Wymer, L., and Haugland, R.A. (2009). Comparison of fecal indicator bacteria densities in marine recreational waters by QPCR. Water Qual. Expo. Health *1*, 203–214.

Chern, E.C., Siefring, S., Paar, J., Doolittle, M., and Haugland, R.A. (2011). Comparison of quantitative PCR assays for *Escherichia coli* targeting ribosomal RNA and single copy genes. Lett. Appl. Microbiol. *52*, 298–306.

Craig, D. l., Fallowfield, H. j., and Cromar, N. j. (2004). Use of microcosms to determine persistence of *Escherichia coli* in recreational coastal water and sediment and validation with *in situ* measurements. J. Appl. Microbiol. *96*, 922–930.

Darakas, E., Koumoulidou, T., and Lazaridou, D. (2009). Fecal indicator bacteria declines via a dilution of wastewater in seawater. Desalination *248*, 1008–1015.

Davies-Colley, R.J., Bell, R.G., and Donnison, A.M. (1994). Sunlight inactivation of Enterococci and fecal coliforms in sewage effluent diluted in seawater. Appl. Environ. Microbiol. *60*, 2049–2058.

Dick, L.K., and Field, K.G. (2004). Rapid estimation of numbers of fecal *Bacteroidetes* by use of a quantitative PCR assay for 16S rRNA genes. Appl. Environ. Microbiol. *70*, 5695–5697.

Dick, L.K., Stelzer, E.A., Bertke, E.E., Fong, D.L., and Stoeckel, D.M. (2010). Relative decay of *Bacteroidales* microbial source tracking markers and cultivated *Escherichia coli* in freshwater microcosms. Appl. Environ. Microbiol. *76*, 3255–3262.

Geeraerd, A.H., Herremans, C.H., and Van Impe, J.F. (2000). Structural model requirements to describe microbial inactivation during a mild heat treatment. Int. J. Food Microbiol. *59*, 185–209.

Green, H.C., Shanks, O.C., Sivaganesan, M., Haugland, R.A., and Field, K.G. (2011). Differential decay of human faecal *Bacteroides* in marine and freshwater. Environ. Microbiol. *13*, 3235–3249.

Haugland, R.A., Siefring, S.C., Wymer, L.J., Brenner, K.P., and Dufour, A.P. (2005). Comparison of *Enterococcus* measurements in freshwater at two recreational beaches by quantitative polymerase chain reaction and membrane filter culture analysis. Water Res. *39*, 559–568.

Haugland, R.A., Varma, M., Sivaganesan, M., Kelty, C., Peed, L., and Shanks, O.C. (2010). Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected *Bacteroidales* species and human fecal waste by qPCR. Syst. Appl. Microbiol. *33*, 348–357.

Jeanneau, L., Solecki, O., Wéry, N., Jardé, E., Gourmelon, M., Communal, P.-Y., Jadas-Hécart, A., Caprais, M.-P., Gruau, G., and Pourcher, A.-M. (2012). Relative decay of fecal indicator bacteria and human-associated markers: A microcosm study simulating wastewater input into seawater and freshwater. Environ. Sci. Technol. *46*, 2375–2382.

Korajkic, A., McMinn, B.R., Harwood, V.J., Shanks, O.C., Fout, G.S., and Ashbolt, N.J. (2013a). Differential decay of enterococci and *Escherichia coli* originating from two fecal pollution sources. Appl. Environ. Microbiol. *79*, 2488–2492.

Korajkic, A., Wanjugi, P., and Harwood, V.J. (2013b). Indigenous microbiota and habitat influence *Escherichia coli* survival more than sunlight in simulated aquatic environments. Appl. Environ. Microbiol. *79*, 5329–5337.

Korajkic, A., McMinn, B.R., Shanks, O.C., Sivaganesan, M., Fout, G.S., and Ashbolt, N.J. (2014). Biotic interactions and sunlight affect persistence of fecal indicator bacteria and

microbial source tracking genetic markers in the upper Mississippi River. Appl. Environ. Microbiol. *80*, 3952–3961.

Kreader, C.A. (1998). Persistence of PCR-Detectable *Bacteroides distasonis* from human feces in river water. Appl. Environ. Microbiol. *64*, 4103–4105.

Liang, Z., He, Z., Zhou, X., Powell, C.A., Yang, Y., Roberts, M.G., and Stoffella, P.J. (2012). High diversity and differential persistence of fecal *Bacteroidales* population spiked into freshwater microcosm. Water Res. *46*, 247–257.

Ludwig, W., and Schleifer, K.-H. (2000). How quantitative is quantitative PCR with respect to cell counts? Syst. Appl. Microbiol. *23*, 556–562.

Mieszkin, S., Yala, J.-F., Joubrel, R., and Gourmelon, M. (2010). Phylogenetic analysis of *Bacteroidales* 16S rRNA gene sequences from human and animal effluents and assessment of ruminant faecal pollution by real-time PCR. J. Appl. Microbiol. *108*, 974–984.

