Isolating the impact of septic systems on fecal pollution in streams of suburban watersheds in Georgia, United States

Robert A. Sowah[1*], Mussie Y. Habteselassie[1], David E. Radcliffe[2], Ellen Bauske[3], Mark Risse[4]

[1]Crop and Soil Sciences, The University of Georgia Griffin Campus, 1109 Experiment St, Griffin, GA 30223

[2]Crop and Soil Sciences, The University of Georgia, 3111 Carlton St, Athens, GA 30602

[3]Center for Urban Agriculture, The University of Georgia Griffin Campus, 1109 Experiment St, Griffin, GA 30223

[4]The University of Georgia, Marine Extension and Georgia Sea Grant, 1030 Chicopee Building, Athens, GA 30602

Abstract

The presence of multiple sources of fecal pollution at the watershed level presents challenges to efforts aimed at identifying the influence of septic systems. In this study multiple approaches including targeted sampling and monitoring of host-specific *Bacteroidales* markers were used to identify the impact of septic systems. Twenty four watersheds with septic density ranging from 8 – 373 septic units/km^2 were monitored for water quality under baseflow conditions over a 3-year period. The levels of the human-associated HF183 marker, as well as total and ruminant *Bacteroidales*, were quantified using quantitative polymerase chain reaction. Human-associated *Bacteroidales* yield was significantly higher in high density watersheds compared to low density areas and was negatively correlated (r = -0.64) with the average distance of septic systems to stream in the spring season. The human marker was also positively correlated with the total *Bacteroidales* marker, suggesting that the human source input was a significant contributor to total fecal pollution in the study area. Multivariable regression analysis indicate that septic systems, along with forest cover, impervious area and specific conductance could explain up to 74% of the variation in human fecal pollution in the spring season. The results suggest septic system impact through contributions to groundwater recharge during baseflow or failing septic system inputs, especially in areas with >87 septic units/km^2. This study supports the use of microbial source tracking approaches along with traditional fecal indicator bacteria monitoring and land use characterization in a tiered approach to isolate the influence of septic systems on water quality in mixed-use watersheds.

Key words: *Bacteroidales* markers, water quality, fecal pollution, septic density, HF183, baseflow

1. Introduction

Septic systems are used widely for wastewater treatment in southeastern U.S. It is estimated that 37 – 48% of all housing units in North and South Carolina, Georgia and Alabama use septic systems for wastewater treatment (U.S. EPA, 2002). This usage rate exceeds the national average of 23% according to the same report. It is also estimated that more than 33% of all new housing units in the U.S. are on septic systems, which make septic systems second only to centralized systems, in terms of the number of households served, in the wastewater management infrastructure (U.S. EPA, 2002). A significant number of these septic systems are in suburban areas, with some reports showing that the majority of septic systems are now located in suburban communities compared to rural areas (MNGWPD, 2006; U.S. EPA, 2002). The upward trend in septic systems' use has coincided with widespread fecal pollution of surface waters across the U.S., raising questions about the potential contribution of septic systems to water quality impairment at the watershed level (Verhougstraete et al., 2015; Sowah et al., 2014).

Data from the United States Environmental Protection Agency implicates fecal pathogens as the leading cause of water quality impairment in the country (U.S. EPA, 2016). Frequently, the sources of fecal matter impacting surface water bodies have proved difficult to isolate especially in urbanizing areas with mixed land uses (Liang et al., 2013). Typically, fecal pollution of surface water resources derives from two or more sources within a watershed (Jent et al., 2013; Chin et al., 2009). These sources may include wildlife, livestock, manure applications and human inputs (wastewater treatment facility discharges, faulty septic systems and leaky sewers). The inputs from these sources can also vary on both temporal and spatial scales which further complicate pollution management. The traditional approach of monitoring fecal indicator bacteria (FIB) such as *Escherichia coli* (*E. coli*), enterococci and fecal coliforms is not capable

of differentiating sources of fecal pollution at the watershed scale (Tran et al., 2015; Sauer et al., 2011; Field and Samadpour, 2007). This has led to new approaches that rely on FIB monitoring along with analysis of watershed characteristics (such as land use, imperviousness and septic density) and environment conditions (including precipitation and rainfall intensity) in addition to microbial source tracking (MST) with human-specific markers to discern fecal pollution sources at the watershed scale (Verhougstraete et al., 2015; Tran et al., 2015; Sowah et al., 2014; Peed et al., 2011).

Understanding the link between septic systems and fecal pollution of surface waters is critical to improving water quality. It is a well-known fact that individual septic systems can contribute to fecal pollution of ground and surface waters if the systems are not properly designed, sited and maintained (Schneeberger et al., 2015; Habteselassie et al., 2011; Humphrey et al., 2011). Until recently, there was limited information on the cumulative effects of septic systems at the watershed scale leading to inadequate regulation and improper management of septic systems throughout the U.S. (Gregory et al., 2013; Carey et al., 2012; Farnleitner et al., 2011; Swann, 2001). Recent studies have attempted to bridge this knowledge gap by examining the relationships between septic system use and fecal bacteria levels at the watershed level (Verhougstraete et al., 2015; Sowah et al., 2014; Peed et al., 2011). Using host-specific MST markers in combination with low-order head-watershed sampling and precipitation information, Peed et al. (2011) observed a positive correlation between the concentration of human-specific marker and septic systems following precipitation. Predictive models developed by Sowah et al. (2014) indicated that septic system density and average distance of septic to streams were significant factors driving fecal pollution in suburban watersheds of Georgia. More recently, Verhougstraete et al. (2015) used classification and regression tree analysis (CART) to

82 demonstrate a link between increasing septic system numbers and bacterial concentrations. The
83 findings of these studies show that MST methods employing human-associated markers such as
84 the HF183 (Seurinck et al., 2005), when combined with land use information, can be a powerful
85 tool for isolating septic system influence at the watershed level.

86 As promising as these studies are for water resources management, questions still remain
87 about the relationship between septic system density and fecal pollution under different
88 hydrologic and climatic conditions. The applicability of MST methods to different geographic
89 areas and land use scenarios need to be assessed to improve confidence in this approach. This
90 information is critical in suburban watersheds with mixed land uses and a pool of fecal sources
91 including failing septic systems, leaky sewers, livestock, pets and wildlife that can present
92 challenges to watershed managers. Also, previous studies utilizing MST methods have focused
93 on relatively low to medium density septic impacted watersheds, that is, areas with <100 septic
94 units/km^2 (Verhougstraete et al., 2015; Peed et al., 2011). The current study provides a
95 comprehensive assessment of septic system impacts in watersheds with a gradient of septic
96 densities ranging from 8 – 373 septic units/km^2. In addition to examining septic system impacts,
97 this study will also address questions about the influence of leaky sewer pipes on fecal pollution
98 loads.

99 The overarching goal of this study therefore was to evaluate the impact of septic systems on
100 fecal pollution loads in suburban streams. Our main objective was to determine the relationship
101 between septic system variables and fecal pollution loads at the watershed level. The present
102 study would use MST approaches together with targeted monitoring to capture septic system
103 influence across a wide range of septic system densities typical of suburban areas in the
104 southeastern U.S. Additionally, seasonal and temporal trends in fecal pollution loads as impacted

105 by increasing septic system density, land use characteristics and water quality parameters will be
106 examined. Finally, the utility of MST for identifying septic system influence on water quality
107 will be evaluated as part of a tiered approach to fecal source identification at the watershed level.

