

1 **Precipitation thresholds for fecal bacterial indicators in the Chesapeake Bay**

2 A.K. Leight^{1,2*}, R. R. Hood²

3 ¹NOAA National Ocean Service, Cooperative Oxford Laboratory, 904 South Morris Street, Oxford, MD
4 21654

5 ²University of Maryland, Horn Point Laboratory, 2020 Horn Point Road, Cambridge MD 21613

6 *Corresponding author: ak.leight@noaa.gov

7

8 Keywords: fecal coliforms, shellfish closures, indicator bacteria, water quality prediction

9

10 **Abstract**

11 Many coastal states of the United States restrict harvest of shellfish from select areas based on
12 some environmental trigger. Such areas are classified as being conditionally approved. In Maryland, the
13 trigger is an inch or more of rainfall that has fallen in the last 24 hours. This study used 11 years of
14 monitoring data to test the relationship between daily rainfall totals and densities of fecal indicators in
15 Maryland shellfish harvest waters. Precipitation and fecal coliform (FC) water monitoring data from
16 2004 to 2014 were matched by date and watershed. The influence of antecedent rainfall conditions (i.e.
17 rainfall in the preceding days or weeks) and the distance of each monitoring station to land were
18 compared to the percent of samples exceeding the FDA criterion for managing shellfish harvest areas.
19 Sample stations beyond 1000m from land had FC densities consistently below the FDA criterion and
20 were excluded from further analysis. Rainfall events greater than an inch tended to result in
21 significantly elevated FC for the following two days, followed by lower levels thereafter. The total
22 amount of rain in the last three weeks was positively related to the proportion of samples with FC
23 greater than the FDA criterion. Bay-wide, the percent of samples exceeding the FDA criterion rose from

24 seven percent for rainfall less than an inch to 37% following one or more inches of rain. Watersheds
25 were classified based on the percent of FC densities over the criterion when rainfall was an inch or
26 more, with 41 of 81 watersheds showing FC responses indicative of potential conditionally approved
27 areas, those shellfish growing areas where the one inch precipitation trigger may be applied. These
28 areas largely overlapped the current conditionally approved areas defined by Maryland. The percent of
29 open water, wetlands, and poorly drained soils explained a significant amount of the variability ($R^2 =$
30 0.72) in the difference in percent of samples exceeding the FDA criterion when rainfall was greater than
31 an inch and when it was less than an inch. Logistic regression analysis showed that the current trigger of
32 one inch of rain in 24 hours is predictive of FC densities over the FDA criterion, though the appropriate
33 threshold will most likely depend on how far the particular shellfish growing area is from land and
34 antecedent rain conditions. In watersheds with relatively high percentages of open water to total
35 watershed size, higher rainfall thresholds might be appropriate. The approach taken in this study could
36 be applied to individual stations and sub-watersheds, potentially allowing the reclassification of some
37 shellfish harvest areas.

38

39 **1. Introduction**

40 Fecal indicator bacteria in estuarine waters are used as indicators of fecal pollution and an
41 increased risk of encountering human pathogens in nearby shellfish (Ashbolt et al. 2001, FDA 2015). In
42 areas where non-point sources are present or stormwater overflows occur, fecal pollution in estuaries is
43 often related to rainfall (FDA 2015, Kelsey et al. 2004). Thus, in the United States (US), under the Food
44 and Drug Administration (FDA) guidance for management of shellfish harvest (FDA 2015) and
45 Environmental Protection Agency (EPA) guidance for recreational water use (EPA 1986), access to these
46 natural resources may be restricted after a specified level of precipitation. For shellfish, such areas are
47 classified as 'conditionally approved' meaning that the area being open to harvest is conditional upon

48 some set of environmental conditions. Shellfish harvest in many US coastal states is prohibited from
49 conditionally approved beds whenever a specified amount of rain has fallen in the last 24 hours.

50 In Maryland, the classification of conditionally approved shellfish beds is based on field studies
51 that identified conditions generally meeting the FDA criteria except after large rainfall events. The bulk
52 of these studies occurred in the 1980's. The threshold for closure, one inch of rain in the last 24 hours,
53 was established by quantifying fecal coliform (FC) densities in surface waters during and after rain
54 events in select areas of the Bay in 1987. The level of FC in surface waters was measured at several
55 times over several days following rain events (Kathy Brohawn, MDE, personal communication). The
56 resulting management decision was to close conditionally approved areas for three days following a rain
57 event over and inch in 24 hours. Although this process provided evidence of the link between rainfall
58 and FC densities in water and the extent of time that elevated FC densities occurred, it was limited in
59 the number of samples, watersheds, and rain events tested. Therefore, the amount of rain necessary to
60 produce significant runoff may not have been thoroughly investigated, nor were antecedent rainfall
61 conditions (i.e. rainfall that occurred in the days or weeks preceding the current rain event) always taken
62 into account. Additionally, field studies designed to assess the impact of rain events did not quantify
63 related factors, such as land use or soil types, which may affect levels of fecal pollution.

64 Studies in other aquatic systems have assessed the relationship of fecal indicator bacteria, such
65 as FC, in surface waters to a large number of environmental variables and, in some cases, have been
66 able to develop models with moderate capability to predict fecal indicator bacteria densities (EPA
67 2010a, Gonzalez et al. 2012, Kelsey et al. 2010, Maimone et al. 2007, Mallin et al. 2001). Predictive
68 models for fecal indicator bacteria in recreational waters at beaches has been of particular focus (EPA
69 2010a). Variables found to be predictive in previous studies included rainfall, wind velocity, turbidity,
70 water temperature, and riverflow (Campos et al. 2013, EPA 2010a, Kelsey et al. 2010, Maimone et al.
71 2007, Mallin et al. 2001). These studies primarily focused on using linear regression models (Ferguson et

72 al. 1996, Kelsey et al. 2010, Maimone et al. 2007) and decision trees (Maimone et al. 2007) to provide
73 guidance for risk of fecal bacterial densities exceeding established criteria, though some use of logistic
74 regressions has attempted to predict probability of occurrence (Eleria and Vogel 2005). The predictive
75 power of rainfall and/or riverflow (typically a function of rainfall) for fecal indicator densities relates to
76 the land-based source of most fecal bacteria (Kelsey et al. 2004). In general, the ability to predict
77 concentrations of fecal indicator bacteria in natural water bodies with low uncertainty has proved
78 challenging (EPA 2010a, Novotny and Olem 1994). For convenience, the use of a rainfall threshold that
79 is predictive of excessive fecal bacteria may serve as a tool for shellfish managers to make decisions
80 about shellfish bed closures based on the relative risk of having fecal pollution in the growing waters.
81 An assessment of FC densities at California beaches following large storm events underscored the utility
82 of precipitation thresholds (Ackerman and Weisberg 2003).

83 In this study, empirical data was used to examine the relationship between FC densities and
84 precipitation in Maryland's estuarine waters. The null hypotheses were that the level of precipitation
85 necessary to result in FC densities in excess of the FDA criterion is at least an inch and is uniform
86 between various small watersheds across Maryland's portion of the Chesapeake Bay. Supporting
87 hypotheses were that the relationship between precipitation and FC levels was not influenced by the
88 distance of the monitoring stations in each watershed to land, antecedent rain conditions, wind speed,
89 and air temperature (as a proxy for seasonal patterns). The response of FC densities to rainfall was
90 further compared to characteristics of the watershed, such as the percent of open water, impervious
91 surface and soil types. Logistic regressions were used to assess the amount of rainfall resulting in a
92 significant probability of fecal densities exceeding the FDA criterion for management of shellfish harvest
93 areas.

94

95 **2. Materials and Methods**

96 2.1 Meteorological Data

97 Several sources of precipitation data were considered, including daily estimates from National
98 Weather Service weather stations as well as estimated rainfall based on Doppler radar images.
99 Ultimately, rainfall estimates produced by the Middle Atlantic River Forecast Center (MARFC) called
100 Multi-Sensor Precipitation Estimates (MPE) (http://www.weather.gov/marfc/Multisensor_Precipitation)
101 were chosen, primarily due to their use by the Maryland Department of the Environment (MDE) to
102 regulate closures of their conditionally approved shellfish beds and the relatively fine spatial coverage
103 (grid size is approximately 16km²). Some MPE data was excluded from our analysis on the advice of the
104 MARFC (Jason Nolan, MARFC, personal communication). MPE data prior to 2004 was excluded based on
105 a lower level of confidence in the estimates, and data for the months January through March were
106 excluded because the radar precipitation estimates were not as accurate for frozen precipitation and
107 suffer from 'bright banding' - where melting snow registers as large raindrops. The archived MPE data
108 represents 24 hour estimates of total precipitation in inches from 8:00pm to 8:00pm (UTC-5), the same
109 time span used by MDE for conditionally approved shellfish area closures. Antecedent rainfall amounts
110 were calculated by summing previous rainfall amounts for each day up to a week and then by week up
111 to a month prior to the target date.

112 Wind speed and air temperature data were gathered from the Global Historical Climatology
113 Network-Daily (GHCN-D) database through the National Centers for Environmental Information (NCEI)
114 (www.ncei.noaa.gov/, accessed 4/7/2016). Mean daily air temperature and wind speed data were
115 chosen from NOAA weather stations based on their completeness of record and geographic locations.
116 Air temperature data came from weather stations at the Conowingo Dam (USC00182060) and Royal Oak
117 (USC00187806) while wind data came from Baltimore/Washington National Airport (USW00093721) and
118 Salisbury/Wicomico Regional Airport (USW00093720). Temperature and wind data were averaged
119 between stations. To account for sharp changes in air temperature between consecutive days, an

120 average of the air temperature for the day of sampling and the previous day was used. Averaging of
121 weather station spatially and temporally has been conducted routinely by NCEI, such as the monthly
122 mean air temperatures for the climate divisions of the U.S. ([www.ncdc.noaa.gov/monitoring-
123 references/maps/us-climate-divisions.php](http://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-divisions.php)).

