Water Research 104 (2016) 262–271

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Bayesian meta-analysis to synthesize decay rate constant estimates for common fecal indicator bacteria

Lauren E. Brooks, Katharine G. Field*

Oregon State University, Department of Microbiology, 226 Nash Hall, Oregon State University, Corvallis, OR, 97331, USA

ARTICLE INFO

Article history:
Received 2 May 2016
Received in revised form
29 July 2016
Accepted 1 August 2016
Available online 3 August 2016

Keywords:
Fecal indicator bacteria
Microbial source tracking
Persistence
Decay
Inactivation

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. *Escherichia 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 culturable 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 *Escherichia 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 (Fig. 1) and the significance of environmental conditions make it difficult to forecast how FIB will persist in the

* Corresponding author.
E-mail addresses: brooksla@onid.orst.edu (L.E. Brooks), kate.field@oregonstate.edu (K.G. Field).

http://dx.doi.org/10.1016/j.watres.2016.08.005
0043-1354/

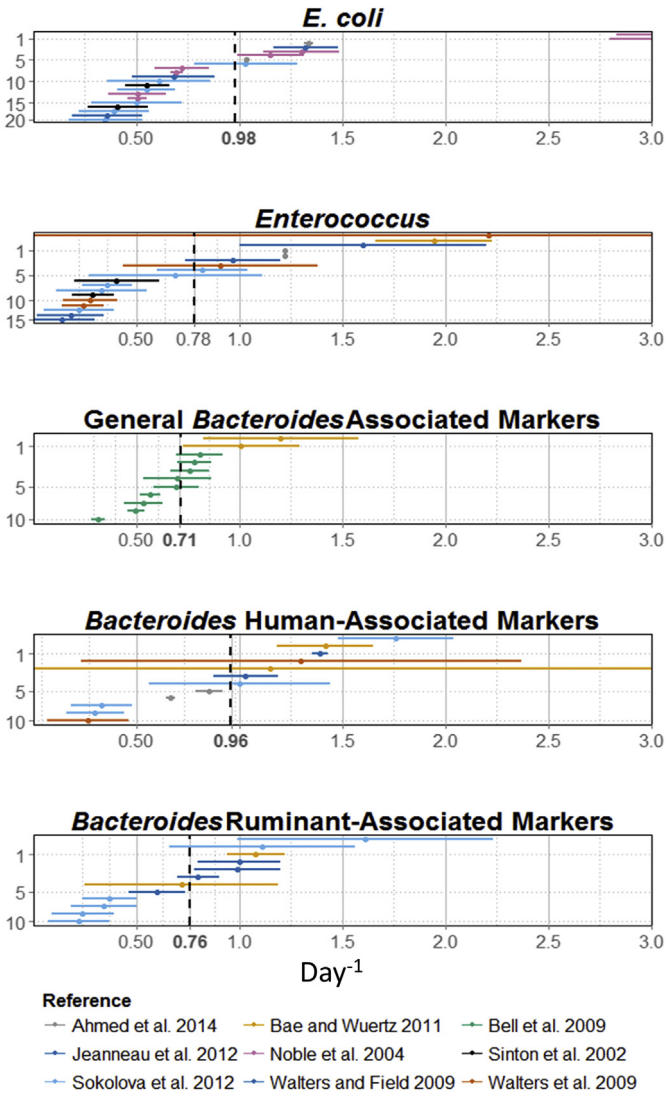

Fig. 1. First order decay rate estimates and standard errors from published papers for selected FIB reveal inconsistency in the literature among decay rate estimates.

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 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:

2.1.1. 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.

2.1.2. 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.

2.1.3. Metadata

Another criterion for inclusion in this study was the reporting of metadata. Data were compiled for a number of possible covariates (Supplemental Table 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:

$$\mu \sim Norm(0.5\ ,\ 10\ )$$

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

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)

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

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

Prior Densities:

$$\theta_i \sim Norm(0.5,\ 10)$$

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

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

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 autocorrelation, 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). 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* has 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 (Fig. 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

Table 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 Wuertz 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 Wuertz 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 Wuertz 2011 Bell et al., 2009	Light[b] Temperature[c] Experiment Length[c] Primer set[c]	Bae and Wuertz, 2009 Bae and Wuertz 2015 Dick et al., 2010 Schulz and Childers 2011 Tambalo et al., 2012
Bacteroides Human-Associated Markers	Ahmed et al., 2014 Bae and Wuertz 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 Wuertz 2015 Dick et al., 2010 Liang et al., 2012 Sokolova et al., 2012 Tambalo et al., 2012
Bacteroides Ruminant-Associated Markers	Bae and Wuertz 2011 Sokolova et al., 2012 Walters and Field 2009	Light[b] Temperature[c] Experiment Length[c]	Bae and Wuertz, 2009 Bae and Wuertz 2015 Liang et al., 2012 Tambalo et al., 2012

Table 2
Model output and statistics for overall and specific combinations for each target.