108 2. Materials and methods

109 2.1. *Study Area*

110 The study area is in Gwinnett County, northeast of Atlanta, GA and was previously described in
111 Sowah et al. (2014) and Landers and Ankcorn (2008). It has a mean annual rainfall of about
112 1245 mm. The study area consists of 24 watersheds which range in size from 0.2 to 8.8 km^2. A
113 summary of watershed characteristics is provided in Table 1. The selected watersheds are in the
114 Ocmulgee and Oconee River basins, which drain to the Altamaha River, and ultimately into the
115 Atlantic Ocean. A map of the study area and watershed boundaries is presented in Figure 1. The
116 selected watersheds are typical of suburban watersheds along the Interstate 85 corridor in the
117 southeastern Piedmont region of the U.S. The watersheds represent a gradient of land use
118 characteristics from low to high density of septic systems representing low density residential to
119 medium density residential land use. The watersheds are classified into two groups: low density
120 (LD) watersheds and high density (HD) watersheds. An arbitrary threshold of <38 septic
121 units/km^2 was set for LD watersheds and >77 septic units/km^2 for HD watersheds. The criteria
122 for watershed classification was based on U.S. EPA's designation of areas with >15 septic
123 units/km^2 as regions of potential groundwater contamination (U.S. EPA, 1977). The LD
124 threshold was raised to account for improvements in septic system technology and regulation in
125 the last two decades. The predominant soil types in the study area are the moderately permeable
126 Appling and Pacolet soil series (Web Soil Survey, 2016). The soils in this region are typically
127 underlain by saprolite with saturated hydraulic conductivities (K_{sat}) in groundwater ranging from

128 0.7 – 62.4 cm/day (Amoozegar et al., 1993). Other watershed characteristics were determined
129 from spatial datasets of land cover (NLCD, 2006), septic systems (GCBC, 2013) and sewer lines
130 (GCDPU, 2004).

131 2.2. *Sample collection*

132 Surface water samples from streams in the 24 selected watersheds were collected during
133 baseflow on 9 synoptic sampling events spanning November, 2011 to July, 2014. Synoptic
134 sampling coincided with the spring (n = 72; March 2012, April 2013 and March 2014), summer
135 (n = 72; July 2012, 2013 and 2014) and fall (n = 72; November 2011, 2012 and 2013) seasons.
136 Baseflow conditions were determined using antecedent precipitation and hydrograph from two
137 USGS gage stations near the study area (USGS, 2016). Also, baseflow sampling coincided with
138 periods of zero precipitation at least 72 hours prior to the sampling event. At each monitoring
139 station, samples were collected in duplicate in 1-L sterile high-density polypropylene bottles.
140 Samples were kept on ice and transported to the laboratory for analysis of FIB (usually within 6
141 hours of sample collection). Sample collection and analysis followed guidelines of the National
142 Field Manual for the Collection of Water-Quality Data (USGS, variously dated). Baseflow
143 discharge (m^3/s) was measured at each monitoring point during sampling events by our project
144 partners from the United States Geological Survey (USGS) Georgia Water Science Center in
145 Atlanta. The velocity-area method (Rantz, 1982) was used for discharge measurements.

146 2.3. *Standard water quality analysis*

147 Data for *E. coli* and enterococci concentrations and water quality parameters such as pH,
148 temperature, dissolved oxygen and specific conductance which covers synoptic sampling from
149 Nov. 2011 – Nov. 2013 are presented in Sowah et al. (2014). In this study, two additional

synoptic samples (March 2014 and July 2014) were analyzed for *E. coli* and enterococci using the Colilert-18 and Enterolert kits (IDEXX Laboratories Inc., Westbrook, ME). The methods used followed similar protocols outlined in Sowah et al. (2014). Standard water quality parameters including pH, temperature, dissolved oxygen and specific conductance were also measured during sampling with a calibrated Quanta multi-parameter probe (HYDROLAB, Loveland, CO).

2.4. *Bacterial DNA extraction*

Bacterial DNA was concentrated by filtering 100 ml of water sample through 0.40 µm polycarbonate filters (GE Whatman, Pittsburgh, PA). Filters were placed in microcentrifuge tubes and stored for up to 2 years at -80^{o}C prior to DNA extraction. Filters were directly extracted using the PowerFecal DNA isolation kits (MOBIO Laboratories, Carlsbad, CA) according to manufacturer recommendations. At least one extraction blank was processed along with each batch of synoptic surface water samples. The method blanks were extracted from sterile ultrapure water. DNA was eluted to a final volume of 50 µl and stored for not more than 3 months at -20^{o}C before analysis.

2.5. *Quantitative Polymerase Chain Reaction (qPCR)*

To identify septic influence in our watersheds, we performed qPCR assays targeting the HF183 human-associated *Bacteroidales* marker (Seurinck et al., 2005) in extracted water samples. The MST markers used in this study were selected on the basis of their widespread use and validation in different geographical regions. Method comparison studies in the U.S indicate that the HF183 marker is one of the high performing markers in terms of specificity and sensitivity (Boehm et al., 2013; Layton et al., 2013; Shanks et al., 2010). A ruminant-associated marker BacR

(Reischer et al., 2006) and total *Bacteroidales* marker AllBac (Layton et al., 2006) were also enumerated to identify drivers of fecal pollution in the watershed. Similar to the HF183 marker, the AllBac marker has seen widespread application in different locations and demonstrated high sensitivity to fecal material from a wide range of animals. Recently, the specificity of the AllBac and other generic Bacteroidales markers have been questioned due to their cross-reaction with environmental *Bacteroidales* strains (Vierheilig et al., 2012; van der Wielen and Medema, 2010). In the absence of new and more efficient general *Bacteroidales* marker - the development of a new marker was outside the scope of this study, we believe the AllBac marker can provide useful information for MST analysis and furthermore the data generated in this study can be compared to previous studies that employed this marker. Finally, the BacR marker has proved effective for discriminating ruminant sources from other sources of fecal pollution as attested to by method validation studies in the U.S. (Boehm et al., 2013; Raith et al., 2013). Human and ruminant sources of fecal pollution were the focus of this study as land use information suggests that these were the likely contributors to total fecal pollution in streams. These genetic markers target regions of the 16S rRNA genes of *Bacteroidales* in representative host groups. Detailed description of qPCR methods, primer sequences and quality control measures is provided in the supplemental material.

2.6. *Data Presentation and Statistical Analysis*

The data was summarized on a seasonal basis for presentation and statistical analysis. Data for spring, summer and fall represent the geometric mean of synoptic samples collected over the study period. For statistical analysis, samples that were below the limit of quantification and qPCR non-detects were replaced with values imputed using robust regression on order statistics (ROS). Regression on order methods of imputing non-detects are reported to provide better

195 results compared to other methods such as the maximum likelihood elimination approach or
196 substitution with detection limits (Helsel, 2010; Wong et al., 2009; Helsel, 2005). The ROS
197 method was used within the U.S. EPA's ProUCL Statistical tool to impute values of non-detects
198 based on a log-normal distribution of the detected values (U.S. EPA, 2013). The copies of total
199 and host-associated markers were expressed as marker yield in gene copies per second per square
200 kilometers (copies/s.km^2) by accounting for streamflow and watershed area. The yield of MST
201 markers was log-transformed to achieve normality for use in statistical tests. A closer look at the
202 watersheds in Figure 1 will show that nine of the watersheds (sites 3, 4, 5, 9 in LD watersheds
203 and sites 14, 18, 19, 20, 23 in HD watersheds) were nested within larger watersheds. In order to
204 meet independence assumption of statistical tests, we performed Durbin Watson test to check for
205 autocorrelation in *Bacteroidales* markers in the nested watersheds (Little et al., 2008). The
206 Durbin Watson test results showed that sites 3, 5, 14, 18 and 23 were auto-correlated with
207 watersheds downstream and as such were excluded from statistical analysis.

