Development of Regression-Based Models to Predict Fecal Indicator Bacteria at the Illinois River Basin, Arkansas and Oklahoma

UNIVERSITY OF ARKANSAS
DIVISION OF AGRICULTURE

Brian E. Haggard and Morgan M. David

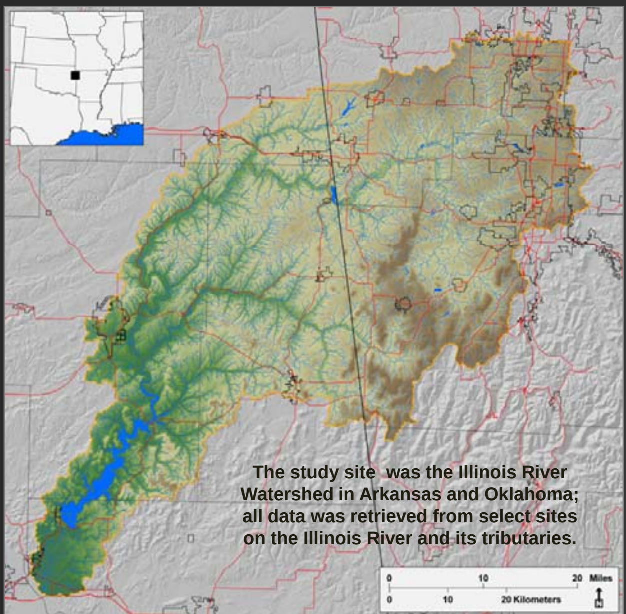
Department of Biological and Agricultural Engineering, University of Arkansas, Fayetteville, Arkansas, Email: haggard@uark.edu

Introduction

- **Fecal bacteria enter the aquatic environment through direct deposition, runoff from fields receiving manure application and municipal wastewater treatment plant effluent discharges.**
- **The presence of fecal bacteria, particularly E. Coli, in aquatic systems is an indicator of pathogens, as well as a potential or possible health risk for humans exposed to the water.**
- **Fecal bacteria presence and transport in the fluvial channel can be difficult and expensive to intensively monitor; however,** some studies have shown that real-time monitoring of surrogates can be used to predict fecal bacteria numbers in streams.

Methods

- **Archived water quality data from the USGS NWIS database was separated by flow conditions using Base Flow Index (BFI Model) .**
- **Base flow samples were defined as samples collected when base flow discharge was greater than or equal to 70% of mean daily discharge (total stream flow); storm flow samples represented other data.**
- **Several water quality parameters (e.g., conductivity, DO, temperature, sediment, etc.) were compared with bacteria concentrations (i.e., fecal coliform and E. Coli) at base flow and storm event conditions.**
- **A synthetic model was developed to identify the relationships between fecal bacteria and other measured physico-chemical parameters using all data, and that separated by flow conditions.**



The study site was the Illinois River Watershed in Arkansas and Oklahoma; all data was retrieved from select sites on the Illinois River and its tributaries.

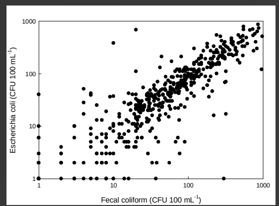
Objectives

- **Develop regression models that accurately predict fecal bacteria concentrations in the Illinois River Watershed using parameters measured in real-time during base flow and storm even conditions.**


Surface runoff (Left) and base flow (Right) conditions at Baron Fork Creek at Eldon, Oklahoma in the Illinois River Watershed during 2001.

Results

- **Fecal bacteria and E. coli concentrations were highly correlated during baseflow and storm events, except at low numbers.**
- **Fecal bacteria numbers were generally low during base flow conditions (see Tables below).**
- **Across the Illinois River and its tributaries, fecal bacteria numbers increased during surface runoff conditions, where fecal bacteria increased with increasing Q.**



Relationship between fecal coliform and E. coli at all sites during base flow and surface runoff event conditions.