124

125 2.2 Station and Watershed Data

126 Water quality stations monitored by the Maryland Department of Environment (MDE) were
127 selected for this study based on the frequency and duration of sampling. All monitoring stations in
128 Maryland's portion of the Chesapeake Bay sampled consistently from 2004 to 2014 were considered, in
129 order to match the time period of the precipitation data. This provided data from 509 stations, with an
130 average sample count of 107 per station and a total of 54,580 observations. MDE typically collects data
131 from these stations twice a month, though data gaps exist, due primarily to winter ice conditions and
132 extreme weather events. Distance from each MDE station to the closest point of land was calculated by
133 plotting the stations in ArcMap (version 10.2.2; ESRI, Inc.) and using the Distance tool and a map of the
134 Chesapeake Bay waterline made by the Chesapeake Bay Program (CBP 2003).

135 United States Geological Survey (USGS) Hydrodynamic Unit Code 12-Digit watersheds (HUC12)
136 were used to compare precipitation and FC levels (<http://water.usgs.gov/GIS/huc>). These watersheds
137 are the smallest of the watershed delineations included in the USGS Watershed Boundary Dataset
138 (<http://nhd.usgs.gov/wbd.html>) and are the closest in size to the current conditionally approved
139 shellfish harvest areas in Maryland (average size approximately 3km²). MDE monitoring stations were
140 associated with the USGS HUC12 watershed in which they fell (ESRI ArcMAP; spatial join tool). These
141 stations were located within 81 HUC12 watersheds, with an average of six stations per watershed. MPE
142 precipitation grids were associated with USGS HUC12 watersheds in which any part of the grid fell. The
143 Chesapeake Bay mainstem (020600010000), Eastern Bay (020600020609), Lower Choptank River

144 mainstem (020600050508), and the Lower Potomac River (020700111001) HUC12 watersheds were
145 excluded because their watersheds consisted of greater than 99% open water and they likely receive
146 riverine and tidal inputs from larger areas than the other watersheds.
147 Land use characteristics and soils data for each watershed were collected from the National Oceanic and
148 Atmospheric Administration (NOAA) Coastal Change Analysis Program (C-CAP)
149 (<https://coast.noaa.gov/digitalcoast/tools/lca>) and the United States Department of Agriculture (USDA)
150 Soil Survey (<http://websoilsurvey.nrcs.usda.gov>), respectively.

151

152 2.3 Data Analysis

153 Precipitation and FC data were matched by HUC12 watershed and date (see Supplemental
154 Figure 1 for example). Median daily precipitation levels were calculated for each watershed and day.
155 The US FDA criterion of 49MPN/100mL for FC was used to assess response to rainfall, where MPN
156 stands for an estimate of bacterial concentration called the most probable number (FDA 2015).
157 Maryland uses the three-tube dilution method for estimating FC densities, which involves inoculation of
158 selective liquid culture media (A1) at three different volumes of sample water (0.1, 1.0, and 10.0 mL)
159 with three tubes per dilution followed by incubation in waterbaths at 44.5°C (APHA 1998). When using
160 this three-tube dilution method, the median FC density estimate must be below 14MPN/100mL, while
161 the calculated 90th percentile estimate must fall below 49MPN/100mL for the area to remain open to
162 harvest.

163 Management of existing conditionally approved areas is slightly different than other areas in
164 that the classification is based on the most recent 15 samples collected during the open status (without
165 rainfall) in comparison to the 49MPN/100mL criterion. The 49MPN/100 mL criterion is used because it
166 is more indicative of conditions resulting from a rain event, such as runoff from land or stormwater
167 overflow caused by a large rainfall (FDA 2015), whereas the median value is more likely to exceed the

168 14MPN/100mL criterion in the absence of an event. If less than 10% of samples collected when rainfall
169 was not present are below 49MPN/100mL, the area remains classified as conditionally approved.
170 Because the focus of this study was to assess responses at the watershed level, the percent of samples
171 over 49MPN/100mL for a watershed was used instead of a station-based percentile, resulting in an
172 average of 677 data points per watershed.

173

174 2.3.1 Frequency Analysis

175 The relationship of rain and FC was first examined by comparing the proportion of samples above the
176 49MPN/100mL FDA criterion with the number of days since the last daily rainfall amount over an inch.
177 The potential effect of antecedent rainfall on the levels of FC resulting from a rain event greater than an
178 inch was tested by comparing the amount of rainfall in the previous one, two, three and four weeks,
179 exclusive of rainfall over the previous two days, to FC levels after a daily rainfall greater than an inch.
180 Based on elevated fecal coliform levels for 48 hours following a rainfall over an inch (see results below)
181 and the random chance of water sampling occurring after a rainfall, a RainEvent was considered to have
182 occurred if there was rainfall over an inch in either of the previous two days (Rain2Day). This criterion
183 was chosen over a cumulative amount of rain for the previous two days because it is in keeping with the
184 current management strategy of closing conditionally approved beds after a 24hr total for rainfall.

185 In order to assess the use of rainfall as an appropriate variable for conditional closure of
186 Maryland shellfish beds, the percent of FC densities over 49MPN/100mL following a RainEvent, both at
187 the bay-wide scale (the entire study area) and by HUC12 watershed were compared. A set of decision
188 points was used to classify HUC12 watersheds based on these frequencies (Supplemental Figure 2). The
189 first decision point was whether or not there were at least three RainEvents sampled. For those
190 watersheds with three or more RainEvents, the next decision point was whether or not the percentage
191 of samples over 49MPN/100mL when a RainEvent had not occurred was less than 10%, a threshold

192 similar to the 10%-of-samples management rule used to classify shellfish growing areas (FDA 2015). For
193 those watersheds with less than 10% FC exceedance of the FDA criteria without RainEvents, those with
194 greater than 10% exceedances after RainEvents were considered responsive watersheds, while those
195 that had less than 10% exceedances after RainEvents were considered not responsive to rainfall.

196

197 2.3.2 Comparison of Frequencies to Watershed Characteristics

198 For those watersheds determined to be responsive to rainfall, the difference in percentage of
199 samples over the 49MPN/100mL criterion when a RainEvent occurred versus the percentage when
200 rainfall was less than an inch (hereafter called PDiff) was compared by Pearson correlation tests to the
201 percent of various land use classifications and the percentage of soil groups within each watershed.
202 Classification and regression tree (CART) analysis was also conducted to identify the most important
203 land use and soil-type variables and corresponding thresholds that best categorize the range of PDiff
204 values. Percent land use was estimated for each of the watersheds by comparing the amount of land
205 area classified by C-CAP (<https://coast.noaa.gov/digitalcoast/tools/lca>) for each of the 21 land cover
206 categories to the total amount of land in each USGS HUC12 watershed. A correlation matrix revealed
207 strong linear relationships ($r > 0.70$) between types of developed land in the watersheds (i.e. high
208 intensity, medium intensity, etc) and between cropland and pasture, so these were combined into total
209 development and total cropland percentages, respectively, before CART analysis. The appropriate
210 number of branches for the tree (pruning) was determined using cross-validation ($K=10$) and by
211 choosing the tree that had the minimum sums of square error (Lemon et al. 2003). The results from the
212 frequency analysis were also used to select watersheds for inclusion in the logistic regression analysis
213 described below.

214

215 2.3.3 Logistic Regression Analysis

216 Logistic regressions were conducted to assess the amount of rainfall (Rain2Day) leading to
217 exceedance of the FDA criterion. Specifically, the 50% probability of a surface water sample to contain
218 FC exceeding 49MPN/100mL was calculated as a function of the greater amount of rainfall for the
219 previous two days (Rain2Day), the total amount of rainfall in the previous three weeks (excluding the
220 previous two days), the distance of each sample from the nearest shoreline, air temperature, and the
221 average wind speed for the previous day. For this analysis, the FC density for each sample was
222 converted to a binary response variable as either above or below the 49MPN/100mL criterion. An
223 analysis of all bay-wide data from within 1000m of land was first conducted. Models were compared
224 using Akaike's Information Criteria (AIC) and the area under the Receiver Operating Curve (ROC) (Hastie
225 et al. 2009). The AIC is a measure of model parsimony, which is the amount of variance explained
226 relative to the number of variables included in the model. The area under the ROC is a measure of
227 correct classification of observations. Smaller AIC and larger ROC values are indicators of more
228 parsimonious and accurate models, respectively. Parameter estimates were assessed using
229 bootstrapping techniques with 1000 iterations of model run (Kelsey et al. 2010). In order to further
230 evaluate the ability of the chosen model to correctly predict probability of exceeding the FDA criteria,
231 data were binned by model predicted probability to the nearest 0.5 percent and compared to the
232 observed frequency of occurrence for the raw data in each bin (Jacobs et al. 2014). The predicted
233 probability and percent occurrence values were then compared using linear regression. Logistic
234 regression analyses using the same predictor variables from the selected bay-wide model were then
235 conducted for each of the watersheds that were considered to be responsive to rainfall based on the
236 decision points used in the frequency analysis. Logistic regression results were reported for watersheds
237 with significant relationships between rainfall and the proportion of samples over 49MPN/100mL
238 ($p < 0.05$), with a percent concordance greater than 75%, and sufficient data to calculate a rainfall
239 threshold with 95% confidence limits.

240 The logistic regression approach was further tested by application to an existing conditionally
241 approved area as characterized by MDE. Although many of the existing conditionally approved areas
242 were established decades ago and included targeted sampling efforts around precipitation events and
243 sampling stations that no longer exist, sufficient data existed to test the logistic regression approach for
244 several of them. One of these is the St. Mary's River. There were only two stations located in or beside
245 the conditionally approved area at the head of the east branch of the St. Mary's River that were
246 consistently monitored from 2004 to 2014 (Supplemental Figure 3). One hundred fifty six samples,
247 including those from 6 RainEvents, were collected over that time span. Because logistic regression
248 results may be biased by such small sample numbers, the Firth bias correction method (Firth 1993) was
249 used to estimate model parameters.