208 Two-way analysis of variance (ANOVA) was performed to determine the effect of septic
209 density and season on total and host-associated marker yield. The generalized linear model
210 procedure (proc GLM) was used in SAS 9.3 (SAS Institute, Cary, NC) to examine variations in
211 marker yield due to septic density and seasonal factors. The Tukey post-hoc multiple comparison
212 test was used to determine main effects and simple main effects following statistically significant
213 difference or interaction. Spearman rank correlations were performed to determine the
214 relationships between total and host-specific marker and septic density, average distance of
215 septic systems to streams, sewer line density, land use characteristics, FIB, and standard water
216 quality parameters. Additionally, multivariable linear regression models were developed to
217 determine the combination of land use and water quality variables driving human-associated

218 *Bacteroidales* levels in suburban streams. Variable selection in the regression analysis was based
219 on the backward elimination method using the proc REG procedure in SAS. Inclusion of
220 variables in the model depended on the variables meeting a significance threshold of $p = 0.05$ to
221 avoid over-parametrizing the regression models, and a variance inflation factor of <10 to reduce
222 multicollinearity of model variables (Gonzalez et al., 2012; Hathaway et al., 2010). Moreover,
223 residuals for response and explanatory variables were plotted and checked to confirm that
224 normality assumptions of the models were not violated. All statistical analysis was performed
225 with SAS 9.3 (SAS Institute, Cary, NC) and statistical significance was defined at $p \leq 0.05$
226 unless stated otherwise.

227 **3. Results**

228 3.1. *Analysis of qPCR Inhibition and limit of quantification*

229 Comparison of qPCR inhibition in water samples collected from the study area with sterile water
230 showed no statistically significant difference in C_T values for the spiked human-specific marker
231 (Table 2). Moreover, dilution of samples that tested positive and negative in qPCR tests did not
232 result in significant differences following re-run of the samples. This leads us to conclude that
233 qPCR inhibition did not significantly affect results from this study. Assay limits of quantification
234 determined using the standards were 3 gene copies per reaction for the human marker and 30
235 gene copies per reaction for the total and ruminant-specific markers.

236 3.2. *Distribution of total and host-associated Bacteroidales markers*

237 The *Bacteroidales* markers targeted in this study were widely distributed in streams impacted by
238 LD to HD of septic systems. Figure 2 shows the distribution of total, human and ruminant
239 associated markers grouped by season and septic density. The total *Bacteroidales* marker, which
240 captures fecal inputs from human, bovine, canine and swine among others, was detected in 100%

241 (n = 216) of samples. Total *Bacteroidales* concentration ranged from 3.3 – 6.7 $\log_{10}$ copies/100
242 ml whilst the yield of the total marker was between 4.7 – 8.8 $\log_{10}$ copies/s.km^2. The highest and
243 lowest concentrations of total *Bacteroidales* were recorded in fall in LD areas whereas the yield
244 was highest in spring in LD watersheds and lowest in fall in HD areas (Figure 2). The human
245 marker was quantifiable in 57% (n = 216) of the samples collected from the study area. The
246 frequency of detection, based on the number of samples that were quantifiable, was 63% for HD
247 watersheds and 51% for LD watersheds. The human marker varied from non-detect or below the
248 limit of quantification to a maximum of 3.7 $\log_{10}$ copies/100 ml. The highest concentration was
249 observed in spring in HD watersheds. Similarly, the maximum yield of human marker was
250 observed in HD areas in the spring. The ruminant marker was quantifiable in 65% (n = 192) of
251 surface water samples: 61% in HD watersheds compared to 68% in LD watersheds. A total of
252 192 samples (representing 8 synoptic sampling events) were included in the analysis of the
253 ruminant marker due to low detection (3 out of 24 samples) of the marker in samples collected in
254 November 2011. In order to impute non-detect values, more than 50% of the samples have to be
255 detected (ITRC, 2013). The ruminant marker varied from non-detect to maxima of 5.9 $\log_{10}$
256 copies/100 ml and 8.1 $\log_{10}$ copies/s.km^2 in concentration and yield respectively. The average
257 concentration and yield of the total, human and ruminant-associated markers in HD and LD
258 watersheds followed a seasonal trend, with low levels of the markers observed in fall in
259 comparison to spring and summer seasons. Statistical tests were performed with the yield of total
260 and host-associated *Bacteriodales* markers since the yield provides a robust estimate of marker
261 distribution, accounting for differences in hydrologic conditions and watershed area.

262 Results from two-way ANOVA tests show variations in the influence of septic system and
263 season on the yield of total and host-associated markers (Table 3). Analysis of the effect of septic

density and season on total *Bacteroidales* yield indicated no significant interaction between the factors on marker yield. Septic density also did not appear to influence total *Bacteroidales* yield in the watersheds. However, seasonal changes played a key role in the yield of total *Bacteroidales* with statistically significant differences ($p < 0.001$) between spring, summer and fall. Examination of the main effects showed that total *Bacteroidales* yield was significantly higher in spring and summer compared to fall ($p < 0.001$). There was however no difference in marker yields between spring and summer. For the human-associated marker, there was no significant interaction between septic density and season. The main effects showed statistical significance between levels of septic density ($p = 0.046$) and season ($p < 0.001$) for the human-specific marker. The results showed that the marker was significantly higher in HD watersheds compared to LD areas. Similar to the results for total *Bacteroidales*, the human marker yield was significantly higher in spring and summer compared to fall ($p < 0.001$), but not statistically different between spring and summer. Statistically significant interaction ($p = 0.002$) was observed between septic density and season for the ruminant-specific marker. Simple main effects analysis indicated that the ruminant marker was significantly higher in LD watersheds for all seasons compared to the fall in HD watersheds.

3.3. *Correlation analysis*

Spearman rank correlation for total and host-associated marker yield as influenced by land use characteristics and standard water quality parameters is presented in Table 4. Results indicate that the human marker was negatively correlated with average distance of septic to stream ($r = -0.64$, $p = 0.003$) for samples collected in spring. Septic density and the average distance of septic to stream were not correlated with human marker yield in summer or fall. However, the pooled data showed a significant negative correlation between the average distance of septic to stream

287 and human marker yield (r = -0.52, p = 0.008). The relationship between septic density and
288 ruminant marker, as expected, was negatively correlated (r = -0.76, p < 0.001) in fall. We also
289 observed a strong positive correlation between total *Bacteroidales* and human marker yield (r =
290 0.65, p = 0.005) for the pooled data. However, total *Bacteroidales* was not correlated with
291 ruminant marker in general. Sewer line density, a potential source of human fecal inputs into
292 streams, showed no correlation with human marker yield for all seasons and the pooled dataset.
293 In contrast, sewer line density was correlated with total *Bacteroidales* (r = 0.62, p = 0.006) in the
294 summer. In general, *E. coli* and enterococci yield were positively correlated with human marker
295 for all seasons and the pooled data. No correlations were observed between the total
296 *Bacteroidales* marker and land use and environmental parameters in spring. Overall, the percent
297 of agriculture land cover was a significant predictor of ruminant marker yield in this study (r =
298 0.57, p = 0.009).

299 3.4. *Multivariable linear regression analysis*

300 Multivariable regression models developed in this study were of the format $Y = \beta_0 + \beta_1 X_1 +$
301 $\beta_2 X_2 + \ldots \beta_n X_n$ where Y is the dependent variable, β_0 is the intercept, β_1 to β_n are parameter
302 estimates and X represents the explanatory variables. Using the proc REG procedure in SAS, we
303 examined the influence of land use characteristics and water quality parameters on human
304 marker yield on a seasonal level. The results, summarized in Table 5, show that septic density
305 and average distance of septic to stream were important variables explaining variations in human
306 marker yield in spring and fall. A strong septic system impact was detected in spring samples in
307 line with the observed outcome from correlation analysis. Moreover, the regression model for the
308 pooled data had septic system as the most critical factor driving human fecal pollution in the
309 study area. The percent of impervious cover could also explain some of the variation in human

310 fecal pollution overall. Water quality parameters such as specific conductance and water pH
311 were also significant explanatory variables for human fecal pollution in the study area. None of
312 the land use characteristics and environmental variables examined in this study were predictors
313 of human marker yield in the summer. The adjusted R^2 values for the seasonal models were 0.74,
314 0.31 and 0.47 in spring, fall and pooled data respectively.