- **The percent of samples with fecal bacteria numbers exceeding WQS was greatest at surface runoff.**

USGS Site Number	Flow	n	Escherichia coli (CFU/100 mL) Min	Max	Mean	Median	Primary Exceedence	Secondary Exceedence
07194800	ALL	48	4	5100	500	300	60.4%	8.3%
07195000	ALL	46	12	540	200	100	37.0%	0.0%
07195430	ALL	73	5	72000	4350	1600	43.8%	34.2%
07195500	ALL	36	1	13000	1700	200	22.2%	13.9%
07195865	ALL	60	1	19000	435	30	8.3%	1.7%
07196000	ALL	77	1	40000	1850	100	31.2%	19.5%
07196090	ALL	39	1	20000	3050	2200	56.4%	46.2%
07196500	ALL	80	1	15000	950	200	33.8%	26.3%
07196900	ALL	50	8	14000	550	200	48.0%	12.0%
07197000	ALL	75	1	41000	1500	100	29.3%	21.3%
07197360	ALL	56	1	6400	260	100	12.5%	5.4%
07194800	BF	29	4	750	200	200	51.7%	0.0%
07195000	BF	34	17	540	200	100	35.3%	0.0%
07195430	BF	36	12	210	100	70	8.3%	0.0%
07195500	BF	25	1	300	100	30	8.0%	0.0%
07195865	BF	33	1	900	70	30	9.1%	0.0%
07196000	BF	35	1	1500	100	40	11.4%	0.0%
07196090	BF	14	2	16000	1200	20	14.3%	7.1%
07196500	BF	35	1	2000	100	30	5.7%	0.0%
07196900	BF	16	29	660	200	100	31.3%	6.3%
07197000	BF	33	1	440	100	30	6.1%	0.0%
07197360	BF	31	1	88	20	10	0.0%	0.0%
07194800	SRO	19	84	5100	800	300	73.7%	21.1%
07195000	SRO	12	12	450	200	100	41.7%	0.0%
07195430	SRO	37	5	72000	8600	3200	78.4%	67.6%
07195500	SRO	11	1	13000	3300	300	54.5%	45.5%
07195865	SRO	27	1	19000	800	30	7.4%	3.7%
07196000	SRO	42	2	40000	3600	200	47.6%	35.7%
07196090	SRO	25	1	20000	4900	4400	80.0%	68.0%
07196500	SRO	45	3	15000	1800	300	55.6%	46.7%
07196900	SRO	34	8	14000	900	200	55.9%	14.7%
07197000	SRO	42	1	41000	2900	100	47.6%	38.1%
07197360	SRO	25	1	6400	500	100	28.0%	12.0%

USGS Site Number	Flow	n	Fecal Coliform (CFU/100 mL) Min	Max	Mean	Median	Primary Exceedence	Secondary Exceedence
07194800	ALL	48	17	4500	400	200	45.8%	8.3%
07195000	ALL	46	10	520	150	100	32.6%	0.0%
07195430	ALL	73	4	110000	4400	2000	41.1%	32.9%
07195500	ALL	36	1	18000	2300	100	22.2%	13.9%
07195865	ALL	42	1	17000	575	40	9.5%	2.4%
07196000	ALL	59	1	42000	2450	200	37.3%	30.5%
07196090	ALL	39	2	21000	3400	1600	56.4%	41.0%
07196500	ALL	62	1	19000	1500	400	46.8%	30.6%
07196900	ALL	49	4	7400	400	200	42.9%	8.2%
07197000	ALL	58	1	50000	2200	300	37.9%	24.1%
07197360	ALL	56	1	4200	200	25	10.7%	3.6%
07194800	BF	29	17	470	200	100	34.5%	0.0%
07195000	BF	34	10	520	100	100	32.4%	0.0%
07195430	BF	36	6	220	100	40	2.8%	0.0%
07195500	BF	25	1	510	100	30	8.0%	0.0%
07195865	BF	25	1	250	50	30	8.0%	0.0%
07196000	BF	25	1	560	100	30	4.0%	4.0%
07196090	BF	14	3	18000	1300	10	14.3%	7.1%
07196500	BF	26	1	360	50	20	7.7%	3.8%
07196900	BF	16	20	1100	200	100	18.8%	0.0%
07197000	BF	26	1	250	40	20	7.7%	0.0%
07197360	BF	31	1	120	20	10	0.0%	0.0%
07194800	SRO	19	62	4500	600	300	63.2%	21.1%
07195000	SRO	12	17	430	200	100	33.3%	0.0%
07195430	SRO	37	4	110000	8700	3900	78.4%	64.9%
07195500	SRO	11	5	18000	4500	200	54.5%	45.5%
07195865	SRO	17	1	17000	1100	40	11.8%	5.9%
07196000	SRO	34	6	42000	4800	300	61.8%	50.0%
07196090	SRO	25	2	21000	5500	3200	80.0%	60.0%
07196500	SRO	36	2	19000	3000	800	75.0%	50.0%
07196900	SRO	33	4	7400	600	200	54.5%	12.1%
07197000	SRO	32	1	50000	4400	500	62.5%	43.8%
07197360	SRO	25	1	4200	400	40	24.0%	8.0%

Fecal coliform (left) and E. coli (right) concentrations during combined (ALL), base flow (BF), and surface runoff events (SRO). Primary % exceedence: bacteria > 200 CFU per 100 mL. Secondary % exceedence bacteria > 100 CFU per 1000 mL.