250

251 2.3.4 Analysis Software

252 Datasets were combined and queried using SAS (SAS Institute, Inc., Cary, NC). The frequency
253 analysis was also conducted in SAS. Logistic regressions, including Firth bias correction, and
254 bootstrapped parameter estimates were performed using R (The R Foundation for Statistical Computing,
255 version 3.3.1). Classification and Regression Tree Analysis (CART) were also conducted using R.

256

257 3. Results

258 A decrease in FC levels was noted with increasing distance from land, with all stations beyond
259 1000m having less than 10% of their samples above 49MPN/100mL, the FDA criterion for restricting
260 harvest (Figure 1). A distinct, significant relationship between FC and rainfall was seen when comparing
261 the number of days between a RainEvent and the date of surface water sampling for samples collected
262 within 20 days of a RainEvent (Chisq $p < 0.001$; Figure 2, Panel A). Greater than 30% of all samples within
263 1000m of land contained FC above the 49MPN/100mL criterion when a RainEvent occurred in the

264 previous two days. However, this percentage decreased sharply after two days to less than 20 percent
265 and continued to decrease until it reached a level below 10 percent after five days. Antecedent rainfall,
266 the total amount of rain over the previous three weeks exclusive of the last two days, was also
267 significantly and positively related to FC densities that were detected following a RainEvent
268 (Supplemental Table 1), though the R^2 values indicate antecedent rainfall explained very little of the
269 variation in the FC data. The amount of rainfall within the previous three weeks (Figure 2, Panel B) had
270 the highest R^2 value (0.16) of the time periods considered. The larger percentage of violations that were
271 found when total 3-week rainfall was 9 inches resulted from extremely high FC densities in samples
272 collected directly after Super Storm Sandy in 2011. Wind was associated with increased proportions of
273 samples over the FDA criterion for several speeds over 10mph (Figure 2, Panel C). Average air
274 temperature for the day samples were collected and the previous day had a generally negative
275 relationship with the proportion of samples containing FC over the FDA criterion (Figure 2, Panel D).
276 However, a significant drop in the proportion of samples over the criterion occurred in the 35°F bin
277 relative to the bins on either side of it. This large discrepancy results from earlywinter samples
278 (November and December) with wet and dry winters generally falling into different temperature bins.

279

280 3.1 Frequency Analysis

281 Seven percent of water samples, collected within 1000m of land, contained FC exceeding the
282 49MPN/100mL criterion when rainfall for each of the previous two days was less than one inch. In
283 contrast, 37% of the samples, nearly a six-fold increase, exceeded the criterion after a RainEvent. When
284 analyzed at the watershed (HUC12) scale, 41 watersheds contained less than 10% exceedances of FC
285 over the criteria with rainfall less than an inch and greater than 10% after a rainfall over an inch. These
286 watersheds were classified as being responsive to rainfall (Figure 3). The difference in percent
287 exceedance between samples above and below one inch of rainfall (PDiff) ranged from 3% to 68% for

288 watersheds classified as responsive (Figure 4). Twenty-seven of the 34 conditionally approved shellfish
289 harvest areas, as designated by MDE, fall within these responsive watersheds (Supplemental Table 2).
290 Five of the remaining areas classified by MDE as conditionally approved were located in watersheds
291 where there were insufficient numbers of RainEvents sampled (<3) to make a reasonable assessment.
292 One of the MDE conditionally approved areas, Bolingbroke Creek, resides within watersheds having
293 more than 10% of the samples above the FDA criterion of 49MPN/100mL.

294

295 3.2 Comparison of Frequencies to Watershed Characteristics

296 For each watershed, the difference in the percent of samples with FC densities over the
297 49MPN/100mL criterion when rainfall was less than an inch versus the percent when rainfall was greater
298 than an inch (RainEvent) was not significantly related to the number of RainEvents, the average amount
299 of rainfall for all Rainevents, or the maximum rainfall amount (p value < 0.05). However, PDiff was
300 significantly correlated ($p < 0.05$) to just six land use and soil type characteristics (Table 1). The strongest
301 relationship was a negative correlation to the percent of open water in the watershed, followed by a
302 negative relationship to the percent of unconsolidated shoreline, and a positive relationship to the
303 percent of cultivated crops. Unconsolidated shores are those lacking permanent vegetation and
304 composed of sediment which is prone to inundation and redistribution by waves
305 (<https://coast.noaa.gov/digitalcoast/tools/lca>). Classification and regression tree analysis also showed
306 that the percent of open water was the strongest indicator of how FC densities respond to rainfall and
307 that a percentage of open water for a watershed above 25.8% tended to result in lower levels of FC
308 following RainEvents (Figure 5). For those watersheds with less than 25.8% open water, the next most
309 important variable was a significant negative relationship to the percent of palustrine forested wetland
310 in the watershed. Based on the wetlands classification scheme used by NOAA's C-CAP (NOAA 2016),
311 palustrine forested wetlands are low salinity, non-tidal wetlands primarily consisting of woody

312 vegetation that is at least 5m in height and total vegetation coverage greater than 20 percent. The next
313 most important variable for classification was the percentage of C/D soils in the watershed. These are
314 soils that have slow to very low infiltration rates and relatively high runoff potentials (Supplemental
315 Table 3). Altogether, the percentages of open water, palustrine forested wetland, and C/D soils account
316 for 72% of the variability in the PDiff for samples exceeding the 49MPN/100mL threshold relative to
317 RainEvents. The average distance of the stations in each watershed to land was not correlated to PDiff,
318 suggesting that distance to land is not important in how many samples exceed the criteria, but may be
319 important in the amount of rainfall required to result in FC densities above the FDA criterion.

320

321 3.3 Logistic Regression Analysis

322 Logistic regression analysis of pooled data from all responsive watersheds revealed a predictable
323 increase in the probability of exceeding the FDA criterion with increasing amounts of rainfall, increasing
324 antecedent rainfall, decreasing air temperature, and decreasing distance from land. Analysis of data
325 from all responsive watersheds combined produced a logistic model with a concordance of 76.5% when
326 including all of these independent variables, and a concordance of 60.2% when only including recent
327 rainfall (Rain2Day) in the model (Table 2). Tests for variance inflation indicated that no significant
328 collinearity existed between the predictor variables or between the predictor variables and the intercept
329 (α 0.05, p-value > 0.05). The value for Rain2Day corresponding to a 50% chance of exceeding the
330 49MPN/100mL criterion, under average antecedent rainfall and distance from land, was between 2 and
331 2.2 inches for all models tested. Although the inclusion of air temperature lowered the AIC, it did not
332 significantly increase the area under the receiver response curve. Therefore, the model using only
333 rainfall (Rain2Day and 3wkRainfall), distance from land, and antecedent rainfall was chosen for further
334 testing. For the 50% probability threshold, the percent of false negatives (observations above the FDA
335 criterion predicted to be below the criterion) was low (<10%) while the percent of false positives

336 (observations below the FDA criterion predicted to be above the criterion) was higher (> 10%).
337 However, a plot of the observed proportion of samples over the 49MPN/100mL criterion within each of
338 the predicted probability bins (Figure 6) showed a significant, positive relationship, with some notable,
339 nonlinear scatter for bins close to 50% probability. Bootstrapping showed that there was relatively
340 small error associated with the parameter estimates for this model (Supplemental Table 4).

341 Logistic regression analysis, conducted for each of the responsive watersheds revealed
342 statistically significant models with concordance above 75% and sufficient data to determine 95%
343 confidence limits for 10 of the 41 watersheds (Table 4). The amount of rainfall (Rain2Day)
344 corresponding to a 50% chance of exceeding the FDA criterion, based on average watershed values for
345 all other significant variables, ranged from 1.2 to 2.4 inches (Table 4). These rainfall probability
346 threshold values had a significant, positive relationship to the percentage of open water in the
347 watershed and to highly impervious soils (type D) (Figure 7), but were not related to the time of year
348 that the exceedances occurred ($p > 0.05$).

349 Application of the logistic approach to an existing MDE conditionally approved area in the St.
350 Mary's River produced superior results to the watershed-scale analysis. Testing the probability of
351 exceeding the 49MPN/100mL criterion confirms the findings for this watershed and provides strong
352 evidence for the dominant influence of rainfall on FC densities in this conditionally approved area (Table
353 5). With an odds ratio of 50.9, and no significant variable other than Rain2Day remaining in the model,
354 the relationship between FC densities and rainfall was highlighted.

355

356 **4. Discussion**

357 The findings from this study underscore the importance of rainfall in effecting FC densities in
358 estuarine surface waters and provide evidence that the effect of rainfall on these densities varies
359 between different watersheds. The amount of rainfall in the last 2 days (Rain2Day) provided a

360 convenient and useful metric for separating watersheds based on the frequency of encountering high
361 densities of FC and was consistently more important (highest odds ratio) than antecedent rainfall, wind
362 speed, and air temperature in determining the probability of high FC densities. These findings both
363 confirm the importance of rainfall in surface water conditions and supports the use of rainfall as a
364 conditional closure tool in select areas. The drop in FC densities after two days post RainEvent also
365 suggests the current management strategy of closing conditionally approved areas for a defined number
366 of days is protective from a human health risk perspective. However, the proportion of samples below
367 the FDA criterion did not fall below 10% for five days after a RainEvent. If all of the stations less than
368 1000m from land were part of conditionally approved areas then there would likely be some times
369 and/or places when the three day closure rule was not fully protective. Any establishment of a new
370 conditionally approved area would need to assess the rate of FC dilution and decay for that particular
371 area.