315 **4. Discussion**

316 The overall goal of this study was to examine the influence of septic systems on fecal pollution
317 in suburban watersheds by tracking the sources of human fecal inputs impacting streams of
318 watersheds with varying septic system density. In a previous study, Sowah et al. (2014) assessed
319 septic system influence by monitoring the streams in these watersheds for standard FIB. The
320 results from the previous study which included comparison of FIB loads in HD and LD
321 watersheds, correlation and multivariable regression analysis pointed to the contribution of septic
322 systems to total fecal pollution in streams. Validation of the results from the previous study and
323 similar studies implicating septic systems in water quality impairment require the use of multiple
324 approaches as recommended in the literature (Noble et al., 2006; Boehm et al., 2003). The
325 present study employs the widely used HF183 human-associated *Bacteroidales* marker to track
326 the influence of septic systems on fecal pollution in streams. This marker has shown stability in
327 different geographical locations with comparatively better specificity and sensitivity in method
328 comparisons studies here in the U.S (Boehm et al., 2013; Stewart et al., 2013; Shanks et al.,
329 2010;), Europe (Gawler et al., 2007; Gourmelon et al., 2007) and other parts of the world
330 (Ahmed et al., 2009; Jenkins et al., 2009).

331 Seasonal trends in total and host-associated marker distribution in the study area compare to
332 the observed distribution of FIB from Sowah et al. (2014). Overall, FIB and *Bacteroidales*

marker yields were higher in samples collected in spring and summer compared to fall. The seasonal trends are also symptomatic of the underlying seasonal changes in hydrologic conditions. Table S2 in the supplemental material shows the changes in climatic and hydrologic conditions in the study area. Average baseflow discharge in this study was relatively low in fall compared to summer and spring (Table S2 in supplemental material). This pattern in flow and fecal pollution indicators suggests that the sources of fecal pollution impacting the streams in this study are both temporally and seasonally stable in the study area. This seasonal trend, in relation to human-associated *Bacteroidales*, is suggestive of a continuous source of fecal pollution such as wastewater treatment discharges, leaky sewer pipes, failing septic systems or septic effluent transported through groundwater into streams (Carroll et al., 2005). Our research, however, shows that none of the streams in this study are impacted by permitted wastewater treatment discharges (GAEPD, 2016). Moreover, sewer pipes appear to be an insignificant source of human fecal inputs in these watersheds from our correlation and regression analysis. This leaves septic systems as the primary continuous source of human fecal pollution in the study area.

Results of ANOVA and correlation analyses support our hypothesis that septic systems are a significant source of human fecal pollution in HD areas. The observed increase in human marker yield in streams of HD watersheds compared to LD areas, and strong correlation of human marker with average distance of septic to stream are suggestive of pronounced septic system impact in areas with septic densities above 87 septic units/km^2. The linear increase in human marker with decrease in septic distance to stream was particularly strong in spring which coincides with the seasonal shallow groundwater table in the study area. Combined with moderate to high hydraulic conductivity of saprolite in the saturated zone, the shallow groundwater can act as a conduit for the transport of effluent from the dense network of septic

356 drainfields into nearby streams. The influence of septic systems on baseflow water quality has
357 been the subject of recent studies including work by Peed et al. (2011) that found no correlation
358 between septic systems and human-associated *Bacteroidales* marker under low flow conditions.
359 This study however confirms that under baseflow conditions, the influence of septic systems
360 depends on seasonal trends in hydrologic conditions within our study area. The marked septic
361 system impact in spring is not surprising considering the reported interconnectivity of
362 groundwater and surface water systems in the Southern Piedmont region (Clarke and Peck,
363 1991). Evaluation of the relationships between *Bacteroidales* markers showed a positive
364 correlation between human and total *Bacteroidales* yield for the pooled data, suggesting the
365 significant contribution of human sources to total fecal pollution in the study area. The relative
366 contribution of the human marker to total *Bacteroidales* yield in streams was examined in this
367 study due to the differential persistence of *Bacteroidales* markers in the aquatic environment, and
368 inadequate information about the copies of *Bacteroidales* genetic markers in different host
369 animals and humans (Tambalo et al., 2012; Dick et al., 2010; Walters and Field, 2009).

370 Our data shows significant differences in ruminant influence in LD and HD watersheds,
371 with LD areas recording higher ruminant marker yield. This result was expected as stream walks
372 of the watersheds in the summer of 2015 revealed that livestock, particularly cattle had access to
373 the streams in watershed numbers 6, 7, and 8. The positive correlation of ruminant marker with
374 total *Bacteroidales* in summer highlights the seasonal nature of animal impacts, especially in the
375 LD watersheds with higher agricultural and forest cover. The prominent agricultural land uses
376 include hay and pastures which are commonly grazed by cattle and horses. The strong negative
377 correlation between the ruminant marker and septic density in fall was expected as increasing
378 septic footprint was associated with low agricultural land use. The diversity of animal hosts

379 contributing to the ruminant signature in streams and the potential differences in persistence of
380 this marker may explain the seasonal differences in the influence of land use factors such as
381 forest cover and agricultural activities (Table 4). The pooled data in contrast, isolates agriculture
382 as the predominant animal source of ruminant marker in the study area.

383 In general, poor correlations have been reported between human *Bacteroidales* markers and
384 FIB levels in previous MST studies (Sauer et al., 2011; Edge et al., 2010; Okabe et al., 2007a).
385 These studies have, however, been largely focused on pollution incidents arising from storm
386 runoff. Storm induced pollution can however originate from diverse sources at the watershed
387 level which leads to confounding results. In this study we found moderate to strong positive
388 correlations between human marker and *E. coli* and enterococci levels. The overall correlation
389 coefficients between the human marker and *E. coli* and enterococci were 0.71 and 0.57
390 respectively. It has been reported that FIB and *Bacteroidales* markers differ in their persistence
391 profiles in the aquatic environment (Ballesté et al., 2010; Okabe et al., 2007b). The observed
392 positive correlations between FIB and the human marker therefore suggest a continuous source
393 (e.g. septic systems) of human fecal pollution in the study area. In addition to septic systems,
394 forest cover was also significant in driving variations in host-associated marker yield in suburban
395 streams. The effect of forest cover was particularly pronounced in spring for both human and
396 ruminant markers. With respect to the human marker, forest cover was negatively correlated,
397 supporting our observation of increasing human inputs from areas with a greater septic system
398 footprint and low forest cover. Surprisingly, forest cover was negatively correlated with the
399 ruminant marker in spring which suggests that ruminant animals that were not directly associated
400 with forest cover may be contributing to ruminant inputs in the streams.

401 The regression models developed in this study provide further evidence of the significant
402 contribution of septic systems to human fecal pollution at the watershed level. Combined with
403 land use characteristics such as impervious cover and percent forest cover, the density of septic
404 systems and average distance of septic to stream could explain a significant amount of human
405 fecal inputs, especially in spring. The contribution of septic systems to variations in human fecal
406 pollution in fall and pooled data was also significant, stressing the apparent influence of septic
407 systems on fecal pollution in suburban streams. Water quality indicators such as specific
408 conductance and pH, which can affect microbial survival and persistent, were also important
409 explanatory variables of human fecal pollution in the study area. However, the contribution of
410 septic systems to human fecal pollution in summer was not obvious from correlation and
411 multivariable statistical analysis.

412 **5. Conclusions**

413 • The findings from this study suggest the influence of septic systems, specifically the
414 density of septic systems and average distance of septic to stream, on fecal pollution at
415 the watershed level in areas with >87 septic units/km^2.

416 • Our study also showed that septic systems were more likely to impact water resources in
417 the study area during the spring season which is associated with shallow groundwater
418 table and high baseflow conditions.

419 • Apparently, the density of sewer pipes in these watersheds did not affect the yield and
420 distribution of human-associated *Bacteroidales* marker which makes septic systems the
421 predominant source of human fecal pollution in the study area.

- This study supports the use MST approaches together with traditional FIB monitoring and land use characterization in a tiered approach to isolate the influence of septic systems on water quality in mixed use watersheds.
- Future research should consider monitoring other human-associated markers and pathogens as multiple lines of evidence to elucidate septic system impacts.
- Finally, the results of this study can be used by watershed managers and stakeholders to promote septic system management at the watershed level.