Motulsky, H., and Christopoulos, A. (2004). Fitting models to biological data using linear and nonlinear regression: A practical guide to curve fitting (Oxford University Press, USA).

Noble, R.T., Lee, I.M., and Schiff, K. c. (2004). Inactivation of indicator micro-organisms from various sources of faecal contamination in seawater and freshwater. J. Appl. Microbiol. *96*, 464–472.

Raith, M.R., Kelty, C.A., Griffith, J.F., Schriewer, A., Wuertz, S., Mieszkin, S., Gourmelon, M., Reischer, G.H., Farnleitner, A.H., Ervin, J.S., et al. (2013). Comparison of PCR and quantitative real-time PCR methods for the characterization of ruminant and cattle fecal pollution sources. Water Res. *47*, 6921–6928.

Rhodes, M.W., and Kator, H. (1988). Survival of *Escherichia coli* and *Salmonella spp.* in estuarine environments. Appl. Environ. Microbiol. *54*, 2902–2907.

Schulz, C.J., and Childers, G.W. (2011). Fecal *Bacteroidales* diversity and decay in response to variations in temperature and salinity. Appl. Environ. Microbiol. *77*, 2563–2572.

Seurinck, S., Defoirdt, T., Verstraete, W., and Siciliano, S.D. (2005). Detection and quantification of the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. Environ. Microbiol. *7*, 249–259.

Shanks, O.C., Atikovic, E., Blackwood, A.D., Lu, J., Noble, R.T., Domingo, J.S., Seifring, S., Sivaganesan, M., and Haugland, R.A. (2007). Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. Appl. Environ. Microbiol. *74*, 745–752.

Shanks, O.C., Kelty, C.A., Sivaganesan, M., Varma, M., and Haugland, R.A. (2009). Quantitative PCR for genetic markers of human fecal pollution. Appl. Environ. Microbiol. *75*, 5507–5513.

Siefring, S., Varma, M., Atikovic, E., Wymer, L., and Haugland, R.A. (2008). Improved real-time PCR assays for the detection of fecal indicator bacteria in surface waters with different instrument and reagent systems. J. Water Health *6*, 225–237.

Sinton, L., Hall, C., and Braithwaite, R. (2007). Sunlight inactivation of *Campylobacter jejuni* and *Salmonella enterica*, compared with *Escherichia coli*, in seawater and river water. J. Water Health *5*, 357–365.

Sinton, L.W., Hall, C.H., Lynch, P.A., and Davies-Colley, R.J. (2002). Sunlight inactivation of fecal indicator bacteria and bacteriophages from waste stabilization pond effluent in fresh and saline waters. Appl. Environ. Microbiol. *68*, 1122–1131.

Sokolova, E., Åström, J., Pettersson, T.J.R., Bergstedt, O., and Hermansson, M. (2012). Decay of *Bacteroidales* genetic markers in relation to traditional fecal indicators for water quality modeling of drinking water sources. Environ. Sci. Technol. *46*, 892–900.

Solecki, O., Jeanneau, L., Jardé, E., Gourmelon, M., Marin, C., and Pourcher, A.M. (2011). Persistence of microbial and chemical pig manure markers as compared to faecal indicator bacteria survival in freshwater and seawater microcosms. Water Res. *45*, 4623–4633.

Tambalo, D.D., Fremaux, B., Boa, T., and Yost, C.K. (2012). Persistence of host-associated *Bacteroidales* gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. Water Res. *46*, 2891–2904.

Walters, S.P., and Field, K.G. (2009). Survival and persistence of human and ruminant-specific faecal *Bacteroidales* in freshwater microcosms. Environ. Microbiol. *11*, 1410–1421.

Wanjugi, P., and Harwood, V.J. (2013). The influence of predation and competition on the survival of commensal and pathogenic fecal bacteria in aquatic habitats. Environ. Microbiol. *15*, 517–526.

Whitman, R.L., Nevers, M.B., Korinek, G.C., and Byappanahalli, M.N. (2004). Solar and temporal effects on *Escherichia coli* concentration at a Lake Michigan swimming beach. Appl. Environ. Microbiol. *70*, 4276–4285.

Figures and Tables:

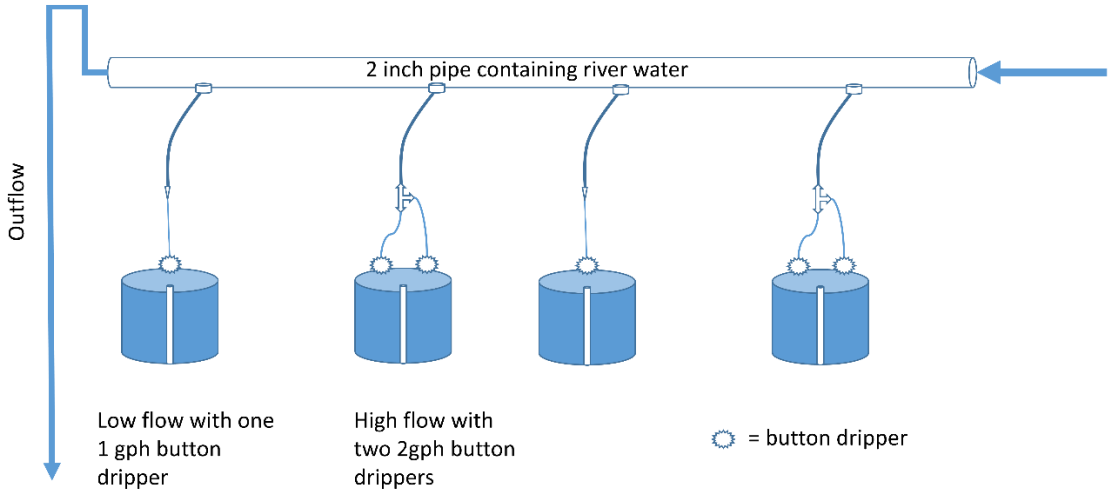

Figure 5.1 Schematic diagram of the low-flow and high-flow CTSRs. Water from the Willamette River passed from a holding tank through a series of connectors into the tanks, through either a single 1 gph button dripper (low flow tanks), or two 2 gph drippers (high flow tanks). A standpipe in each tank kept the volume within the tanks constant at 100L.

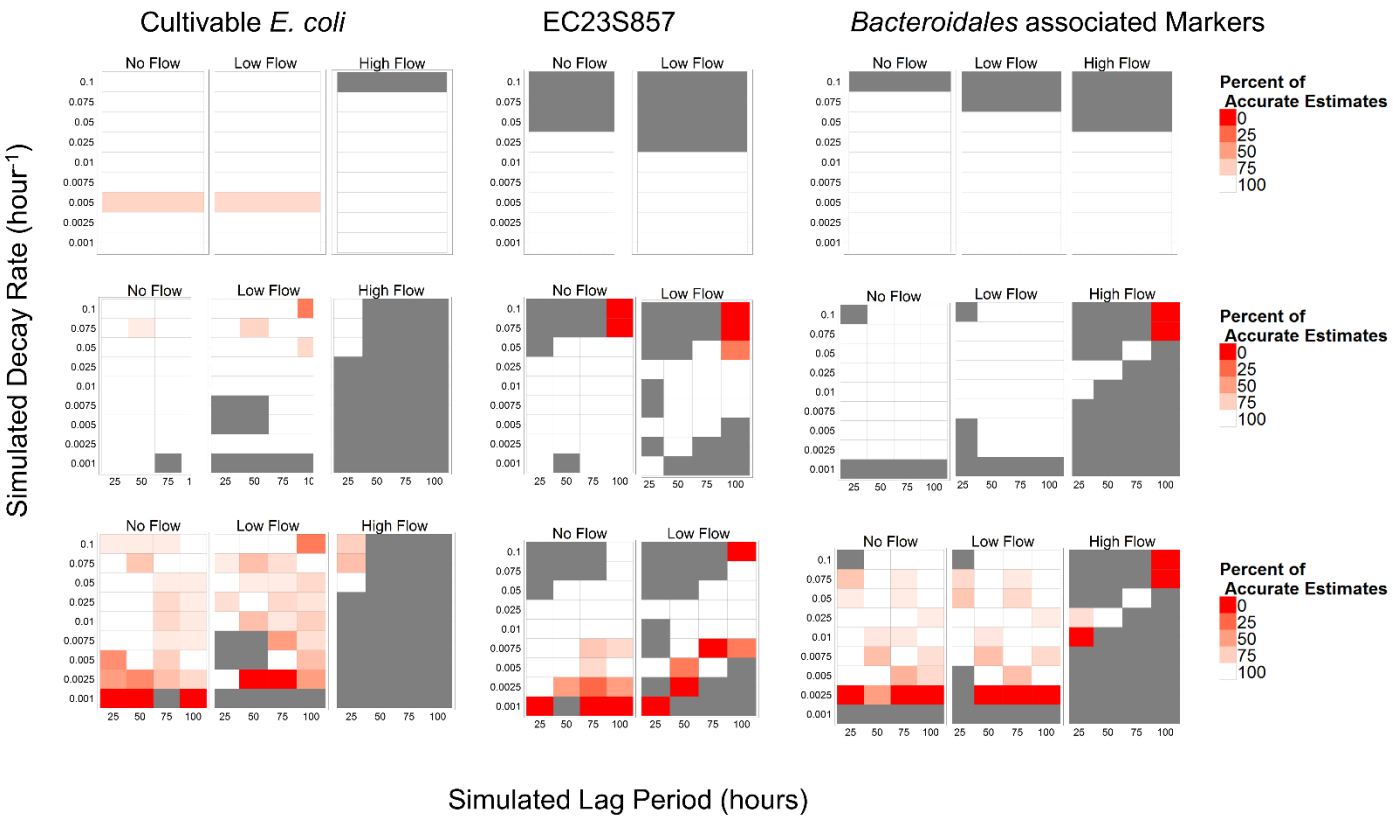

Figure 5.2 Counts of inaccurate estimates of model parameters detected when simulating combinations of model parameters from the Chick Watson (decay rate estimate), Delayed Chick Watson decay rate estimate, and lag period estimate. Estimates were classified as inaccurate when the simulated value was not within the confidence intervals for the estimates. Y axis represents simulated decay rates, x axis represents simulated lag durations (0 for Chick Watson Model). Gray blocks mean no values could be calculated for that combination of parameters.