372 The frequency analysis also provided key insights into the influence of land-based conditions on
373 FC densities in shellfish harvest waters. Other studies (DiDonato et al. 2009, Glasoe and Christy 2004,
374 Mallin et al. 2001) have found statistically significant relationships between impervious surfaces and/or
375 urban development on indicator bacteria densities in coastal systems. These land use types may help
376 explain the excessive number of samples over the FDA criteria regardless of rain in the heavily urbanized
377 watershed draining Baltimore City and two agricultural watersheds on the Eastern Shore. However, this
378 study found that the frequency of FC densities exceeding the FDA criterion was highly variable within
379 region and had more connection to characteristics of the water body (percent open water) and land
380 margin (percentages of palustrine forested wetland) than to dominant upland characteristics. The
381 percent of soils with moderate to high runoff potential (C/D soil) and the percent of cultivated crops
382 were also significantly positively related to the frequency of exceedances (PDiff), but explained less
383 variability than the percentage of open water. The strong relationship to open water, despite the

384 exclusion of stations greater than 1000m from land, indicates dilution of FC-laden runoff in tidal systems
385 with relatively large proportions of water to land. This finding is supported by the Total Maximum Daily
386 Load (TMDL) modeling efforts of MDE which show that inputs of FC from land are often diluted to very
387 low levels within several kilometers of where they are introduced (MDE 2006). More importantly, the
388 significant relationship to open water implies that watersheds with small water-to-land ratios are more
389 prone to rainfall-driven FC densities in exceedance of the FDA criterion. The positive correlation to
390 percentage of C/D soils in the watershed suggests that soils with decreased infiltration rates and higher
391 runoff potential may increase runoff introduction of FC following rain events.

392 The significant negative relationship of FC to certain types of wetlands might be expected as
393 both natural and constructed wetlands have been shown to reduce or retain fecal bacteria and
394 pathogens (Green et al. 1997, Knox et al. 2008). However, others (Grant et al. 2001, Huang 2010) have
395 detected a significant positive relationship between wetlands and numbers of indicator bacteria in
396 estuarine waters. Huang (2010) hypothesizes this positive relationship results from the use of wetlands
397 by warm-blooded animals, such as geese and deer. One of the watersheds in this study, the Monie Bay
398 watershed on the lower eastern shore, had high FC densities regardless of rainfall in extensive estuarine
399 wetlands (18% estuarine emergent wetlands, <https://coast.noaa.gov/digitalcoast/tools/lca>) indicating
400 that types of wetlands, based on salinity levels and the kind of vegetation, might be important in their
401 relationship to fecal bacteria in adjacent waters.

402 The results of the logistic regression analysis for individual watersheds lead to several important
403 conclusions. Foremost, FC densities in exceedance of the FDA criterion in these watersheds are strongly
404 related to rainfall. Secondly, the level of rainfall required to produce exceedances of the FDA criterion
405 may vary between watersheds. In part, this variability is related to the average distance of the sampling
406 stations to land. However, the differences between odds ratios (the odds that an increase in 2DayRain
407 will result in FC concentration of 49MPN/100mL or greater) for Rain2Day between the watersheds

408 suggest different degrees of influence of rainfall on FC densities in these systems. The findings of this
409 study also indicate that the level of rainfall necessary to produce FC densities exceeding the FDA criteria
410 is dependent on a handful of variables, including the distance of the sample stations from land and the
411 proportion of open water that makes up the watershed. These findings agree with those from the
412 frequency analysis, with percentages of open water and soil type influencing not only the frequency of
413 encountering high levels of FC but also the amount of rainfall needed to produce those high levels.

414 Logistic models proved capable of estimating the relative importance of environmental variables
415 on FC densities in surface waters and in estimating the level of rainfall associated with a 50% chance or
416 greater of exceeding the FDA criteria. Although recent rainfall and distance to land had the strongest
417 relationship to probability of exceedance, antecedent rainfall and air temperature also showed
418 significant influence. The negative relationship to air temperature may indicate a seasonal influence
419 with higher probabilities of exceeding the FDA criterion in cooler times of the year. A previous study of
420 FC data in Maryland found an increase in FDA exceedances late in the year (October through
421 December), with a lag in time between peak precipitation and FC levels (Leight et al. 2016). One
422 potential reason for elevated numbers of exceedances in cool time periods is an increase in the amount
423 of rainfall that leads directly to runoff when trees lose their leaves and ground cover has died back. This
424 phenomenon is a known factor in Chesapeake Bay runoff, such that it is included in the Chesapeake Bay
425 Watershed model used by EPA and relevant states to set criteria for pollution (EPA 2010b). Other
426 possibilities include an increase in larger rain events and therefore greater soil saturation from late year
427 tropical storms leading to increased runoff of FC bacteria, a decrease in deactivation of bacterial cells by
428 ultraviolet radiation, and an increase in the presence of migratory waterfowl, particularly Canada geese
429 (*Branta canadensis*).

430 By selecting those watersheds that were responsive to rainfall, based on the frequency analysis,
431 the logistic regression was targeted at watersheds where exceedances related to rainfall would tip the

432 balance towards violation of the FDA criterion. Both the predictive capability of the logistic model seen
433 in the observed and predicted probabilities, and the odds ratios for Rain2Day support the importance of
434 rainfall driven violations in the watersheds selected. Application of the logistic approach to a defined
435 growing area in the St Mary's River produced low percentages of false negatives and false positives, a
436 very high odds ratio, and very high concordance values, despite the modest number of samples included
437 in the analysis.

438 In development of our logistic regression models, one important consideration was the FC
439 density used for assessing violations. FC densities in surface waters after rain events may peak at levels
440 well above 49MPN/100mL. Increasing the FC concentration considered excessive to some value above
441 49MPN/100mL would increase the concordance and lower the rate of false positives in the logistic
442 regression. However, the 90th percentile criterion used here was chosen based on the existing
443 management structure and understanding of risk for illnesses. Additionally, picking a higher probability
444 of exceedance would lower the percent of false positives, but would increase the percent of false
445 negatives and, therefore, run the risk of not being as protective of human health.

446 The co-location of most areas currently classified as conditionally approved by MDE with the
447 watersheds defined in this study as being responsive to rainfall supports the classification of these areas
448 and the general use of the conditionally approved classification. The frequency analysis also suggests
449 that FC densities in many of the watersheds in Maryland's portion of the Chesapeake Bay are responsive
450 to rainfall such that they do not exceed the 90th percentile over 49MPN/100mL FDA criterion without
451 significant rain events. This finding suggests that targeted sampling would likely uncover new areas
452 suitable for being classified as conditionally approved. Focused analysis, such as that done for St. Mary's
453 River in this study, may be used to target assessments for classification. Any conversion of restricted
454 areas to conditionally approved would have important economic and sociopolitical ramifications for

455 both shellfish harvesters and shellfish aquaculturists. Further, any increases in the precipitation
456 threshold would decrease the number of days that shellfish beds are closed.

457 Employing logistic regressions, such as those developed in this study, could provide important
458 information when considering such management decisions. While distance to land has been shown
459 previously to be an important driver of FC densities (MDE 2006), the inclusion of this variable in a
460 regression model might provide an informative way to make decisions about conditional closure areas
461 using empirical data. Results from the logistic regression analysis underscore the importance of how far
462 the sampling location is from land, with decreasing influence of rainfall on FC densities as distance to
463 land increases. By using the average conditions for distance (for each watershed), antecedent rainfall,
464 and wind speed, the probability values calculated by logistic regression support the use of a one inch
465 rainfall threshold. Models including not only measurements of rainfall in the last 24 hours, but also air
466 temperature and antecedent rainfall could be used to better estimate the current probability of
467 exceeding the threshold.

468 The principal limitation in conducting this study was the relatively modest number of samples
469 collected soon after rain events. Several watersheds with areas currently classified as conditionally
470 approved lacked sufficient rain event data to include in the analyses. With increasing amounts of data,
471 as the MDE monitoring program continues, additional data from samples collected within 48 hours of
472 rain events should improve estimates of rainfall triggers and the appropriate areas for their application.
473 Further, our analysis excluded information from January through March based on the advice of the
474 Middle Atlantic River Forecast office, which produces the rainfall product. Although exclusion of winter
475 precipitation coincides with a large window of oyster harvest from natural beds in Maryland, earlier
476 research (Chapter 2) showed that both precipitation amounts and FC densities in January through March
477 tend to be lower than the rest of the year. One management strategy would be to exclude any frozen
478 precipitation as a trigger for closing conditionally approved shellfish areas. Another important

479 consideration for maintaining conditionally approved areas and understanding rainfall triggers is that
480 other shellfish collections (principally clams) and oysters grown by the burgeoning aquaculture industry
481 in Maryland are permitted year round. This importance is underscored by the current MDE strategy
482 which uses numbers of FC from year-round sampling to make management decisions about shellfish
483 harvest closures.

484

485 **5. Conclusions**

- 486 • Rainfall plays a dominant role in the introduction of fecal coliforms to Maryland shellfish waters
- 487 • Distance to land and antecedent rainfall correspond to the densities of fecal coliforms following
488 rain events
- 489 • The response of fecal coliform densities to rainfall differs between watersheds and is related to
490 characteristics such as the percent of open water, the percent and types of wetlands, and the
491 percent of soils with moderate to high runoff potential
- 492 • The level of rainfall predictive of a 50% or greater chance of fecal coliforms exceeding the
493 closure criterion is typically more than an inch but may vary by watershed

494

495 **Conflict of Interest**

496 The authors declare no conflict of interest.

497

498 **Acknowledgements**

499 All sample FC data was provided by the Maryland Department of the Environment, with special thanks
500 to Kathy Brohawn, Heather Merritt, and Keeve Brine. Bill Beatty, Billy Evans, Rusty McKay, and Quentin
501 Forrest were among the many individuals from MDE responsible for collection of samples. Jason Nolan,

502 at NOAA's Middle Atlantic River Forecast Center processed historic MPE data and provided daily
503 estimates of precipitation. Raleigh Hood was supported by grants from NSF, NASA, and NOAA. This is
504 contribution 5448 from the University of Maryland Center for Environmental Science.

505

506 **Disclaimer**

507 The scientific results and conclusions, as well as any opinions expressed herein, are those of the authors
508 and do not necessarily reflect the views of NOAA or the Department of Commerce. The mention of any
509 commercial product is not meant as an endorsement by the Agency or Department.