Results

- **These tables detail the regression models developed in this study.**

USGS[b] Site Number	Flow	n	Stepwise linear regression equation	R^2	*p*
07194800	ALL	48	$0.443\log_{10}Qi + 0.669\log_{10}T - 2.83\log_{10}SC - 1.71\log_{10}DO + 9.16$	0.67	0.06
07195000	ALL	46	$0.980\log_{10}T - 1.92\log_{10}SC - 2.27\log_{10}DO + 8.13$	0.59	0.01
07195430	ALL	72	$1.58\log_{10}Qi + 2.65\log_{10}SS - 2.48\log_{10}DO - 4.82$	0.73	0.04
07195500	ALL	36	$1.87\log_{10}Qi - 2.68\log_{10}DO - 0.639$	0.62	0.05
07195865	ALL	40	$0.773\log_{10}Qi - 2.18\log_{10}DO + 2.85$	0.19	0.07
07196000	ALL	39	$0.770\log_{10}SS + 2.21\log_{10}pH - 20.0\log_{10}DO + 20.5$	0.64	<0.01
07196090	ALL	39	$1.27\log_{10}Qi - 13.2\log_{10}pH + 10.1$	0.75	0.06
07196500	ALL	62	$1.22\log_{10}Qi - 3.11\log_{10}DO + 1.34$	0.67	<0.01
07196900	ALL	41	$0.823\log_{10}T + 2.04\log_{10}SS - 2.78$	0.19	0.05
07197000	ALL	56	$282\log_{10}T - 98.1\log_{10}SC - 4440\log_{10}pH + 48800$	0.32	0.05
07194800	BF	29	$0.630\log_{10}Qi + 0.933\log_{10}T - 1.45\log_{10}DO + 1.41$	0.52	0.09
07195000	BF	34	$0.970\log_{10}T - 1.83\log_{10}SC - 2.07\log_{10}DO + 7.70$	0.54	0.05
07195430	BF	35	$-1.56\log_{10}DO + 3.19$	0.29	0.79
07195500	BF	25	No Model Developed	-	-
07195865[e]	BF	23	No Model Developed	-	-
07196000[e]	BF	22	$-25.7\log_{10}DO + 24.7$	0.21	0.03
07196090	BF	14	$-17.5\log_{10}SC + 45.0$	0.38	0.02
07196500	BF	26	$-2.85\log_{10}DO + 4.16$	0.23	0.01
07196900[e]	BF	13	No Model Developed	-	-
07197000	BF	25	$-3.80\log_{10}DO + 5.03$	0.33	<0.01
07194800[c]	SRO	19	$-4.86\log_{10}SC - 2.26\log_{10}DO + 16.1$	0.76	0.20
07195000	SRO	12	$-4.18\log_{10}DO + 6.27$	0.70	<0.01
07195430	SRO	37	$1.36\log_{10}Qi + 5.46\log_{10}SS - 11.8$	0.63	<0.01
07195500	SRO	11	$2.61\log_{10}Qi - 5.59$	0.76	<0.01
07195865	SRO	17	$1.72\log_{10}Qi - 2.78\log_{10}DO + 2.57$	0.64	0.06
07196000	SRO	17	$1.49\log_{10}SS + 2.71\log_{10}pH + 2.47$	0.86	<0.01
07196090	SRO	25	$2.23\log_{10}Qi + 4.47\log_{10}SC - 13.5\log_{10}pH - 3.34$	0.78	0.10
07196500	SRO	36	$1.29\log_{10}Qi - 3.36\log_{10}DO + 1.33$	0.62	0.01
07196900	SRO	28	$4.34\log_{10}SS - 1.60\log_{10}DO - 4.84$	0.24	0.07
07197000	SRO	31	$-32.0\log_{10}pH + 30.1$	0.36	0.61