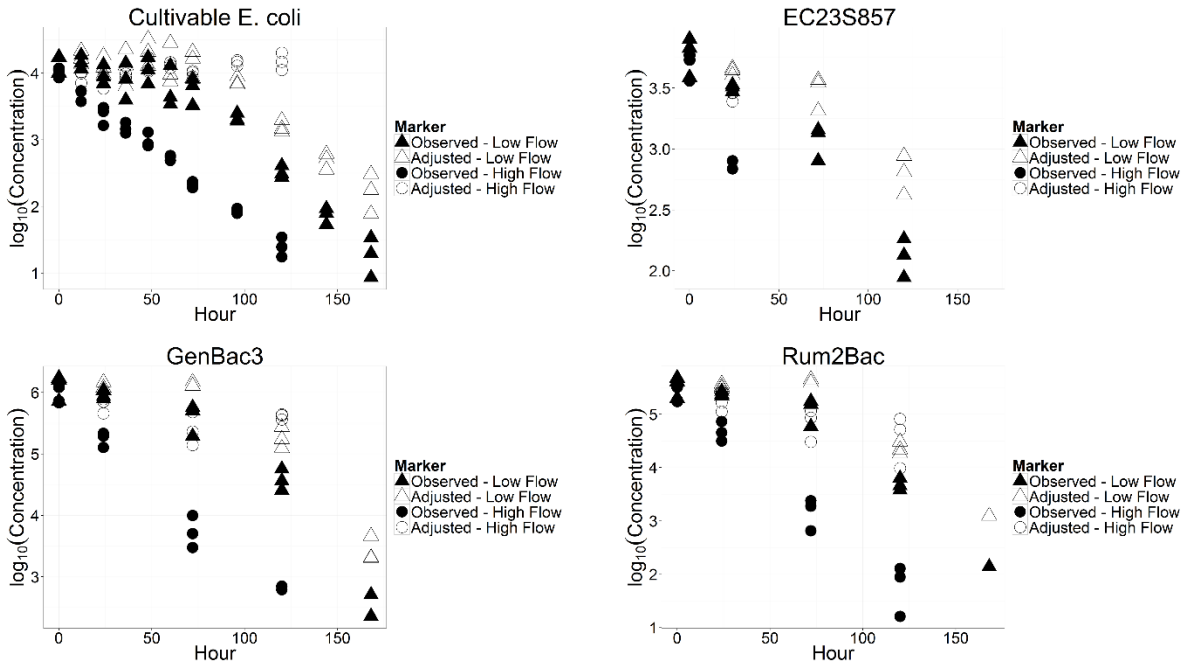

Figure 5.3 Concentrations per 100 ml measured from the open systems before (closed shapes) and after (open shapes) the adjustment to the data.