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	−0.06, 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 spp.	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)

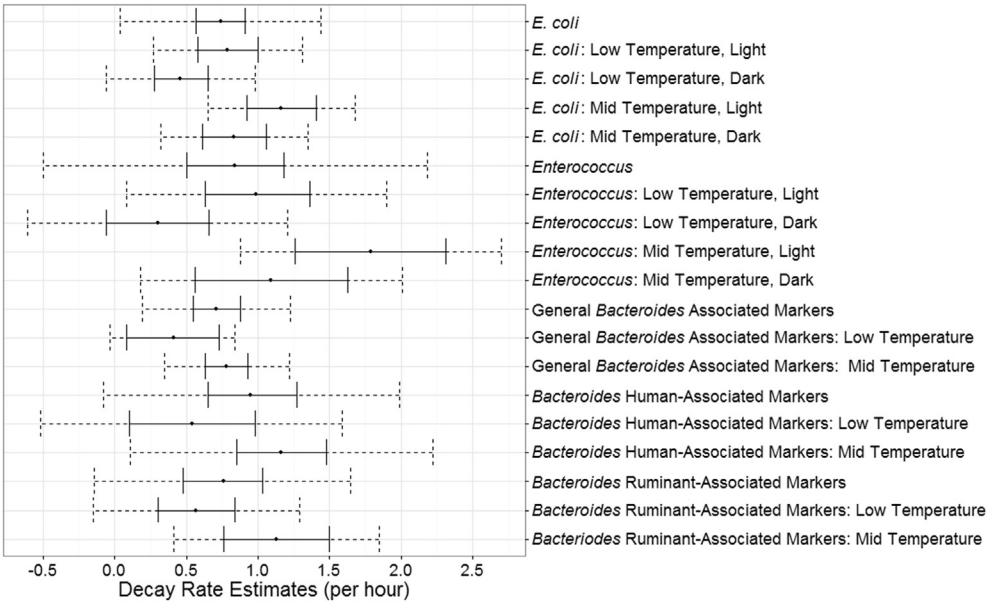

Fig. 2. 95% Credible intervals for the mean (solid line) and distribution (dashed line) generated from each model output.

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).

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 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 ($\mu = 0.84\ day^{-1}$) 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 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 (0.19–1.24), relative to those constructed for *E. coli* and *Enterococcus* spp. (Fig. 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 (Fig. 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 is effective only when contamination is fresh, while variation in decay over time leads 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 (Fig. 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

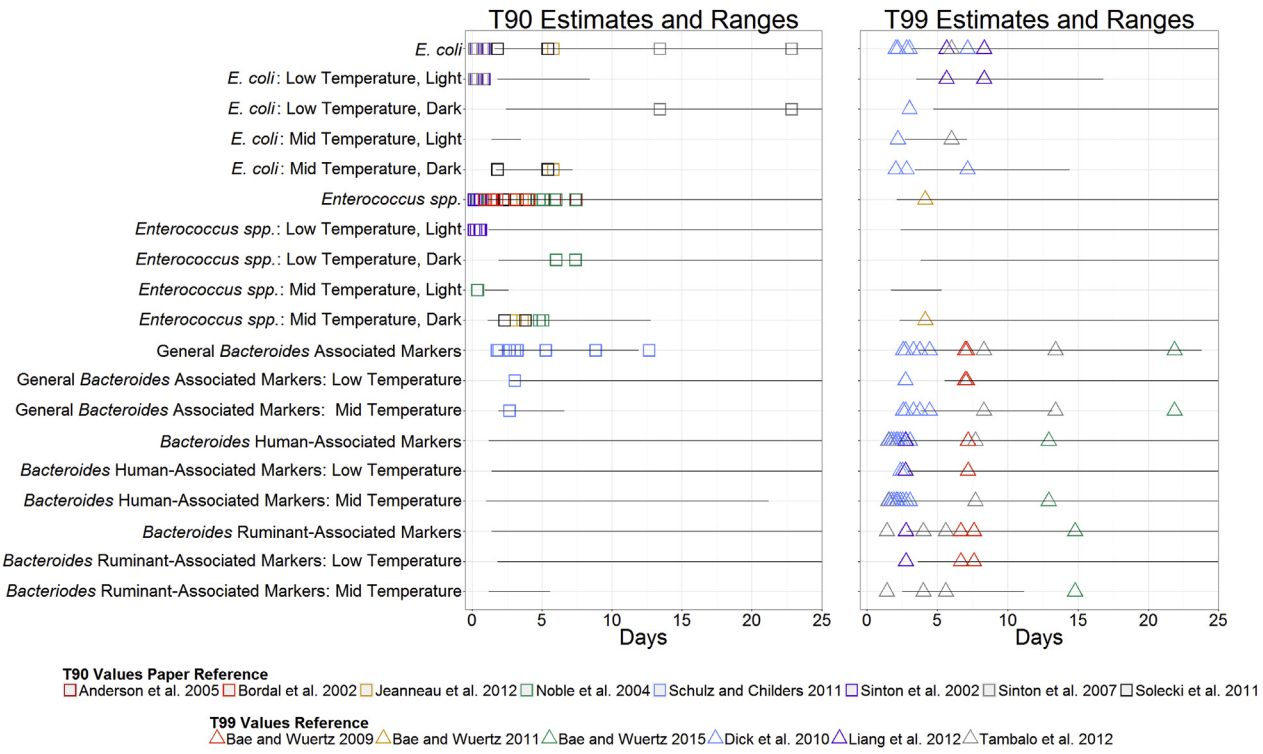

Fig. 3. Data points used for posterior predictive checking and credible T90/T99 ranges generated by model. Lines represent the ranges predicted from the output of the models.