Acknowledgements

This study was funded by USDA NIWQP grant # 2011-05148. We thank the USGS Georgia Water Science Center in Atlanta, Vijayalakshmi Mantripragada, Abha Mundepi, Nahal Hoghooghi and Dr. Noelle Barkley for their diverse contributions to this work.

References

Ahmed, W., Goonetilleke, A., Powell, D. and Gardner, T., 2009. Evaluation of multiple sewage-associated Bacteroides PCR markers for sewage pollution tracking. Water Res. 43(19), 4872-4877.

Amoozegar, A., Hoover, M. T., Kleiss, H. J., Guertal, W. R., & Surbrugg, J. E. (1993). *Evaluation of saprolite for on-site wastewater disposal* (No. PB-93-190635/XAB). North Carolina State Univ., Raleigh, NC (United States). Dept. of Soil Science.

Ballesté, E. and Blanch, A.R., 2010. Persistence of Bacteroides species populations in a river as measured by molecular and culture techniques. Appl. Environ. Microbiol. 76(22), 7608-7616.

Boehm, A.B., Fuhrman, J.A., Mrše, R.D. and Grant, S.B., 2003. Tiered approach for identification of a human fecal pollution source at a recreational beach: case study at Avalon Bay, Catalina Island, California. Environ. Sci. Technol. 37(4), 673-680.

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(18), 6812-6828.

Carey, R.O., Hochmuth, G.J., Martinez, C.J., Boyer, T.H., Nair, V.D., Dukes, M.D., Toor, G.S., Shober, A.L., Cisar, J.L., Trenholm, L.E. and Sartain, J.B., 2012. Regulatory and Resource Management Practices for Urban Watersheds: The Florida Experience. Horttechnology 22(4), 418-429.

Carroll, S., Hargreaves, M. and Goonetilleke, A., 2005. Sourcing faecal pollution from onsite wastewater treatment systems in surface waters using antibiotic resistance analysis. J. Appl. Microbiol. 99(3), 471-482.

Chin, D., Sakura-Lemessy, D., Bosch, D. and Gay, P., 2009. Watershed-scale fate and transport of bacteria. Transactions of the ASABE 52(1), 145-154.

Clarke, J. S., & Peck, M. S. 1991. Ground-water availability in the south metropolitan Atlanta region, Georgia. In Proc. 1991 Georgia Water Resource Conf., pp. 13-16. K. J. Hatcher, ed. Athens, Ga.: University of Georgia, Institute of Ecology..

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(10), 3255-3262.

Edge, T., Hill, S., Seto, P. and Marsalek, J., 2010. Library-dependent and library-independent microbial source tracking to identify spatial variation in faecal contamination sources along a Lake Ontario beach (Ontario, Canada). Water Science and Technology 62(3), 719.

Farnleitner, A.H., Reischer, G.H., Stadler, H., Kollanur, D., Sommer, R., Zerobin, W., Blöschl, G., Barrella, K.M., Truesdale, J.A. and Casarez, E.A., 2011. Agricultural and rural watersheds. Microbial Source Tracking: Methods, Applications, and Case Studies, 399-431.

Field, K.G. and Samadpour, M., 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Res. 41(16), 3517-3538.

GAEPD, 2016. Georgia Environmental Protection Division, Wastewater Permits in Effect, Georgia Environmental Protection Division, accessed April 20, 2016 at: https://epd.georgia.gov/watershed-protection-branch-lists.

Gawler, A.H., Beecher, J.E., Brandão, J., Carroll, N.M., Falcão, L., Gourmelon, M., Masterson, B., Nunes, B., Porter, J. and Rincé, A., 2007. Validation of host-specific Bacteriodales 16S rRNA genes as markers to determine the origin of faecal pollution in Atlantic Rim countries of the European Union. Water Res. 41(16), 3780-3784.

GCBC, 2013. Gwinnett County Board of Commissioners, Gwinnett County GIS Department, accessed January 5, 2014 at: http://gis.gwinnettcounty.com/OnPointWebsite/ WebPages/Map/FundyViewer.aspx.

GCDPU, 2004. Gwinnett County Department of Public Utilities, Gwinnett County GIS Department, accessed December 15, 2015at: G https://www.gwinnettcounty.com/portal/ gwinnett/Departments/InformationTechnologyServices/GeographicInformationSystems/ GISDataBrowser.

Gonzalez, R.A., Conn, K.E., Crosswell, J.R. and Noble, R.T., 2012. Application of empirical predictive modeling using conventional and alternative fecal indicator bacteria in eastern North Carolina waters. Water Res. 46, 5871–5882.

Gourmelon, M., Caprais, M.P., Ségura, R., Le Mennec, C., Lozach, S., Piriou, J.Y. and Rincé, A., 2007. Evaluation of two library-independent microbial source tracking methods to identify sources of fecal contamination in French estuaries. Appl. Environ. Microbiol. 73(15), 4857-4866.

Gregory, L., Blumenthal, B., Wagner, K., Borel, K. and Karthikeyan, R., 2013. Estimating on-site sewage facility density and distribution using geo-spatial analyses. Journal of Natural and Environmental Sciences 4(1), 14-21.

Habteselassie, M.Y., Kirs, M., Conn, K.E., Blackwood, A.D., Kelly, G. and Noble, R.T., 2011. Tracking microbial transport through four onsite wastewater treatment systems to receiving waters in eastern North Carolina. J. Appl. Microbiol. 111(4), 835-847.

Hathaway, J., Hunt, W. and Simmons, O. III, 2010. Statistical evaluation of factors affecting indicator bacteria in urban storm-water runoff. J Environ Eng 136, 1360–1368.

Helsel, D., 2010. Much ado about next to nothing: incorporating nondetects in science. Ann. Occup. Hyg. 54(3), 257-262.

Helsel, D.R., 2005. More than obvious: better methods for interpreting nondetect data. Environ. Sci. Technol. 39(20), 419A-423A.

Humphrey, C. P., O'Driscoll, M. A., and Zarate, M. A., 2011. Evaluation of On-site Wastewater System E. coli Contributions to Shallow Groundwater in Coastal North Carolina. Journal of Water Science and Technology. 63 (4), 789-795.

ITRC (Interstate Technology & Regulatory Council). 2013. Groundwater Statistics and Monitoring Compliance, Statistical Tools for the Project Life Cycle. GSMC-1. Washington, D.C.: Interstate Technology & Regulatory Council, Groundwater Statistics and Monitoring Compliance Team. http://www.itrcweb.org/gsmc-1/.

Jenkins, M.W., Tiwari, S., Lorente, M., Gichaba, C.M. and Wuertz, S., 2009. Identifying human and livestock sources of fecal contamination in Kenya with host-specific *Bacteroidales* assays. Water Res. 43(19), 4956-4966.

Jent, J.R., Ryu, H., Toledo-Hernández, C., Santo Domingo, J.W. and Yeghiazarian, L., 2013. Determining hot spots of fecal contamination in a tropical watershed by combining land-use information and meteorological data with source-specific assays. Environ. Sci. Technol. 47(11), 5794-5802.

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(16), 3701-3715.

Landers, M.N. and Ankcorn, P.D. (2008) Methods to Evaluate Influence of Onsite Septic Wastewater-Treatment Systems on Base Flow in Selected Watersheds in Gwinnett County, Georgia, October 2007, U. S. Geological Survey.

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(6), 4214-4224.

Liang, Z., He, Z., Zhou, X., Powell, C.A., Yang, Y., He, L.M. and Stoffella, P.J., 2013. Impact of mixed land-use practices on the microbial water quality in a subtropical coastal watershed. Sci. Total Environ. 449, 426-433.

Little, C., Soto, D., Lara, A. and Cuevas, J.G., 2008. Nitrogen exports at multiple-scales in a southern Chilean watershed (Patagonian Lakes district). Biogeochemistry 87(3), 297-309.

MNGWPD, 2006. Septic Systems status and issues working paper. Metropolitan North Georgia Water Planning District. Atlanta, Georgia, 37p., Web-only publication accessed March 25, 2016 at: http://www.northgeorgiawater.org/files/district_septic_report_mar2006.pdf.