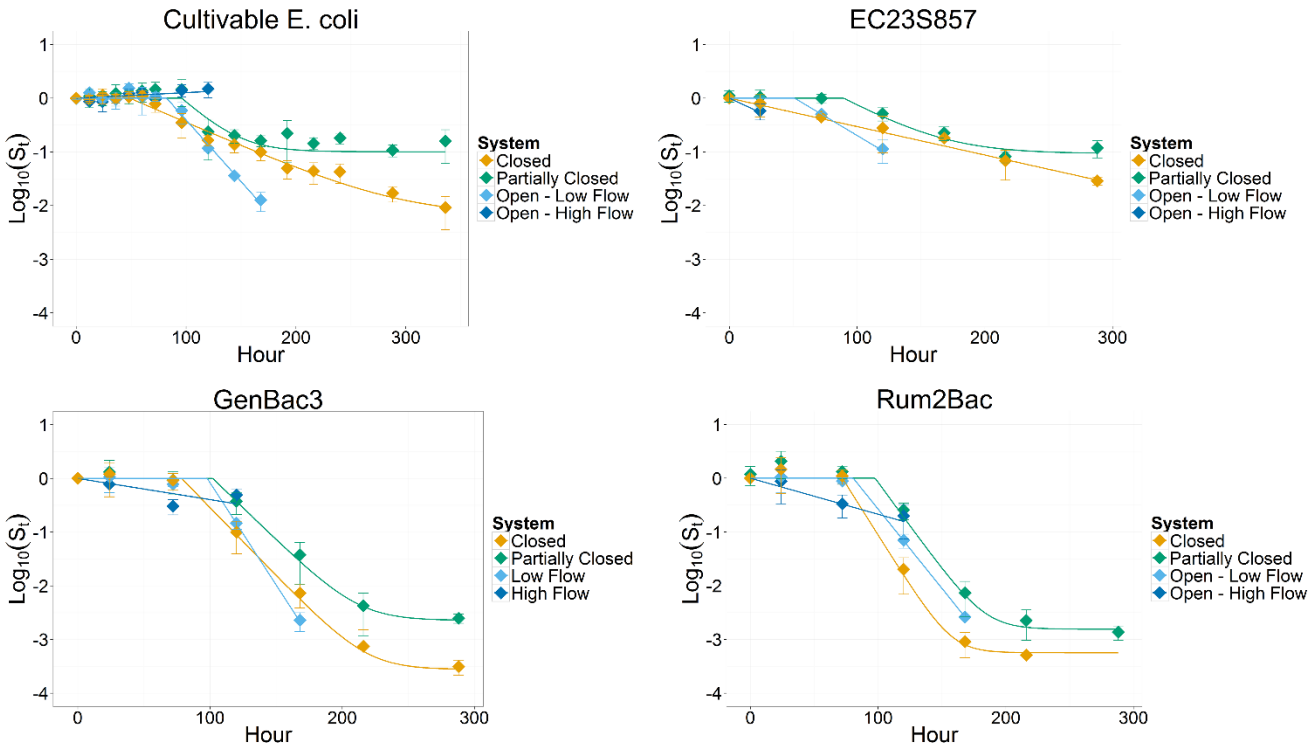

Figure 5.4 Different decay profiles generated from adjusted concentrations measured from each of the four systems as estimated using the best fitting models. Error bars on mean values represent standard deviation estimates for the three replicates. Decay rate estimates were calculated using natural log, but lines displayed here were adjusted and are displayed here on the $\log_{10}$ scale for ease of interpretation.