example, the microbial community has been shown to play a major role in the removal of FIB (Korajkic et al., 2014). Most studies measuring this effect have compared the differences when the environmental microbial community is removed from the system, thereby showing that the microbial community is significant, but not providing information on 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 modeling 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 (Fig. 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 more intense 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 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 even within 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., 2014), 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

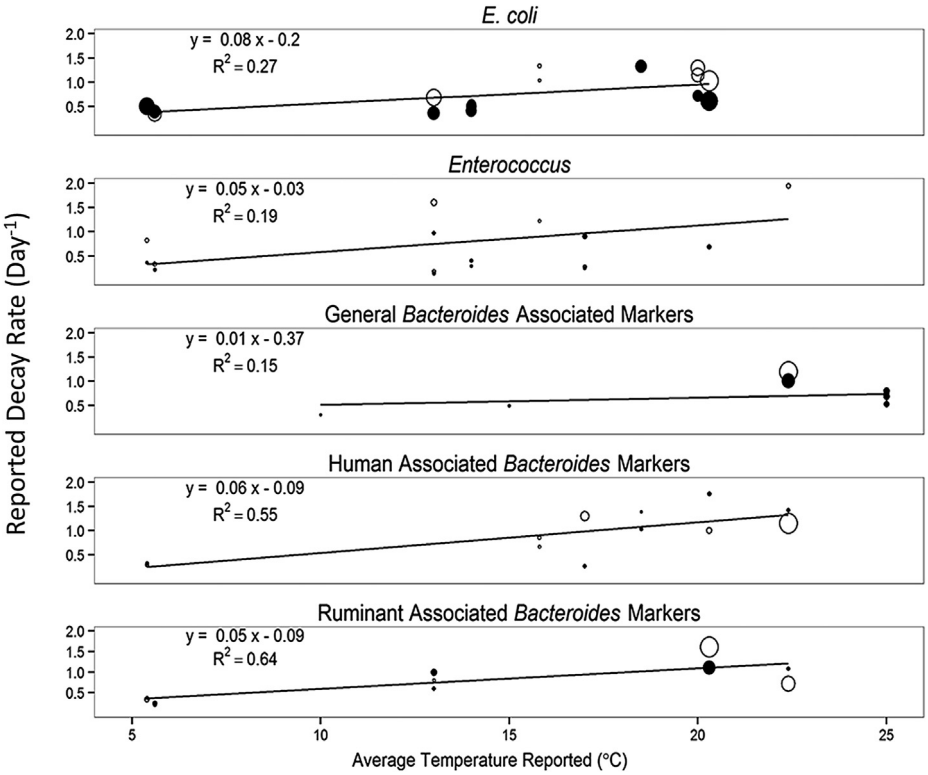

Fig. 4. Simple linear regression showing effect of temperature on decay rate estimates with a distinction for data points generated under light treatments (○) and dark treatment (●). Circle size corresponds to standard error estimates.

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. Factors that could not be included due to lack of data, such as DO, nutrients, TSS, and the microbial community, may be responsible for some of the unexplained variation. Our findings suggest that more data are 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 are 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.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.watres.2016.08.005.

References

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. 59, 348–354.

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.

Bae, S., 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., 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., 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., Sayler, G.S., 2009. Factors influencing the persistence of fecal *Bacteroides* in stream water. J. Environ. Qual. 38, 1224.

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.

Bivand, R., Ono, H., Dunlap, R., Stigler, M., 2015. Package ClassInt: Choose Univariate Class Intervals.

Bordalo, A.A., Onrassami, R., Dechsakulwatana, C., 2002. Survival of faecal indicator bacteria in tropical estuarine waters (Bangpakong River, Thailand). J. Appl. Microbiol. 93, 864–871.

Dick, L.K., 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., 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., 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., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7, 457–472.

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

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.

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.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., 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., 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., 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., 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. AEM.00388–14.

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R., 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.

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.

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.

Meng, X.-L., 1994. Posterior predictive p-values. Ann. Stat. 22, 1142–1160.

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.

Okabe, S., 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., 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., 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.

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.

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.

Su, Y.-S., Yajima, M., 2014. R2jags: a Package for Running Jags from R.

Sutton, A.J., 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., 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., Higgins, J.P.T., 2002. How should meta-regression analyses be undertaken and interpreted? Stat. Med. 21, 1559–1573.

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.

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.

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. 47, 6849–6861.