NLCD, 2006: Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., and Wickham, J., 2011. Completion of the 2006 National Land Cover Database for the Conterminous United States, PE&RS, Vol. 77(9):858-864..

Noble, R.T., Griffith, J.F., Blackwood, A.D., Fuhrman, J.A., Gregory, J.B., Hernandez, X., Liang, X., Bera, A.A. and Schiff, K., 2006. Multitiered approach using quantitative PCR to track sources of fecal pollution affecting Santa Monica Bay, California. Appl. Environ. Microbiol. 72(2), 1604-1612.

Okabe, S., Okayama, N., Savichtcheva, O. and Ito, T., 2007a. Quantification of host-specific Bacteroides–Prevotella 16S rRNA genetic markers for assessment of fecal pollution in freshwater. Appl. Microbiol. Biotechnol. 74(4), 890-901.

Okabe, S. and Shimazu, Y., 2007b. Persistence of host-specific Bacteroides-Prevotella 16S rRNA genetic markers in environmental waters: effects of temperature and salinity. Appl. Microbiol. Biotechnol. 76(4), 935-944.

Peed, L.A., Nietch, C.T., Kelty, C.A., Meckes, M., Mooney, T., Sivaganesan, M. and Shanks, O.C., 2011. Combining land use information and small stream sampling with PCR-based methods for better characterization of diffuse sources of human fecal pollution. Environ. Sci. Technol. 45(13), 5652-5659.

Raith, M. R., Kelty, C. A., Griffith, J. F., Schriewer, A., Wuertz, S., Mieszkin, S. and Holden, P. A., 2013. Comparison of PCR and quantitative real-time PCR methods for the characterization of ruminant and cattle fecal pollution sources. Water Res, 47(18), 6921-6928.

Rantz, S.E., 1982. Measurement and computation of streamflow—Volume 1. Measurement of stage and discharge: U.S. Geological Survey Water-Supply Paper 2175, accessed April 18, 2016 at: http://pubs.usgs.gov/wsp/wsp2175/.

Reischer, G.H., Kasper, D.C., Steinborn, R., Mach, R.L. and Farnleitner, A.H., 2006. Quantitative PCR method for sensitive detection of ruminant fecal pollution in freshwater and evaluation of this method in alpine karstic regions. Appl. Environ. Microbiol. 72(8), 5610-5614.

Sauer, E.P., VandeWalle, J.L., Bootsma, M.J. and McLellan, S.L., 2011. Detection of the human specific Bacteroides genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. Water Res. 45(14), 4081-4091.

Schneeberger, C.L., Charles Humphrey, M., Henry, K., Nancy Deal MS, R. and Seiber, K., Hill, V. R., Zarate-Bermudez, M., 2015. Fate and Transport of Enteric Microbes From Septic Systems in a Coastal Watershed. J. Environ. Health 77(9), 22.

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(2), 249-259.

Shanks, O.C., White, K., Kelty, C.A., Sivaganesan, M., Blannon, J., Meckes, M., Varma, M. and Haugland, R.A., 2010. Performance of PCR-based assays targeting *Bacteroidales* genetic markers of human fecal pollution in sewage and fecal samples. Environ. Sci. Technol. 44(16), 6281-6288.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online at http://websoilsurvey.nrcs.usda.gov/. Accessed 09/12/2016.

Sowah, R., Zhang, H., Radcliffe, D., Bauske, E. and Habteselassie, M.Y., 2014. Evaluating the influence of septic systems and watershed characteristics on stream faecal pollution in suburban watersheds in Georgia, USA. J. Appl. Microbiol. 117(5), 1500-1512.

Stewart, J.R., Boehm, A.B., Dubinsky, E.A., Fong, T.-T., Goodwin, K.D., Griffith, J.F., Noble, R.T., Shanks, O.C., Vijayavel, K. and Weisberg, S.B., 2013. Recommendations following a

multi-laboratory comparison of microbial source tracking methods. Water Res. 47(18), 6829-6838.
Swann, C., 2001. The influence of septic systems at the watershed level. Watershed Protection Techniques 3(4), 821-834.
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(9), 2891-2904.
Tran, N.H., Gin, K.Y.-H. and Ngo, H.H., 2015. Fecal pollution source tracking toolbox for identification, evaluation and characterization of fecal contamination in receiving urban surface waters and groundwater. Sci. Total Environ. 538, 38-57.
U.S. Environmental Protection Agency, 2016. National Summary of Impaired Waters and TMDL Information. Accessed September, 15th 2016 at: https://iaspub.epa.gov/waters10/attains_nation _cy.control? p_report _type=T).
U.S. Environmental Protection Agency, 2013. Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations. Office of Research and Development, Washington, DC 20460.
U.S. Geological Survey, variously dated. National field manual for the collection of water-quality data: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chaps. A1-A9, available online at http://pubs.water.usgs.gov/twri9A.
U.S. Environmental Protection Agency, 2002. Onsite waste-water-treatment systems manual: National Risk Management Research Laboratory Report EPA/625/R-00/008.
U.S. Environmental Protection Agency, 1977. The Report to Congress - Waste Disposal Practices and Their Effects on Groundwater. Washington, DC: U.S. Environmental Protection Agency.
U.S. Geological Survey, 2016. National Water Information System: Web Interface. USGS 02205522 Pew Creek at Patterson Rd, near Lawrenceville, GA (http://waterdata.usgs.gov/ga/nwis/uv/?site_no=02205522); USGS 02207385 Big Haynes Creek at Lenora Road, near Snellville, GA (http://waterdata.usgs.gov/ga/nwis/uv/?site_no=02207385).
van der Wielen, P. W., and Medema, G. (2010). Unsuitability of quantitative Bacteroidales 16S rRNA gene assays for discerning fecal contamination of drinking water. Applied and environmental microbiology, 76, 4876-4881.
Verhougstraete, M.P., Martin, S.L., Kendall, A.D., Hyndman, D.W. and Rose, J.B., 2015. Linking fecal bacteria in rivers to landscape, geochemical, and hydrologic factors and sources at the basin scale. Proceedings of the National Academy of Sciences 112(33), 10419-10424.
Vierheilig, J., Farnleitner, A. H., Kollanur, D., Blöschl, G. and Reischer, G. H., 2012. High abundance of genetic Bacteroidetes markers for total fecal pollution in pristine alpine soils suggests lack in specificity for feces. Journal of Microbiological Methods, 88(3), 433-435.
Walters, S.P. and Field, K.G., 2009. Survival and persistence of human and ruminant-specific faecal *Bacteroidales* in freshwater microcosms. Environ. Microbiol. 11(6), 1410-1421.
Wong, M., Kumar, L., Jenkins, T.M., Xagoraraki, I., Phanikumar, M.S. and Rose, J.B., 2009. Evaluation of public health risks at recreational beaches in Lake Michigan via detection of enteric viruses and a human-specific bacteriological marker. Water Res. 43(4), 1137-1149.