510

511 **Literature Cited**

512 APHA (1998) Standard Methods for the Examination of Water and Wastewater, 20th Edition, American
513 Public Health Association, Washington, DC.

514
515 Ashbolt, N.J., Grabow, W.O.K. and Snozzi, M. (2001) Indicators of microbial water quality. In: Water
516 Quality: Guidelines, Standards and Health. Fewtrell, L. and Bartram, J. (eds), pp. 289-315, IWA
517 Publishing, London, UK.

518
519 Campos, C.J.A., Kershaw, S.R. and Lee, R.J. (2013) Environmental Influences on Faecal Indicator
520 Organisms in Coastal Waters and Their Accumulation in Bivalve Shellfish. *Estuar. Coast.* 36(4), 834-853.

521
522 CBP (2003) Chesapeake Bay Segments, The Chesapeake Bay Program.

523
524 DiDonato, G.T., Stewart, J.R., Sanger, D.M., Robinson, B.J., Thompson, B.C., Holland, A.F. and Van Dolah,
525 R.F. (2009) Effects of changing land use on the microbial water quality of tidal creeks. *Mar. Pollut. Bull.*
526 58(1), 97-106.

527
528 Eleria, A. and Vogel, R.M. (2005) Predicting fecal coliform bacteria levels in the Charles River,
529 Massachusetts, USA. *J. Am. Water Res. Assoc.* 41(5), 1195-1209.

530
531 EPA (1986) Ambient water quality criteria for bacteria - 1986, p. 18, Environmental Protection Agency.

532
533 EPA (2010a) Predictive Tools for Beach Notification Volume I: Review and Technical Protocol, p. 71, US
534 Environmental Protection Agency.

535
536 EPA (2010b) Chesapeake Bay Phase 5.3 Community Watershed Model, U.S. Environmental Protection
537 Agency, Chesapeake Bay Program Office, Annapolis, Maryland.

538
539 FDA (2015) National Shellfish Sanitation Program (NSSP) Guide for the Control of Molluscan Shellfish:
540 2015 Revision. 438.

541
542 Ferguson, C.M., Coote, B.G., Ashbolt, N.J. and Stevenson, I.M. (1996) Relationships between indicators,
543 pathogens and water quality in an estuarine system. *Water Res.* 30(9), 2045-2054.

544
545 Firth, D. (1993) Bias reduction of maximum-likelihood estimates. *Biometrika* 80(1), 27-38.

546

547 Glasoe, S. and Christy, A. (2004) Coastal urbanization and microbial contamination of shellfish growing
548 areas - literature review and analysis, p. 29, Puget Sound Action Team, State of Washington.

549
550 Gonzalez, R.A., Conn, K.E., Crosswell, J.R. and Noble, R.T. (2012) Application of empirical predictive
551 modeling using conventional and alternative fecal indicator bacteria in eastern North Carolina waters.
552 Water Res. 46(18), 5871-5882.

553
554 Grant, S.B., Sanders, B.F., Boehm, A.B., Redman, J.A., Kim, J.H., Mrse, R.D., Chu, A.K., Gouldin, M.,
555 McGee, C.D., Gardiner, N.A., Jones, B.H., Svejkovsky, J. and Leipzig, G.V. (2001) Generation of
556 enterococci bacteria in a coastal saltwater marsh and its impact on surf zone water quality. Environ. Sci.
557 Technol. 35(12), 2407-2416.

558
559 Green, M.B., Griffin, P., Seabridge, J.K. and Dhubie, D. (1997) Removal of bacteria in subsurface flow
560 wetlands. Water Sci. Technol. 35(5), 109-116.

561
562 Hastie, T., Tibshirani, R. and Friedman, J. (2009) The Elements of Statistical Learning: Data Mining,
563 Inference, and Prediction, Springer, New York.

564
565 Huang, J. (2010) Spatial and temporal analysis of fecal coliform distribution in Virginia coastal waters.
566 Dissertation, College of William and Mary.

567
568 Jacobs, J.M., Rhodes, M., Brown, C.W., Hood, R.R., Leight, A., Long, W. and Wood, R. (2014) Modeling
569 and forecasting the distribution of *Vibrio vulnificus* in Chesapeake Bay. J. Appl. Microbiol. 117(5), 1312-
570 1327.

571
572 Kelsey, H., Porter, D.E., Scott, G., Neet, M. and White, D. (2004) Using geographic information systems
573 and regression analysis to evaluate relationships between land use and fecal coliform bacterial pollution.
574 J. Exp. Mar. Biol. Ecol. 298(2), 197-209.

575
576 Kelsey, R.H., Scott, G.I., Porter, D.E., Siewicki, T.C. and Edwards, D.G. (2010) Improvements to shellfish
577 harvest area closure decision making using GIS, remote sensing, and predictive models. Estuar. Coast.
578 33(3), 712-722.

579
580 Knox, A.K., Dahlgren, R.A., Tate, K.W. and Atwill, E.R. (2008) Efficacy of natural wetlands to retain
581 nutrient, sediment and microbial pollutants. J. Environ. Qual. 37(5), 1837-1846.

582
583 Leight, A.K., Hood, R., Wood, R. and Brohawn, K. (2016) Climate relationships to fecal bacterial densities
584 in Maryland shellfish harvest waters. Water Res. 89, 270-281.

585
586 Lemon, S.C., Roy, J., Clark, M.A., Friedmann, P.D. and Rakowski, W. (2003) Classification and regression
587 tree analysis in public health: methodological review and comparison with logistic regression. *Annals*
588 *Behav. Med.* 26(3), 172-181.

589
590 Maimone, M., Crockett, C.S. and Cesanek, W.E. (2007) PhillyRiverCast: A real-time bacteria forecasting
591 model and web application for the Schuylkill River. *J. Water Res. Plan. Manag.* 133(6), 542-549.

592
593 Mallin, M.A., Ensign, S.H., McIver, M.R., Shank, G.C. and Fowler, P.K. (2001) Demographic, landscape,
594 and meteorological factors controlling the microbial pollution of coastal waters. *Hydrobiologia* 460, 185-
595 193.

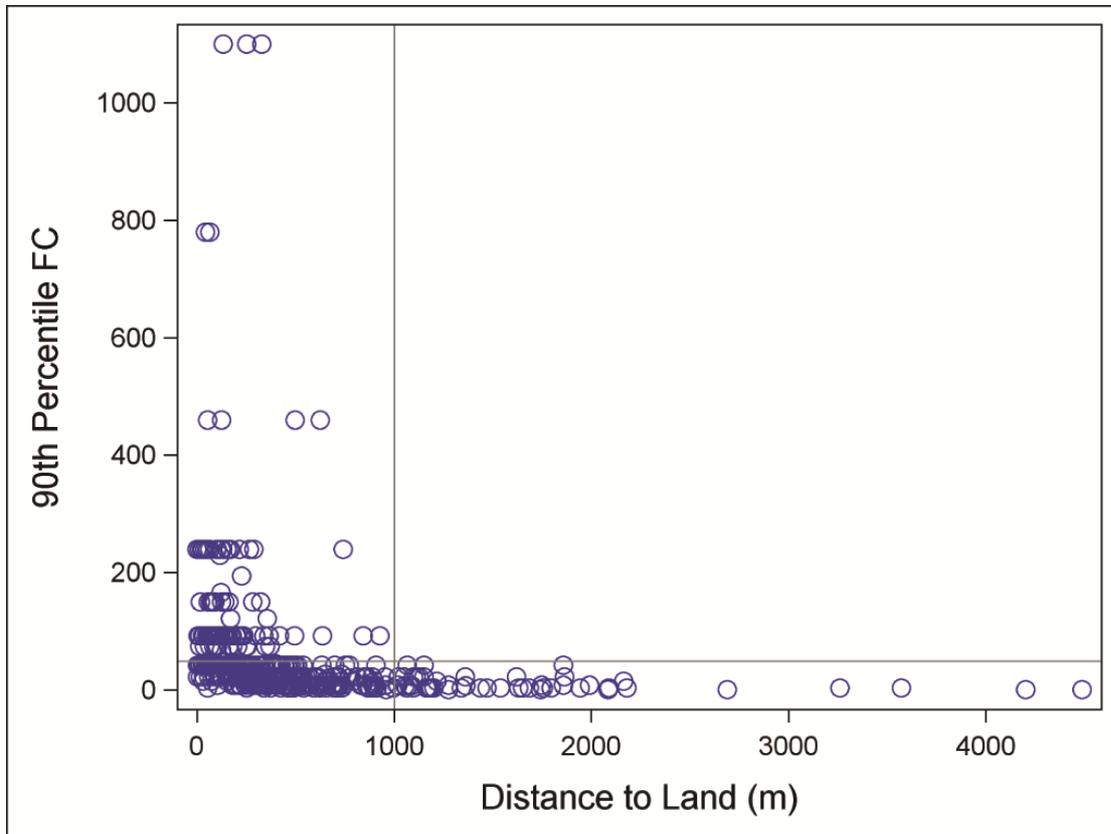
596
597 MDE (2006) Total maximum daily loads of fecal coliform for the restricted shellfish harvest area in the
598 lower Choptank River mainstem in Dorchester and Talbot Counties, Maryland, Maryland Department of
599 the Environment.

600
601 NOAA (2016) <https://coast.noaa.gov/data/digitalcoast/pdf/ccap-class-scheme-regional.pdf> 8/8/16.

602
603 Novotny, V. and Olem, H. (1994) *Water Quality: Prevention, Identification, and Management of Diffuse*
604 *Pollution*, Van Nostrand Reinhold, New York.