USGS[b] Site Number	Flow	n	Stepwise linear regression equation	R^2	*p*
07194800[c]	ALL	48	No Model Developed	-	-
07195000	ALL	46	$0.709\log_{10}T - 2.04\log_{10}SC - 2.21\log_{10}DO + 8.57$	0.46	0.07
07195430	ALL	72	$1.09\log_{10}Qi + 2.05\log_{10}SS - 3.01\log_{10}SC - 2.59\log_{10}DO + 5.10$	0.76	0.10
07195500	ALL	36	$1.84\log_{10}Qi - 4.06\log_{10}DO + 0.784$	0.62	<0.01
07195865	ALL	40	$0.764\log_{10}Qi - 2.99\log_{10}DO + 3.49$	0.18	0.04
07196000	ALL	39	$0.734\log_{10}SS + 2.44\log_{10}pH - 20.0\log_{10}DO + 20.6$	0.51	0.02
07196090	ALL	39	$1.31\log_{10}Qi - 13.3\log_{10}pH + 10.0$	0.75	0.06
07196500	ALL	62	$1.38\log_{10}Qi - 2.91\log_{10}DO + 0.678$	0.63	<0.01
07196900	ALL	40	$-0.934\log_{10}DO + 2.84$	0.08	0.08
07197000	ALL	56	$334\log_{10}T - 124\log_{10}SC - 4930\log_{10}pH + 56800$	0.35	0.06
07194800	BF	29	$0.472\log_{10}Qi + 0.889\log_{10}T + 5.62\log_{10}pH - 1.16\log_{10}DO - 3.69$	0.70	0.04
07195000	BF	34	$-3.25\log_{10}DO - 2.49\log_{10}SC + 11.6$	0.40	0.13
07195430	BF	35	$-2.15\log_{10}SC - 1.81\log_{10}DO + 8.74$	0.36	0.04
07195500	BF	25	$1.46\log_{10}T - 0.416$	0.21	0.02
07195865[e]	BF	23	No Model Developed	-	-
07196000	BF	22	$-25.9\log_{10}DO + 24.7$	0.14	0.09
07196090	BF	14	$-17.2\log_{10}SC + 44.2$	0.37	0.02
07196500	BF	26	$-2.75\log_{10}DO + 3.83$	0.17	0.04
07196900	BF	13	$-0.264\log_{10}Qi + 0.362\log_{10}T - 3.82\log_{10}SC + 11.2$	0.90	<0.01
07197000	BF	25	$-4.04\log_{10}DO + 5.06$	0.36	<0.01
07194800	SRO	19	$-1.86\log_{10}SS - 4.34\log_{10}SC - 1.93\log_{10}DO + 18.2$	0.70	0.06
07195000	SRO	12	$-3.34\log_{10}DO + 5.44$	0.65	<0.01
07195430	SRO	37	$1.48\log_{10}Qi + 5.26\log_{10}SS - 11.9$	0.65	<0.01
07195500	SRO	11	$2.62\log_{10}Qi - 5.58$	0.77	<0.01
07195865	SRO	17	$1.57\log_{10}Qi - 3.73\log_{10}DO + 3.50$	0.47	0.07
07196000[c]	SRO	17	No Model Developed	-	-
07196090	SRO	25	$1.02\log_{10}Qi - 4.02\log_{10}T - 9.11\log_{10}DO + 12.8$	0.82	<0.01
07196500	SRO	36	$1.38\log_{10}Qi + 0.992\log_{10}T + 9.89\log_{10}SS - 22.4$	0.61	0.06
07196900	SRO	27	$9.91\log_{10}pH - 2.38\log_{10}DO - 4.69$	0.26	0.08
07197000	SRO	31	$-3.84\log_{10}SC - 28.1\log_{10}pH + 35.4$	0.48	0.03


Summary

- **The regression models were somewhat different across the sites and the databases used (all data, base flow data, and surface runoff data) at the Illinois River and its tributaries.**
- **The predictive abilities of these regression models was limited during base flow conditions, as several sites showed that no regression model could be developed using these parameters.**
- **The ability of the regression models to predict changes in fecal bacteria numbers differed across sites, where the regression models explained between 8 and 86% of the variation in bacterial numbers.**
- **Several other studies across the south-central USA have developed regression models to predict fecal bacteria numbers, and these models collectively suggest that turbidity should be included.**