631

632 **Tables and Figures**

633 Figure captions

634

635 Figure 1. Location of the study area with watershed boundaries for low and high density of septic

636 systems and monitoring stations in Gwinnett County, GA

637 Figure 2. Distribution of AllBac, HF183 and BacR markers grouped by septic system density and

638 season. Figures (a), (b) and (c) shows concentrations of AllBac, HF183 and BacR markers

639 respectively, whilst (d), (e) and (f) represents the yield of AllBac, HF183 and BacR markers

640 respectively. The data includes imputed non-detect values and the (*) symbol represents the

641 mean of the observations

642

643

Table 1 – Watershed characteristics grouped by septic system density						
Watershed group	High density watersheds			Low density watersheds		
	Mean	Low	High	Mean	Low	High
Watershed area (km^2)	2.0	0.2	8.8	3.0	0.6	8.4
Slope (%)	7.5	5.7	9.1	8.0	4.6	10.6
Septic density (Septic units/km^2)	216	88	373	22	8	37
Sewer line density (m/km^2)	1298	0	3149	633	0	4119
Impervious cover (%)	18	12	26	7	3	15
Agricultural land use (%)	4	0	12	33	10	49
Forest cover (%)	25	11	44	37	15	51
Average distance of septic to stream (m)	96	55	151	128	86	172

644

645

646

647

648

649

650

651

652

653

Table 2 - Analysis of inhibition in extracted DNA spiked with human-specific marker

Samples	PCR cycle threshold (C_T ± margin of error at 95% confidence level)	
	Undiluted extract	10-fold diluted extracts
Sterile water	28 ± 0.6	
Stream water samples	28 ± 0.2 (p = 0.41*)	28 ± 0.2 (p = 0.35*)

654 * p-value represents statistical significance for comparison of C_T values for diluted and undiluted
655 stream water samples and sterile water

656

Table 3 - ANOVA results showing *p*-values for human marker as impacted by septic density and season

Bacteroidales marker	Parameter	*p*-value (α = 0.05)
Total *Bacteroidales*	Density	0.35
	Season	<0.001
	Density*Season	0.42
Human *Bacteroidales*	Density	0.046
	Season	<0.001
	Density*Season	0.76
Ruminant *Bacteroidales*	Density	0.004
	Season	0.006
	Density*Season	0.002

Table 4 - Spearman rank correlation between *Bacteroidales* markers and standard water quality parameters and land use characteristics

Bacteroidales Marker	Parameter	Spearman correlation coefficient (*r*)			
		Spring	Summer	Fall	Pooled Data
Total *Bacteroidales*	Percent impervious cover	-0.20	0.35	-0.19	-0.12
	Septic density	-0.05	0.41	0.02	0.1
	Percent forest cover	0.16	-0.51*	0.14	-0.01
	Percent agricultural land use	0.27	-0.39	0.11	-0.02
	Av. distance of septic to stream	-0.23	-0.25	-0.19	-0.29
	Sewer line density	-0.04	0.62*	0.13	0.36
	Dissolved oxygen	-0.07	0.44	0.02	-0.23
	pH	0.34	-0.04	0.15	0.15
	Water Temperature	-0.03	0.16	0.31	0.4
	Specific conductance	-0.30	0.51*	0.18	0.21
	E. coli	0.23	0.48*	0.63*	0.65*
	Enterococci	-0.10	0.17	0.6*	0.39
	Human *Bacteroidales*	0.25	0.31	0.77*	0.65*
	Ruminant *Bacteroidales*	0.35	0.52*	0.28	0.34
Human *Bacteroidales*	Percent impervious cover	0.27	0.07	-0.24	0.08
	Septic density	0.39	0.25	-0.11	0.29
	Percent forest cover	-0.57*	-0.02	0.24	-0.18
	Percent agricultural land use	-0.23	-0.12	0.18	-0.11
	Av. distance of septic to stream	-0.64*	-0.27	-0.23	-0.52*
	Sewer line density	-0.13	-0.07	-0.01	-0.09
	Dissolved oxygen	0.11	0.39	-0.05	0.26
	pH	0.04	-0.2	-0.14	0.04
	Water Temperature	0.54*	-0.08	0.34	0.35
	Specific conductance	0.44	0.36	0.24	0.41
	E. coli	0.4	0.79*	0.71*	0.71*
	Enterococci	0.52*	0.54*	0.63*	0.57*
	Ruminant *Bacteroidales*	0.43	0.32	0.38	0.4
Ruminant *Bacteroidales*	Percent impervious cover	0.29	-0.26	-0.75*	-0.34
	Septic density	0.37	-0.21	-0.76*	-0.33
	Percent forest cover	-0.53*	-0.03	0.54*	0.07
	Percent agricultural land use	-0.05	0.3	0.85*	0.57*
	Av. distance of septic to stream	-0.45*	0.08	0.51*	0.21
	Sewer line density	0.32	0.11	-0.41	-0.07
	Dissolved oxygen	0.11	0.23	0.09	0.18
	pH	0.12	0.14	-0.03	0.17
	Water Temperature	0.35	-0.15	0.4	-0.05
	Specific conductance	0.46*	0	-0.46*	-0.11
	E. coli	0.37	0.55*	0.54*	0.52*
	Enterococci	0.04	0.48*	0.37	0.07

671 * Significant at $p \leq 0.05$

672

673

674

675

676

677

Table 5 - Output of multivariable linear regression models for human associated *Bacteroidales* marker yield for seasonal and pooled data

Season	Variable	Parameter estimate	Variance inflation factor (VIF)	*p*-value	Intercept	Adjusted R-square (R_a^2)
Spring	Impervious cover	-0.049	4.3	0.003	6.4	0.74
	Septic density	0.003	5.9	0.006		
	Forest cover	-0.029	2.5	0.0006		
	Dist. to stream[a]	-0.006	1.5	0.006		
	SC[b]	-0.01	3.5	0.018		
Summer	None[c]					
Fall	Impervious cover	-0.029	1.4	0.021	10.46	0.31
	Dist. to stream	-0.007	1.6	0.011		
	Water pH	-0.9	1.4	0.02		
Pooled data	Impervious cover	-0.025	3.2	0.025	4.2	0.47
	Septic density	0.002	1.3	0.017		
	Dist. to stream	-0.003	3.7	0.041		

678 [a] Average distance of septic systems to stream
679 [b] Specific conductance
680 [c] No variables met the significant threshold of $p \leq 0.05$