605
606
607

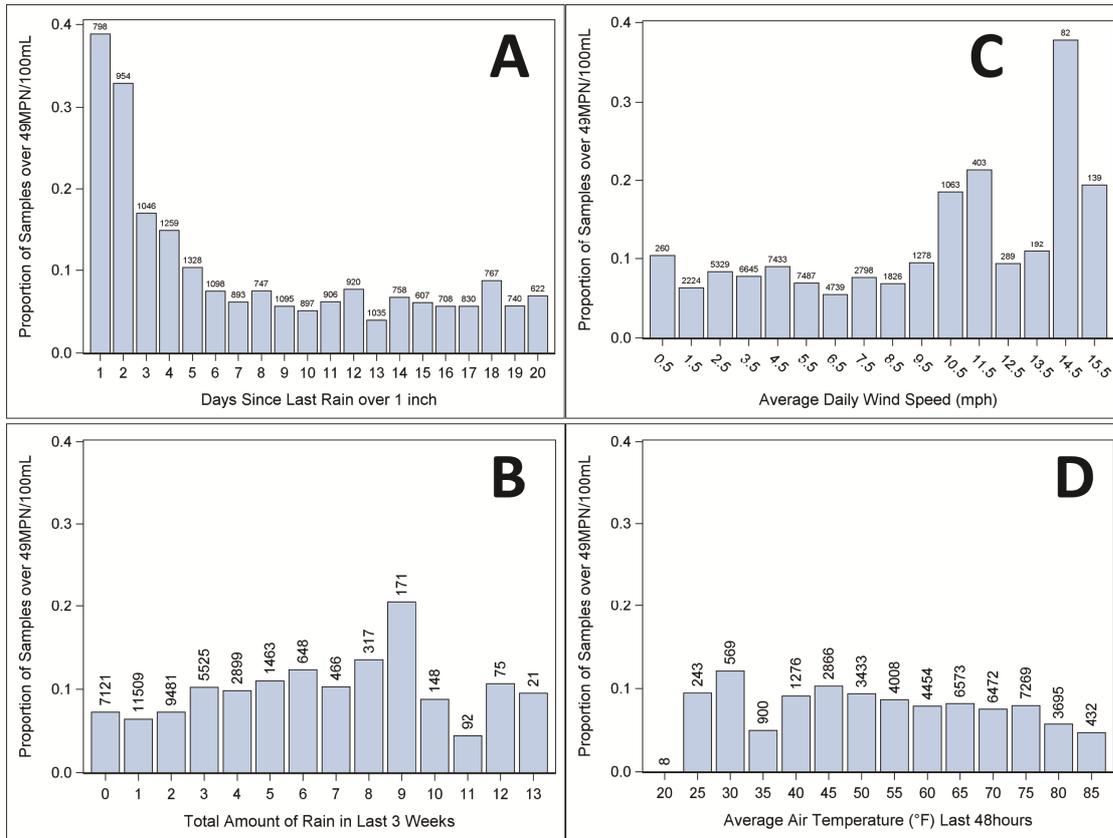
608



609

610 **Figure 1.** Comparison of the 90th percentile of fecal coliform values (MPN/100mL) for each monitoring

611 station with distance of that station to the nearest point of land.



612

613 **Figure 2.** Comparison of the proportion of samples with FC over 49MPN/100mL to the number of days
 614 since last rain over an inch (Panel A), the total amount of rainfall in the previous month (Panel B),
 615 average daily wind speed for the preceding day (Panel C), and average air temperature over the last 48
 616 hours (Panel D). Numbers above bars indicate how many samples were present in each bin.

617

618

619

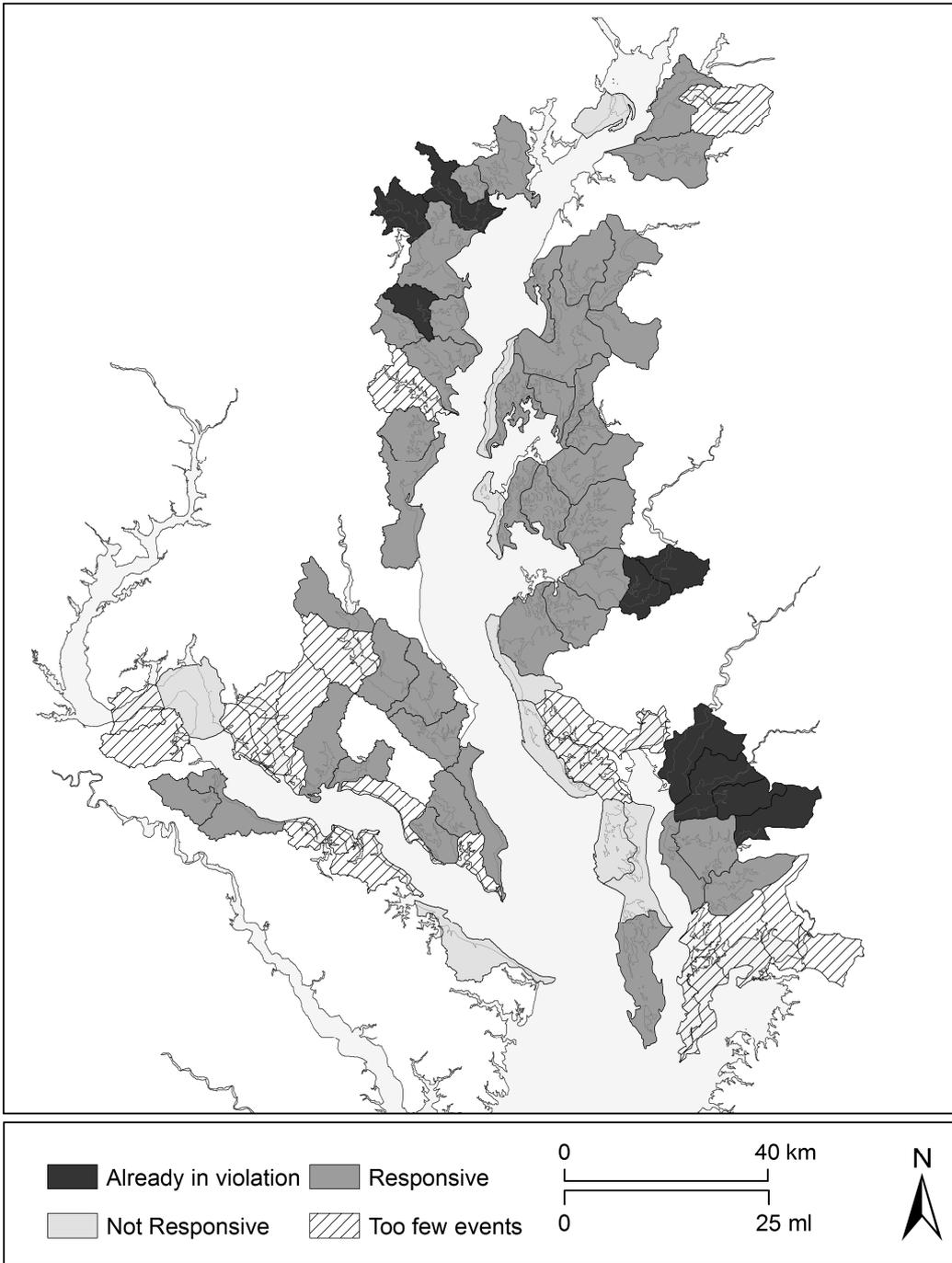
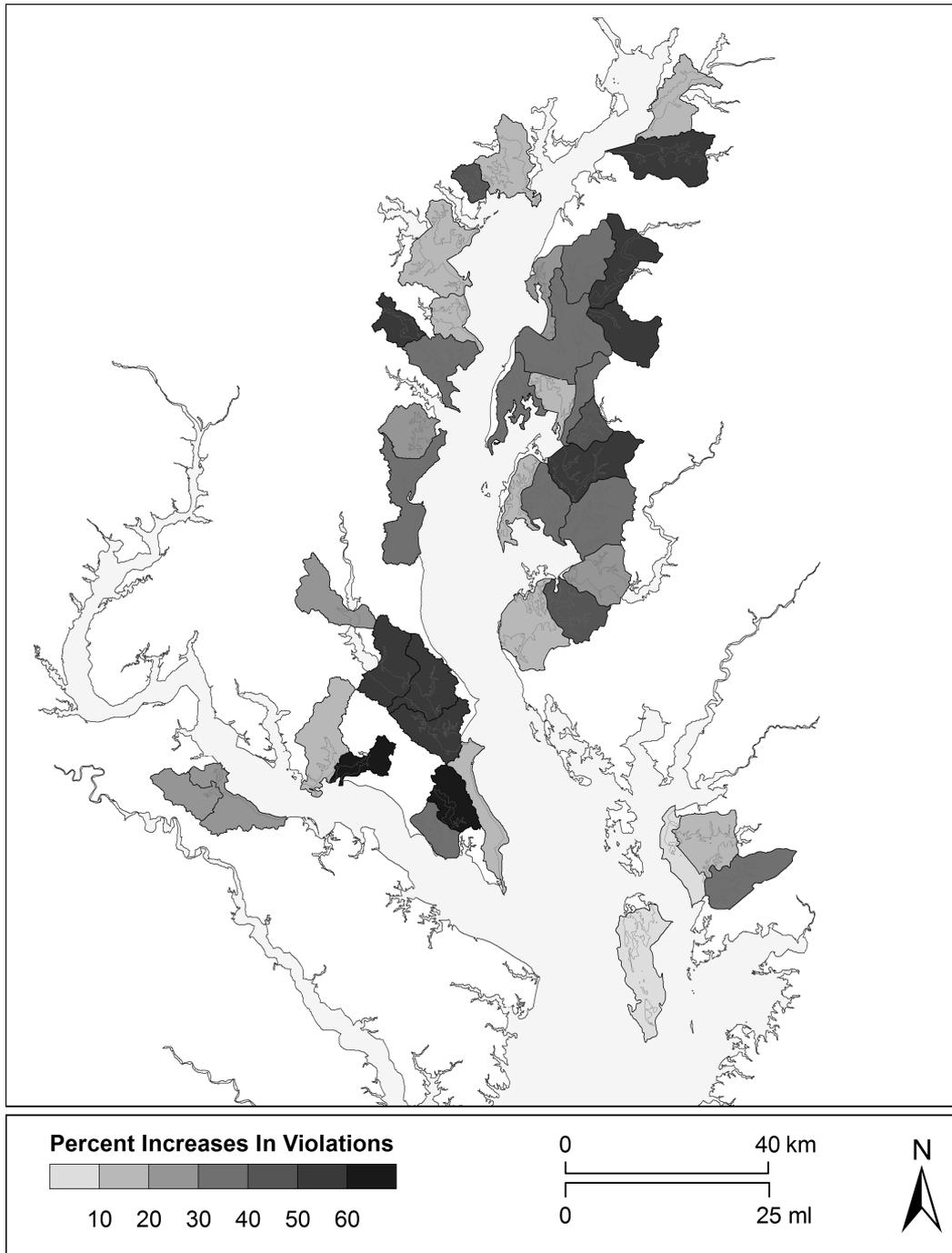


Figure 3. Classification of watersheds based on the change in fecal coliform densities relative to rain events of 1+ inches of rain (Rain2Day). Decision points for the classifications are provided in the methods and displayed in Supplemental Figure 2.