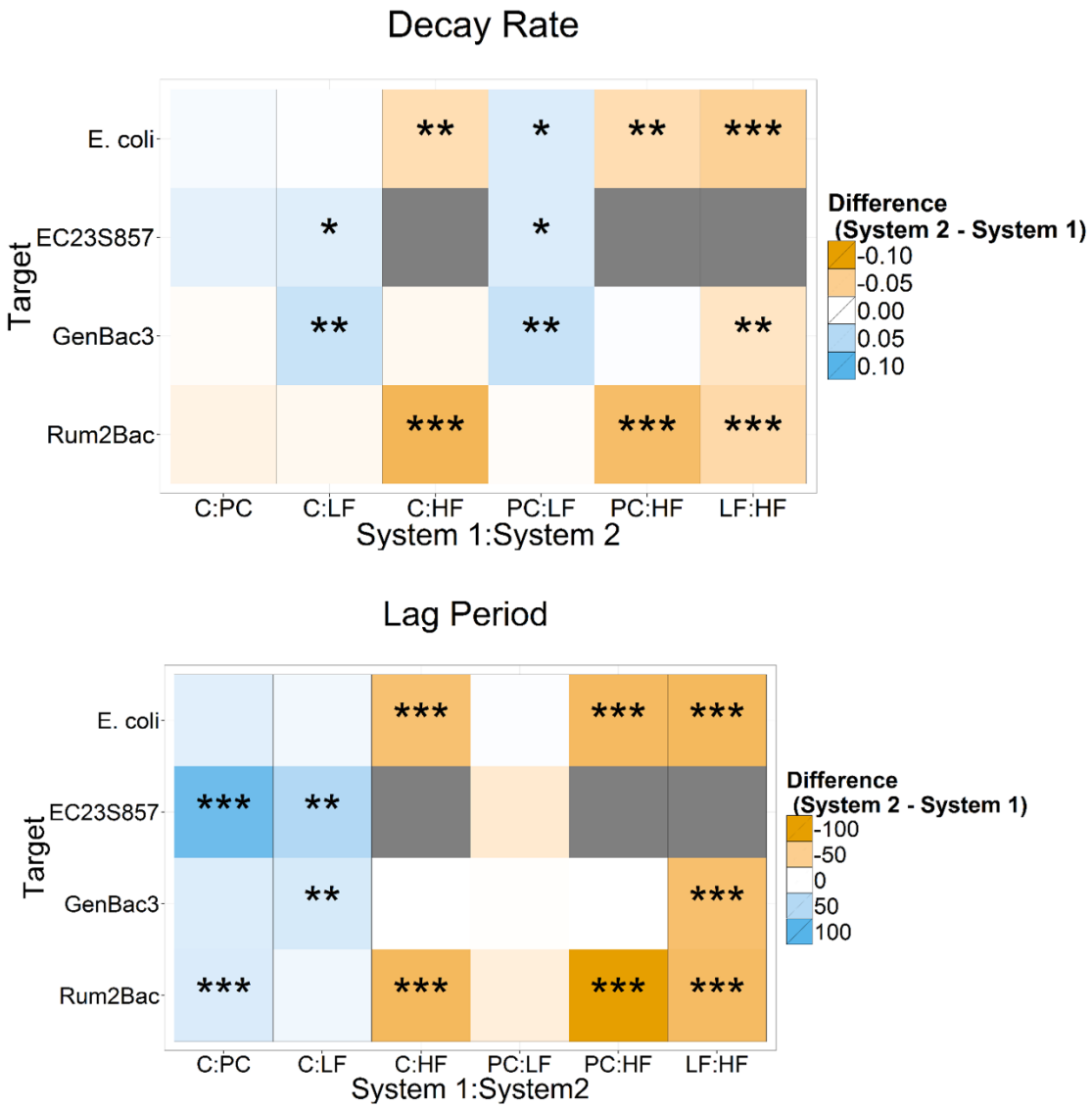

Figure 5.5 Summary of differences observed for all targets across all systems. Stars indicate significance levels (* indicates significance at $\alpha = 0.1$, ** indicates significance at $\alpha = 0.05$, and *** indicates significance at $\alpha < 0.01$)

Table 5.1 Summary results for the best fitting models for each system. Values in parentheses represent standard error of the estimates.

Indicator	System	Best Fitting Model	Log 10 Mean C0 / 100 ml	Lag (Hours)	Decay Rate (Hour^{-1})	Tail (% of C0)	Duration (hours)
Cultivable *E. coli*	Closed	Delayed Chick Watson with Tail	4.16 (3.56)	48.68 (13.99)	.0204 (.0033)	.0066 (.0052)	336
	Partially Closed	Delayed Chick Watson with Tail	4.16 (3.56)	96.0 (.022)	.0444 (.018)	.10 (.00048)	336
	Open – Low Flow	Delayed Chick Watson	4.10 (3.63)	82.28 (8.14)	.0528 (.0076)	--	168
	Open – High Flow	Chick Watson	4.02 (3.22)	1.9e-9 (1.93 e-6)	--	--	120
EC23S857	Closed	Chick Watson	5.54 (4.99)	0.012 (0.00097)	--	--	288
	Partially Closed	Delayed Chick Watson with Tail	5.54 (4.99)	0.030 (0.017)	28 (0.092)	9.2 (3.5)	288
	Open – Low Flow	Delayed Chick Watson	5.49 (5.02)	0.033 (0.012)	51 (19)	--	120
	Open – High Flow	NA	5.39 (4.77)	--	--	--	24
GenBac3	Closed	Delayed Chick Watson with Tail	7.87 (7.30)	0.058 (0.0084)	79 (12)	0.02 (0.01)	288
	Partially Closed	Delayed Chick Watson with Tail	7.87 (7.30)	0.055 (0.093)	102 (9.4)	0.23 (0.088)	288
	Low Flow	Delayed Chick Watson	7.83 (7.43)	0.88 (0.012)	98 (7.1)	--	168
	High Flow	Chick Watson	7.64 (7.17)	0.0091 (0.0029)	--	--	120
Rum2Bac	Closed	Delayed Chick Watson with Tail	7.25 (6.80)	0.087 (0.0087)	72 (0.15)	0.057 (0.029)	288
	Partially Closed	Delayed Chick Watson with Tail	7.25 (6.80)	0.075 (0.015)	98 (9.0)	0.16 (0.049)	288
	Open – Low Flow	Delayed Chick Watson	7.26 (6.87)	0.068 (0.017)	81 (14)	--	168
	Open – High Flow	Chick Watson	7.05 (6.63)	0.015 (0.0029)	--	--	120

Table 5.2 Summary for P-values comparing curves for each combination of systems.

System 1	**System 2**	**Cultivable *E. coli***	**EC23S857**	**GenBac3**	**Rum2Bac**
Closed	Partially Closed	< 0.000001	0.00010	< 0.000001	< 0.000001
Closed	Open – Low Flow	< 0.000001	0.012	0.0072	0.0076
Closed	Open – High Flow	< 0.000001	0.39	0.31	0.000053
Partially Closed	Open – Low Flow	< 0.000001	0.000012	< 0.000001	0.0050
Partially Closed	Open – High Flow	< 0.000001	0.094	0.60	0.014
Open – Low Flow	Open – High Flow	< 0.000001	0.37	0.0038	0.00029

Supplemental Materials:

Appendix 1: qPCR Methods

Quantitative PCR was conducted using an ABI 7500 Fast Real-Time PCR system (Thermo Fisher Scientific, Carlsbad, CA). Individual plates were run for all samples for the first 5 days, and one plate containing the remaining samples was analyzed for days 7, 9 and 12, resulting in 5 runs per marker. A single plasmid was constructed to contain sequences for all targeted sequences. Serial dilutions of the standard were used to construct calibration curves for each run on the instrument. For each assay, a 20 µl reaction consisted of 10 µl TaqMan Fast Universal Master Mix, 500 nmol l^{-1} of each primer, and 250 nmol l^{-1} of probe. All reactions were performed in triplicate in MicroAmp optical 96-well plates with optical adhesive film (Thermo Fisher Scientific, Carlsbad, CA). Cycling parameters for all assays included a 2 minute start at 94°C followed by 40 cycles of 15s at 94°C and 32s at 60°C. Cycle threshold for each run was determined by the software provided with the ABI 7500 Fast Real-Time PCR system and used to calculate the concentration using equations fit from the standard curves (Table S5.1).