681

682

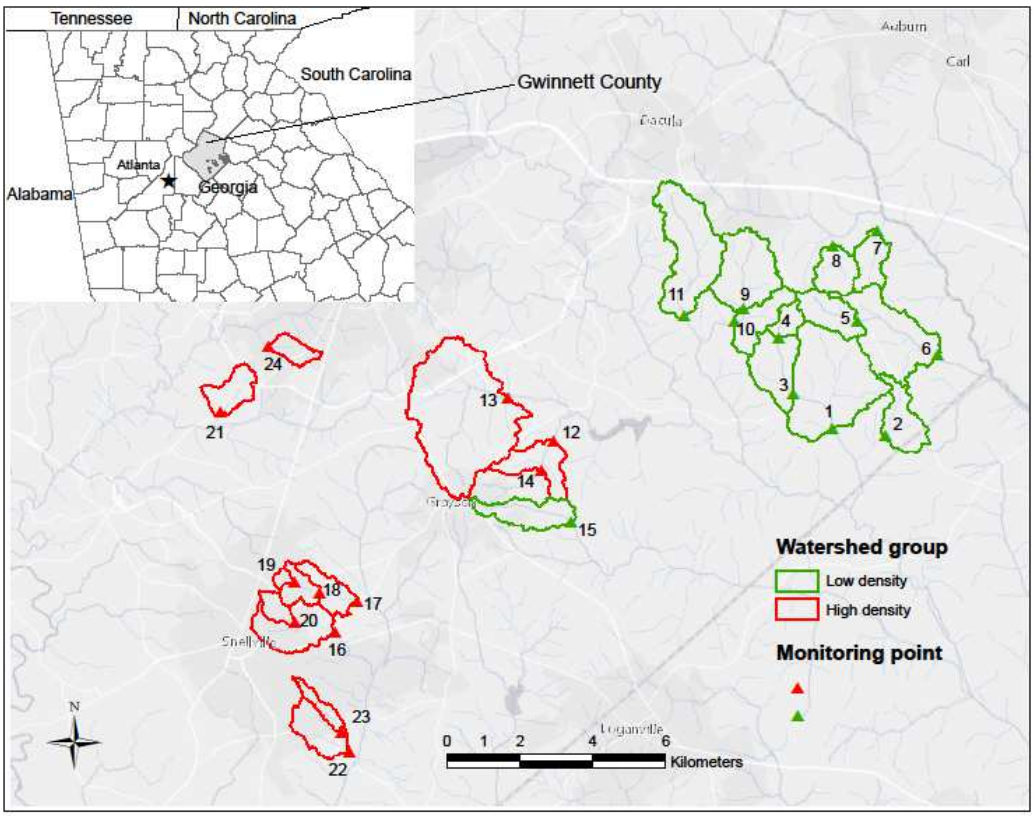
683

684

685

686 Figure 1

687

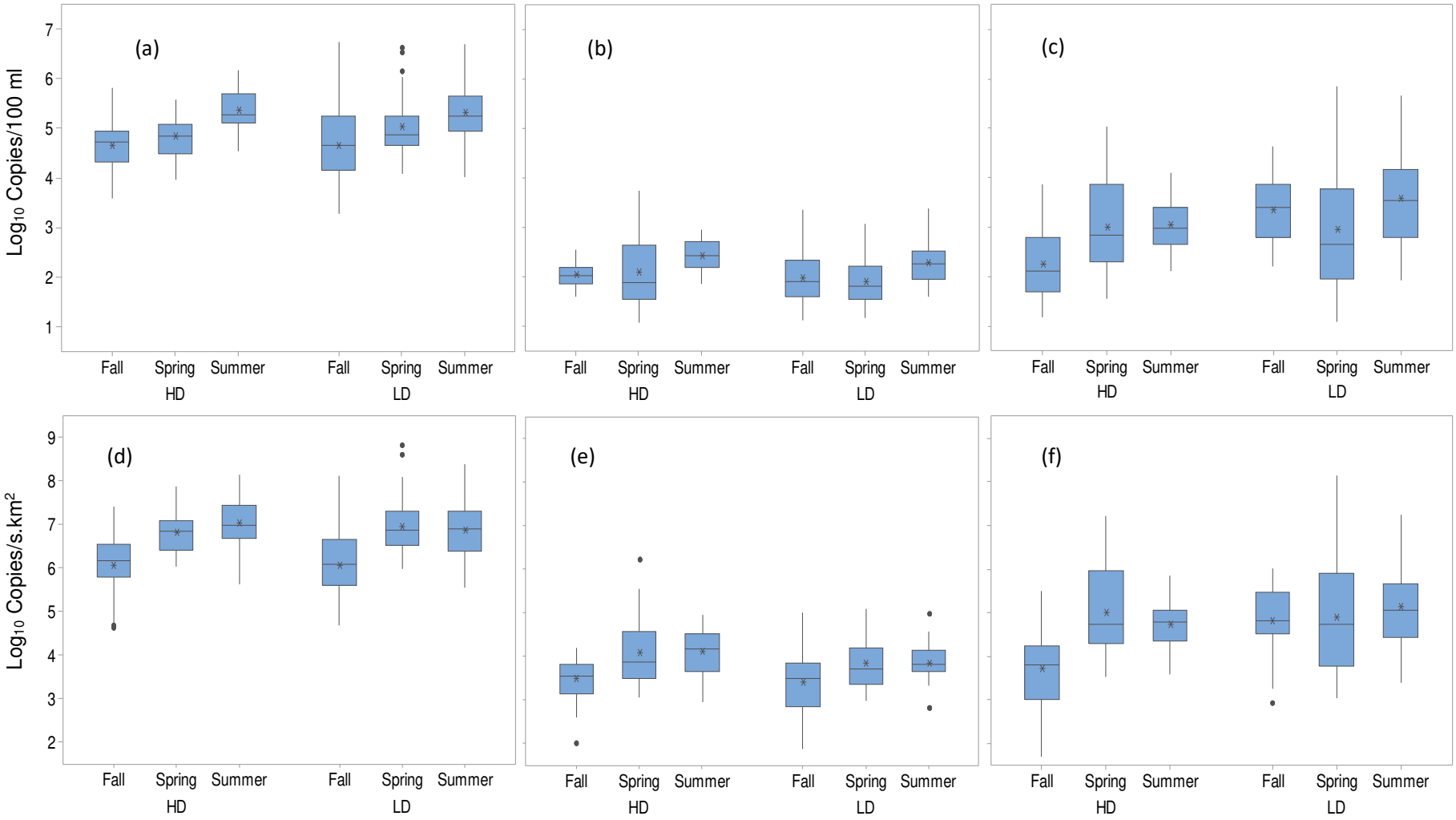


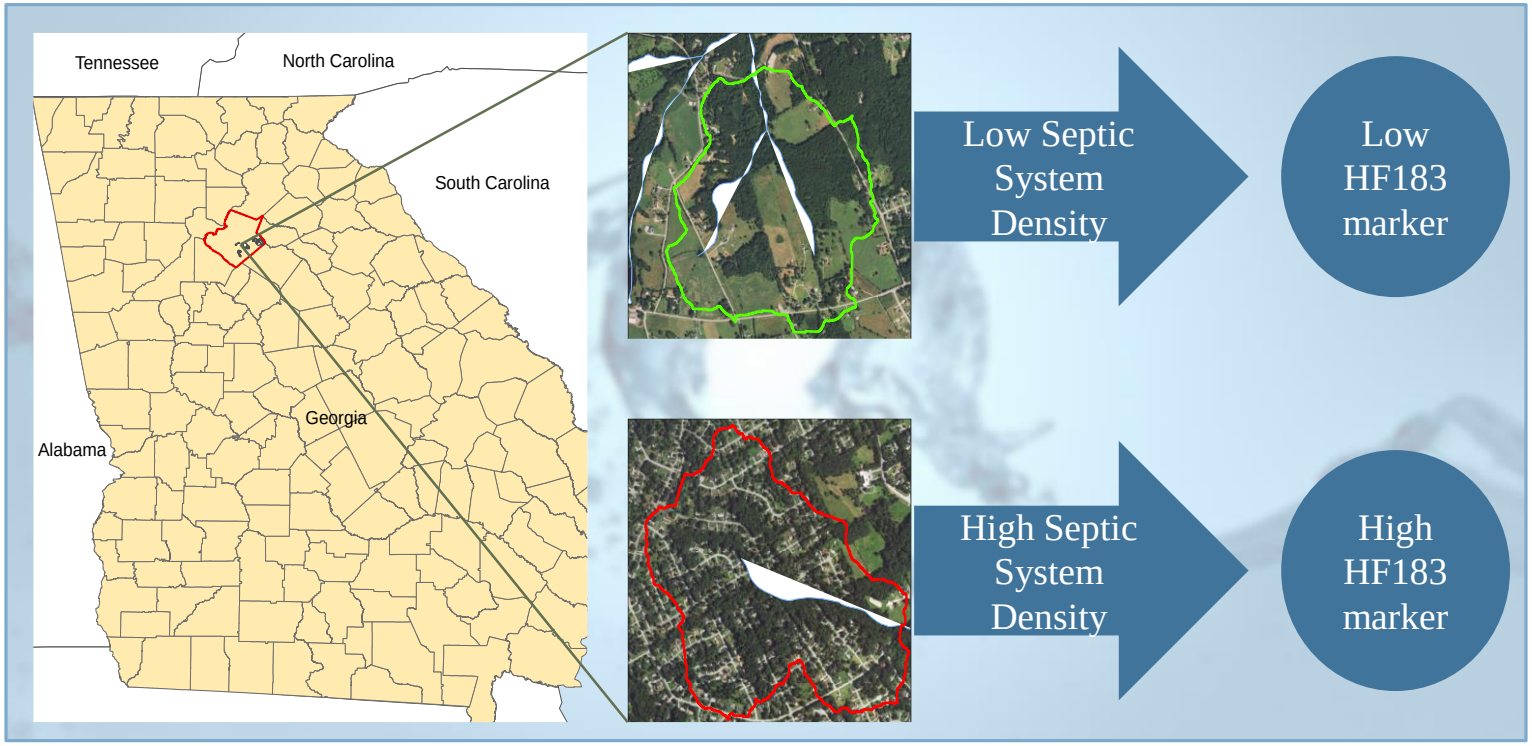
688

689

690 Figure 2.

691




Tennessee
North Carolina
South Carolina
Georgia
Alabama
Low Septic System Density
Low HF183 marker
High Septic System Density
High HF183 marker