625

626 **Figure 4.** For those watersheds classified as responsive to rain (based on decision points in Figure 2), the
 627 percent increase in fecal coliform densities exceeding the FDA 49MPN/100mL criterion after a rain event
 628 of 1+ inches of rain (Rain2Day) from the percent occurring when rainfall was less than 1 inch.

629

630 **Table 1.** Pearson correlation coefficients and p-values associated with comparison PDiff (the difference
631 in percentage of samples over the 49MPN/100mL criterion when a RainEvent occurred versus the
632 percentage when rainfall was less than an inch) to the percent of each land use and soil classification
633 found in each watershed. Significant correlations ($p < 0.05$) are in bold face type.

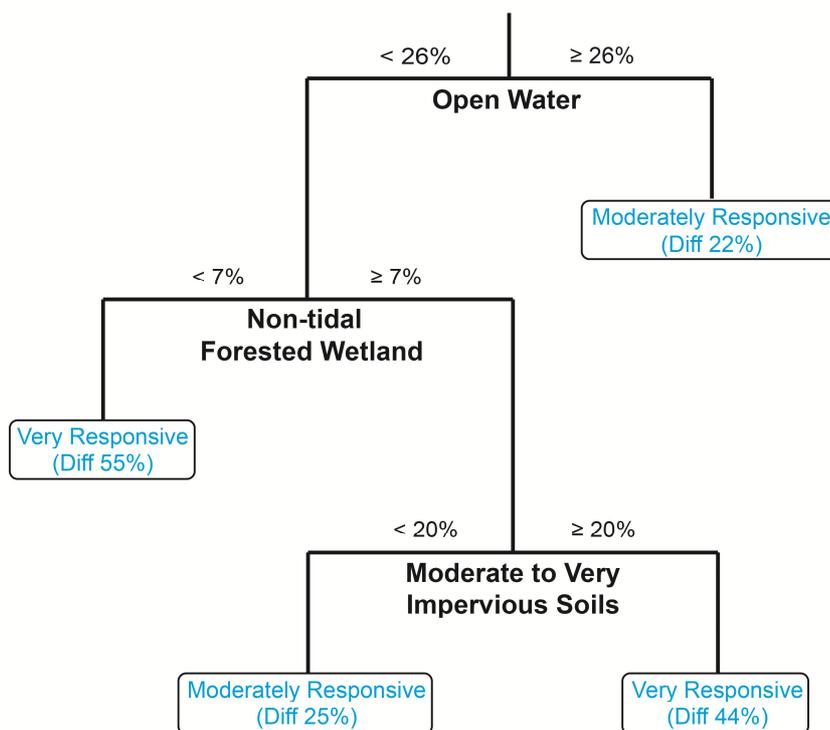
Variable	R ²	p-value
Cultivated Crops	0.370	0.017
Pasture Hay	0.339	0.030
Deciduous Forest	0.316	0.044
Palustrine Aquatic Bed	0.281	0.075
Developed - Open Space	0.234	0.140
Scrub Shrub	0.218	0.171
Type C Soils	0.208	0.192
Mixed Forest	0.199	0.213
Developed - High Intensity	0.190	0.234
Type B Soils	0.185	0.248
Developed - Medium Intensity	0.171	0.286
Type A Soils	0.166	0.298
Developed - Low Intensity	0.145	0.367
Developed - total	0.144	0.368
Type B/D Soils	0.096	0.552
Grassland Herbaceous	0.070	0.665
Evergreen Forest	0.060	0.708
Palustrine Scrub Shrub Wetland	0.053	0.743
Impervious Surfaces	0.005	0.979
Palustrine Forested Wetland	-0.012	0.941
Total Land Area	-0.087	0.589
Bare Land	-0.087	0.589
Palustrine Emergent Wetland	-0.096	0.550
Type C/D Soils	-0.113	0.483
Type D Soils	-0.116	0.470
Estuarine Scrub Shrub Wetland	-0.167	0.298
Type A/D Soils	-0.193	0.227
Estuarine Forested Wetland	-0.295	0.061
Estuarine Emergent Wetland	-0.369	0.018
Unconsolidated Shore	-0.447	0.003
Open Water	-0.512	0.001

634

635

636

637



638

639

640

641

642

643

644

645

Figure 5. Classification and regression tree results for the comparison of PDiff to watershed characteristics. Non-tidal forested wetlands are those classified as Palustrine Forested Wetland by the Coastal Change Analysis Program (C-CAP) (<https://coast.noaa.gov/digitalcoast/tools/lca>). Moderate to very impervious soils are those classified as C/D soils by the United States Department of Agriculture (USDA) Soil Survey (<http://websoilsurvey.nrcs.usda.gov>).

646 **Table 2.** Bay-wide logistic regression results for four different models. The first three models include
 647 only measurements of rainfall (Rain2Day and 3wkRainfall) and distance from nearest
 648 shoreline(Distance). The fourth model also includes average air temperature (Air Temp) and wind
 649 (WindSpeed).

Model Parameters	% Concordant	c	AIC
Intercept 2DayRain	60.2	0.673	12710
Intercept 2DayRain Distance	75.6	0.756	12180
Intercept 2DayRain Distance 3wkRainfall	76.1	0.761	12110
Intercept 2DayRain Distance 3wkRainfall Average Daily Temp WindSpeed	76.5	0.76	12100

650

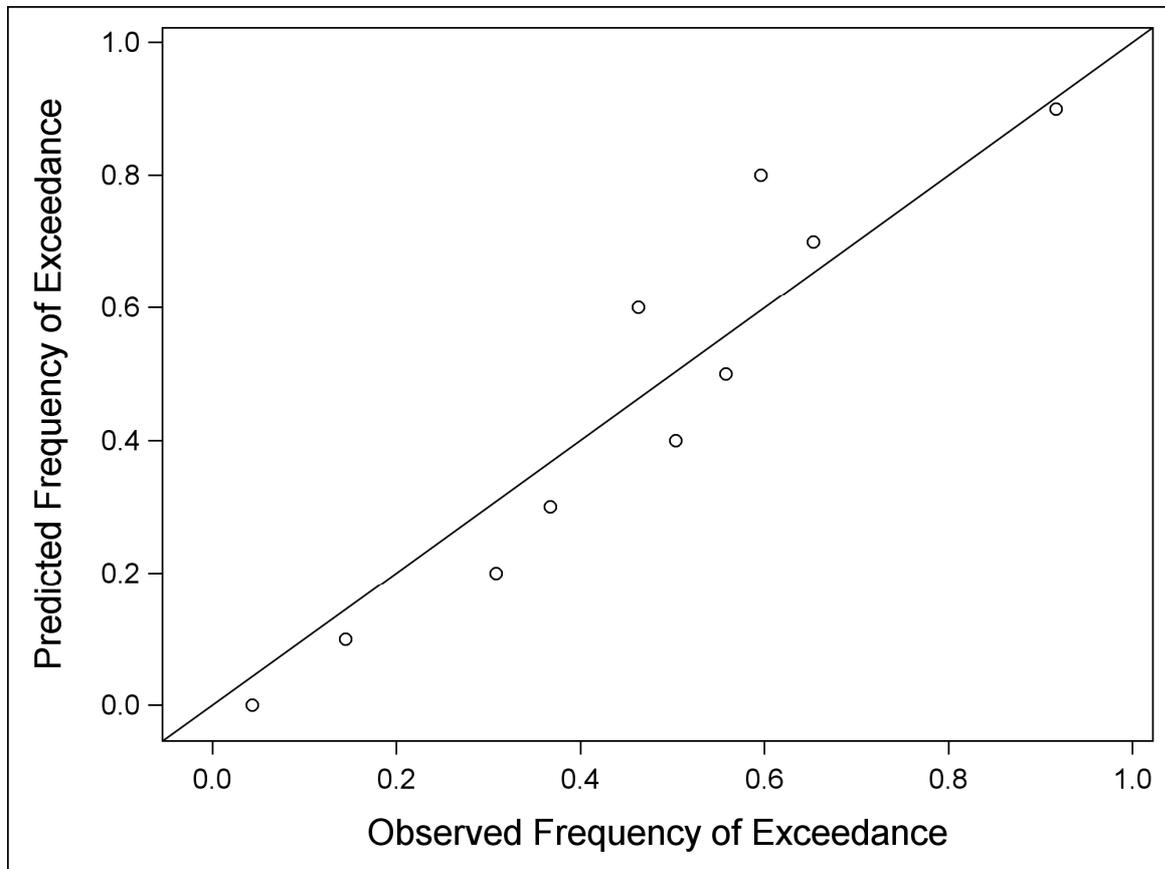
651 **Table 3.** Bay-wide model parameter estimates and results. The estimated rainfall threshold
 652 value is the amount of rainfall (Rain2day) in inches where there is a 50% probability of
 653 exceeding the 49MPN/100mL criterion. It uses the average conditions that were observed for all
 654 other variables besides Rain2Day.