Quality controls were implemented across several steps of the sample processing. Unspiked control tanks monitored for the addition of genetic markers from water added to the open system revealed negligible inputs for Rum2Bac and EC23S857, and a consistently detectable presence of GenBac3 (mean = 25000 markers/100 ml). Extraction blanks were created and analyzed to detect contamination as well as a baseline for the sample processing control. A sample processing control (Haugland et al., 2005) added to the extraction buffer was measured for all samples. Cycle thresholds from extraction blanks were used to create a range equal to the mean +/- 3*standard deviation for extraction blanks for each plate of acceptable cycle thresholds for samples. Cycle thresholds for all samples were within this range. Internal amplification

controls (AIC) were used to measure the effects of inhibition. The GenBac3 assay was conducted as a multiplex reaction, in which 100 copies of IAC (Haugland et al., 2010) was spiked into each sample, and inhibition to the samples was measured by creating acceptable bounds. Some samples were outside the bounds of the acceptable cycle thresholds, but not beyond the extent that was observed for samples at those concentrations, suggesting acceptable levels of inhibition. Negative Template Controls were included for all plates, revealing no contamination of qPCR reactions.

Table S5.1 Primers and Probes used to quantify markers in this study

Target	**Primers and Probes**	**Range of Quantitation / 100 ml**	**Reference**
EC23S857 **(*E. coli*)**	Forward: GGTAGAGCACTGTTTTGGCA Reverse: TGTCTCCCGTGATAACTTTCTC Probe: [FAM]TCATCCCGACTTACCAACCCG-[TAMRA]	10^8-10^3	Chern et al. 2012
GenBac3 **(General *Bacteroidales*)**	Forward: GGGGTTCTGAGAGGAAGGT Reverse: CCGTCATCCTTCACGCTACT Probe: [FAM]CAATATTCCTCACTGCTGCCTCCCGTA[TAMRA]	10^9-10^3	(Dick and Field, 2004; Siefring et al., 2008)
Rum2Bac **(Ruminant Associated *Bacteroidales*)**	Forward: ACAGCCCGCGATTGATACTGGTAA Reverse: CAATCGGAGTTCTTCGTGAT Probe: [FAM]-ATGAGGTGGATGGAATTCGTGGTGT-[BHQ-1]	10^9-10^2	Mieszkin et al. 2009

Appendix 2: Uncertainty Analysis Methods

Based on the results of the decay study, we conducted additional simulations specific for the initial concentration observed, the variability measured in flow rates and fecal indicator measurements, and the sampling times selected for this study. For these simulations, three identical replicate datasets were generated, assuming mean concentrations as calculated from the following decay models:

Equation S1: Chick Watson Model

$C_t = C_0 e^{-kt}$, where k = decay rate (hour^{-1}), t = time (hours)

Equation S2: Delayed Chick Watson Model

If t <= lag: $C_t = C_0$

If t > lag: $C_t = C_0 e^{-k(t-lag)}$, where lag = duration of no decay period (hours),

Equation S3: Delayed Chick Watson with Tail

If t <= lag: $C_t = C_0$

If t > lag: $C_t = (C_0 - Tail)e^{-k(t-lag)} + Tail$, where Tail = proportion of C_0 (unitless, (Copy number/100 ml)/C_0)

Time points and ranges of quantitation for these simulations were identical to our sampling scheme for the simulated indicator. Error terms were then added to each simulated value, assuming a distribution of the errors as Normal(0, .25*C_t). The 25% standard deviation was based on the observed variability in the data generated for the closed system observations from our decay study. Decay models were then fit to the generated datasets to provide a no flow assessment of fit including the variability.

To simulate flow similar to what was observed in our study, we added a flow term to the simulated data based on flow rate values and errors detected in our study. Variable flow rates

were simulated for each hour based on the observed flow rates (i.e. a Normal(5331,122) distribution for the high flow simulations, and Normal(1308,27) for the low flow simulations). We used the average flow rate (5331 or 1308 ml/hour) to adjust the data, which was fit by decay models. Comparisons were made to the accuracy of the model parameter estimates based on the simulated values. The range of accuracy was determined based on the range of decay rate parameters that were fit by the confidence interval for each estimate. Simulations were repeated each time for each combination of parameters.