Model	β^1	Odds Ratios	False Positives	False Negatives	RainFall Threshold ²	
					Threshold	CL
Intercept	-2.587		38.0	6.3	2.16	(2.07, 2.25)
2DayRain	1.525	4.59				
Distance	-0.004	1.00				
3wkRainfall	0.105	1.11				

655
 656 ¹-Parameter Estimates; ²- The estimated rainfall threshold value is the amount of rainfall (Rain2day) in
 657 inches where there is a 50% probability of exceeding the 49MPN/100mL criterion, using the average
 658 conditions that were observed for all other variables besides Rain2Day.

659

660



661

662 **Figure 6.** Comparison of predicted probability of exceedance (binned) to the proportion of observations in
663 each bin that exceeded the criteria. Adjusted R^2 for the linear relationship was 0.88. Predicted
664 probabilities are for all responsive watersheds based on logistic model with RainEvent, antecedent
665 rainfall, and the average distance of stations to shore.

666

667

668

669 **Table 4.** Logistic regression results for responsive watersheds with significant relationships to Rain2Day, concordance above 75%, and rainfall
 670 estimates having 95% confidence limits.

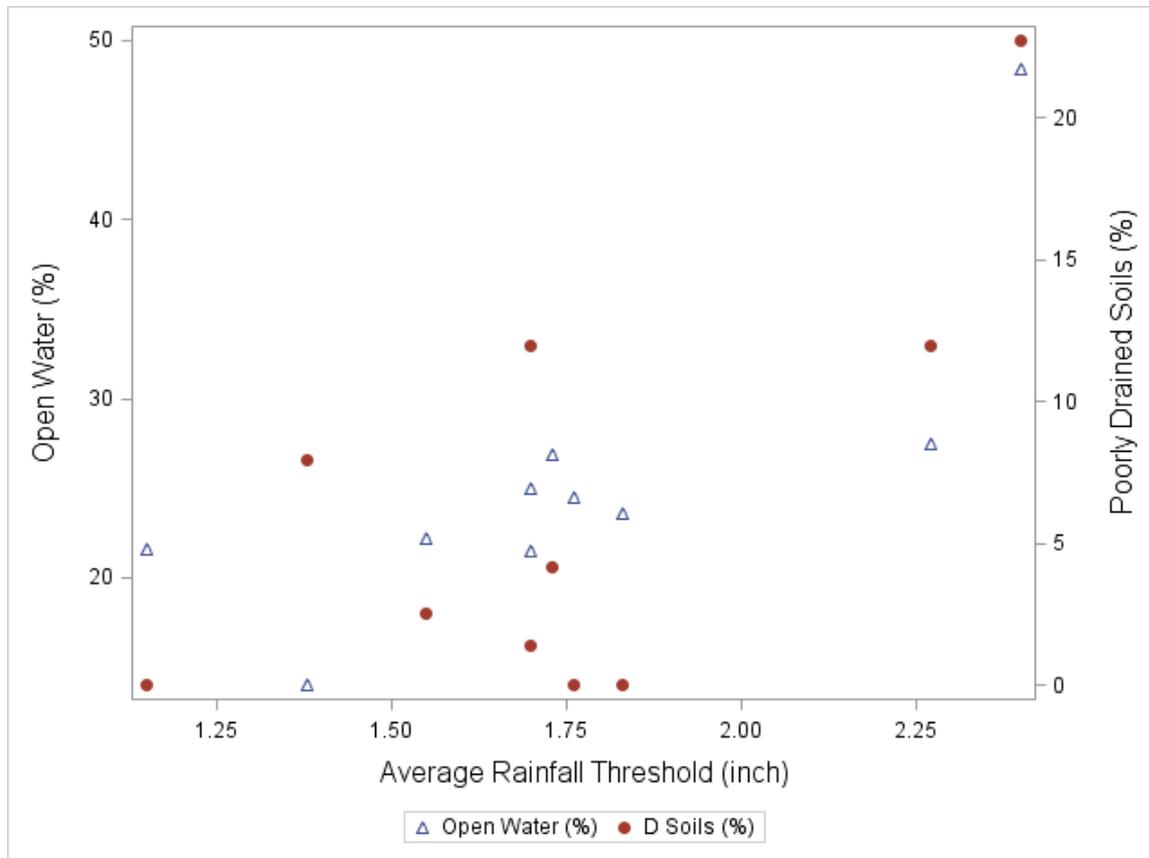
Watershed	# Samples	# RainEvents	Average Rainfall Threshold ¹	Odds Ratio Rain2Day	Concordant (%)	False Positive (%)	False Negative (%)	Distance p-value	Antecedent Rainfall p-value
Lower Wye East River	691	5	1.8 (1.4, 2.5)	5.9	76.1	40	6.5	> 0.050	< 0.001
Round Bay-Severn River	463	8	1.8 (1.4, 2.4)	6.2	84.9	17.6	6.7	< 0.001	0.049
Whitehall Creek-Severn River-Chesapeake Bay	1447	9	2.3 (2.0, 2.7)	5.6	84.8	31	4.5	< 0.001	0.002
Battle Creek-Patuxent River	507	3	1.6 (1.3, 2.0)	16.9	82.5	37.5	3	0.056	0.004
Saint Leonard Creek-Patuxent River	1149	10	1.7 (1.4, 2.0)	5.2	80.2	36.4	8.8	< 0.001	> 0.050
Mill Creek-Patuxent River	944	7	1.7 (1.4, 2.2)	6.2	77.7	31.8	8.4	< 0.001	> 0.050
Glebe Run-Breton Bay	719	9	1.2 (1.0, 1.4)	16.4	84.7	15	8.1	< 0.001	< 0.001
Eastern Branch-Saint Marys River	483	6	1.4 (1.2, 1.7)	16.8	87.1	29.4	3.3	> 0.050	> 0.050
Saint George Creek-Saint Marys River	1182	6	2.4 (2.1, 3.0)	4.9	86.0	52.4	5.3	< 0.001	> 0.050
Big Annemessex River	736	14	1.7 (1.5, 2.1)	10.5	85.7	39.1	4.5	0.003	> 0.050

671

672 ¹The estimated rainfall threshold value is the amount of rainfall (Rain2day) in inches where there is a 50% probability of exceeding the 49MPN/100mL criterion,
 673 using the average conditions that were observed for all other variables besides Rain2Day

674

675



676

677 **Figure 7.** Comparison by watershed of the percent of open water and soils with high runoff potential
678 (Type D) compared to the average rainfall threshold.

679

680 **Table 5.** Comparison of watershed-scale and conditionally-approved area specific logistic regression results for St. Mary's River.

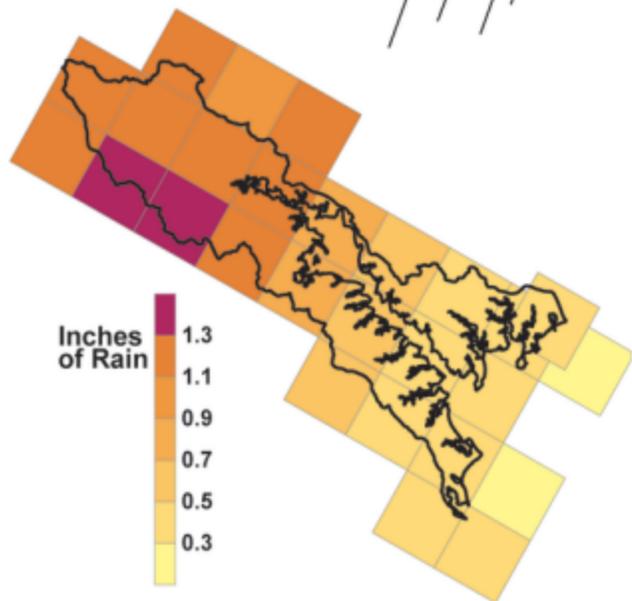
Stations	# Samples	# RainEvents	Rainfall Threshold ¹	Odds Ratio Rain2Day	Concordant (%)	False Positive (%)	False Negative (%)	Distance p-value	3wkRainfall p-value
All Stations in Watershed	483	6	1.4 (1.2, 1.7)	16.6	82.4	29.4	3.3	>0.05	>0.05
2 Stations in Conditionally Approved Area	153	6	1.3 (1.0, 1.7)	50.9	99.1	18.2	1.4	>0.05	>0.05

681

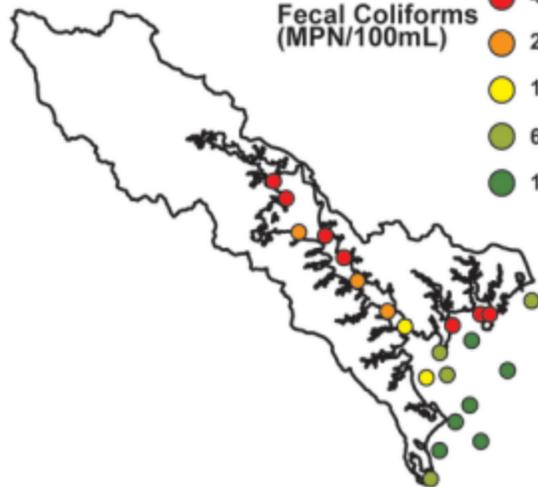
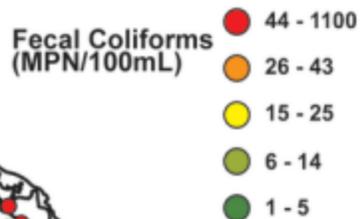
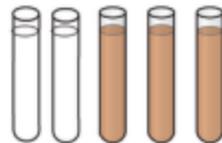
682 ¹The estimated rainfall threshold value is the amount of rainfall (Rain2day) in inches where there is a 50% probability of exceeding the 49MPN/100mL criterion,
683 using the average conditions that were observed for all other variables besides Rain2Day

684

Rain Estimates



Fecal Pollution Estimates



+

=

Predictive Models of Fecal Pollution