Additional Supplemental Tables

Table S5.2 Best fits for datasets shortened to low flow last sampling point

Marker	**Dataset**	**Best Fitting Models**	**Decay Rate (per hour)**	**Lag (hours)**	**Tail**
Cultivable *E. coli*	Open – Low Flow	Delayed Chick Watson	0.053 (0.0076)	82 (8.1)	--
	Closed	Delayed Chick Watson with Tail	0.044 (0.026)	65 (13)	9.3 (5.2)
	Partially Closed	Delayed Chick Watson with Tail	0.068 (0.020)	96 (0.017)	10 (0.26)
EC23S857	Partially Closed	Chick Watson	0.0041 (0.0029)	--	--
	Low Flow	Delayed Chick Watson	0.033 (0.012)	51 (19)	--
	Closed	Chick Watson	0.012 (0.0029)	--	--
GenBac3	Partially Closed	Delayed Chick Watson	0.050 (0.012)	99 (12)	--
	Low Flow	Delayed Chick Watson	0.088 (0.012)	98 (7.1)	--
	Closed	Delayed Chick Watson	0.051 (0.0060)	70 (7.5)	--
Rum2Bac	Partially Closed	Delayed Chick Watson	0.076 (0.012)	102 (7.7)	--
	Low Flow	Delayed Chick Watson	0.068 (0.017)	81 (14)	--
	Closed	Delayed Chick Watson	0.074 (0.0060)	70 (6.13)	--

Table S5.3 Best Fits for datasets shortened to high flow last sampling point. Analyses for EC23S857 are not included as there was not enough data to fit the models. No tailing effects were detected for datasets shortened to this duration.

Marker	**System**	**Best Fitting Model**	**Decay Rate Estimate (hour-1)**	**Lag (Hours)**
Cultivable *E. coli*	Closed	Delayed Chick Watson	0.032 (0.012)	62 (15)
	Partially Closed	Delayed Chick Watson	0.078 (0.012)	96 (0.023)
	Low Flow	Delayed Chick Watson	0.065 (0.024)	87 (9.0)
	High Flow	Chick Watson	$1.9e^{-9}$ ($1.93e^{-6}$)	
GenBac3	Closed	Chick Watson	0.125084	
	Partially Closed	Chick Watson	0.0065 (0.0029)	--
	Low Flow	Delayed Chick Watson	0.035 (0.012)	65 (13)
	High Flow	Chick Watson	0.0091 (0.0029)	--
Rum2Bac	Closed	Delayed Chick Watson	0.090 (0.0084)	72 (0.55)
	Partially Closed	Delayed Chick Watson	0.083 (0.020)	99 (0.33)
	Low Flow	Delayed Chick Watson	0.053 (0.012)	70 (8.0)
	High Flow	Chick Watson	0.015443 (0.029)	--

CHAPTER 6

GENERAL CONCLUSION

Lauren E. Brooks

The interpretation of fecal indicator data gathered for water quality monitoring can depend on the survival characteristics of the selected indicator. Different decay properties for indicators make statements regarding human health risks and identification of contributing fecal sources challenging. Despite a large body of literature, uncertainties regarding indicator decay in the environment limit some applications of indicator data, such as inclusion in spatial modelling efforts for identification of nonpoint sources. Information addressing unidentified sources of variability in decay estimates is essential to making advances in the use of this type of water quality monitoring.

We used Bayesian meta-analysis and meta-regression techniques to synthesize existing decay rate estimates for a number of indicators. This allowed us to identify credible ranges for published decay rate estimates in the environment and identify environmental variables that affect decay rate estimates. Temperature was found to be significant for all indicators, while light was only significant for culture-based indicators. The results show that despite the inclusion of significant variables, high amounts of unexplained variation across decay rate estimates exist, leading to wide credible intervals. Additionally, gaps in the data were detected, especially for ruminant specific markers, meaning application of the synthesized decay rate estimates beyond the specific experimental designs tested previously is not advised.

We observed differential decay for ruminant-associated markers, cattle-associated markers, and general indicators, using a mesocosm seeded with fecal matter from dairy cattle. Differences were noted in the overall shape of the decay profiles using global model fitting, demonstrating the effectiveness of this technique for comparison of nonlinear decay curves. Cattle-associated markers were present at low abundance, and crossed the lower limit of quantitation within the first three days of the experiment, while other markers were detected for

five days, and culturable *E. coli* persisted for seven days. *Bacteroidales* associated markers displayed similar decay profiles to one another, but differed from *E. coli* and CowM2. The decay profile from CowM3 did not differ from any other indicator, possibly due to the limited data available before the concentration passed below the limit of quantitation. These findings highlight the need for caution when selecting markers for source tracking, as differential persistence can lead to different interpretations regarding sources contributing to fecal contamination.

We used simulations to verify the feasibility of conducting decay studies in an open system. We found that while it is possible to adjust for loss caused by removal from the system, thus allowing only decay to be considered, the openness of the system results in a loss of sensitivity, as the duration of measurable concentrations is inherently shortened. Building upon the simulations, we designed an open system and compared decay profiles for several fecal indicators in open, closed, and partially closed systems. We found different decay profiles across systems, with a general trend of increased lag period prior to exponential decay as the system became more open. Our findings demonstrate one example of the effects that experimental design, specifically system selection, has on our understanding of fecal indicator decay.

While a large volume of data is available addressing fecal indicator decay, we have demonstrated that additional variables, such as experimental design, including marker and system selection, as well as techniques chosen for statistical comparisons, influence the results of decay studies, and thus, interpretation of water quality monitoring results. As the use of fecal indicators continues, more research is necessary to continue to identify additional sources of discrepancy in decay study findings and observe decay in more realistic environments.