

Analysis of the Potential of Macro-Invertebrates as Indicators for Microbial Pathogens in Rivers

Rubén Jerves-Cobo*, Gonzalo Córdova-Vela, Xavier Iñiguez-Vela, Catalina Díaz-Granda, Wout Van Echelpoel, Felipe Cisneros, Ingmar Nopens and Peter L. M. Goethals

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4441/10/4/375/s1>, Figure S1: Sampled sites location 2012, 2015 and 2016; Table S1: Verification of the fulfilment of the recreational with primary contact Ecuadorian water use regulations associated with fecal coliforms according the decision tree models (DTMs); Table S2: Verification of the fulfilment of the agriculture and livestock water use regulations associated with fecal coliforms according the decision tree models (DTMs), before the optimization process; Table S3: Verification of the fulfilment of the recreational with primary contact Ecuadorian water use regulations associated with fecal coliforms according the decision tree models (DTMs), with new dataset taken in July of 2015 (dry season) and March of 2016 (rainy season); Table S4: Verification of the fulfilment of the agriculture and livestock Ecuadorian water use regulations associated with fecal coliforms according the decision tree models (DTMs), with new dataset taken in July of 2015 and March of 2016; Table S5: Calculation of the variation of correctly classified instances (CCI) and Kappa statistics in the recreational fecal regulation models, in which three cross validations were manually applied; Table S6: Calculation of the variation of correctly classified instances (CCI) and Kappa statistics in the agriculture fecal regulation models, in which three cross validations were manually applied.

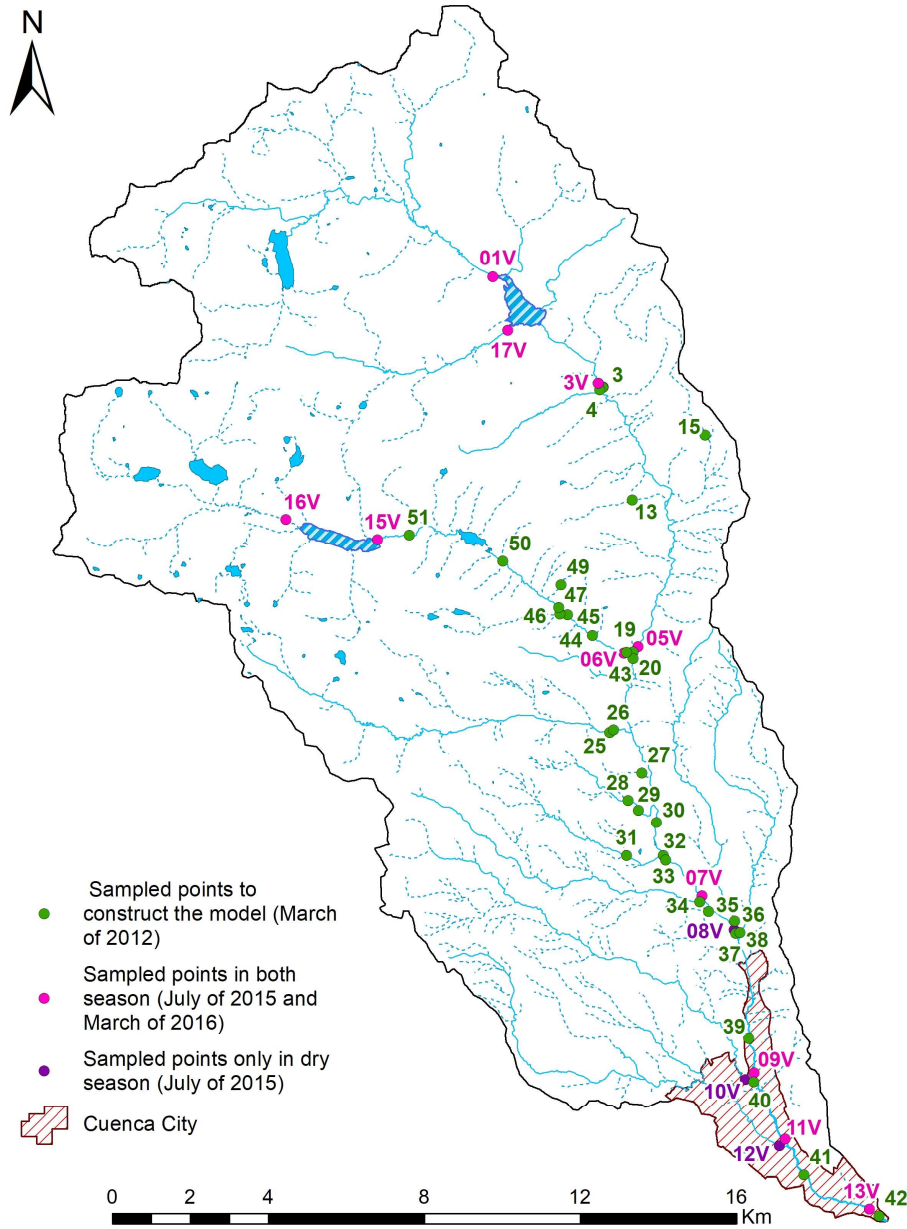


Figure S1: Sampled sites location 2012, 2015 and 2016.

Table S1: Verification of the fulfilment of the recreational with primary contact Ecuadorian water use regulations associated with fecal coliforms according the decision tree models (DTMs).

Sampled Points	Land Use	Recreational fecal coliform regulation			Threshold = 2.00E+02 MPN/100 mL		
		Models: 1a8			Models: 1a5, 1a6 and 1a7.		
		Abundance		Fecal	Abundance		Fecal
		Baetidae	Elminthidae	Coliforms	Baetidae	Scirtidae	Coliforms
3	5	3	2.00	7.0E+01	3	0	7.0E+01
4	5	13	2.00	4.5E+00	13	0	4.5E+00
13	5	37	0.00	3.3E+02	37	0	3.3E+02
15	3	0	3.00	3.3E+02	0	132	3.3E+02
19	5	0	0.00	7.9E+01	0	1	7.9E+01
20	4	2	0.00	4.9E+01	2	1	4.9E+01
24	3	43	16.00	2.4E+02	43	0	2.4E+02
25	5	16	2.00	4.9E+01	16	0	4.9E+01
26	6	2	0.00	4.9E+01	2	2	4.9E+01
27	3	38	3.00	4.6E+02	38	3	4.6E+02
28	6	3	1.00	4.9E+01	3	1	4.9E+01
29	6	2	1.00	3.3E+01	2	0	3.3E+01
30	3	1	5.00	7.9E+01	1	0	7.9E+01
31	6	1	1.00	3.3E+01	1	1	3.3E+01
32	3	0	0.00	4.9E+01	0	0	4.9E+01
33	6	3	2.00	3.3E+01	3	1	3.3E+01
34	2	53	4.00	1.7E+05	53	1	1.7E+05
35	2	7	0.00	1.6E+04	7	4	1.6E+04
36	2	189	1.00	5.4E+05	189	4	5.4E+05
37	2	45	0.00	6.0E+03	45	0	6.0E+03
38	2	108	0.00	1.1E+05	108	2	1.1E+05
39	1	151	0.00	1.7E+04	151	0	1.7E+04
40	1	31	0.00	1.7E+05	31	0	1.7E+05
41	1	17	0.00	1.1E+05	17	0	1.1E+05
42	1	19	0.00	4.6E+04	19	0	4.6E+04
43	5	1	1.00	1.3E+02	1	0	1.3E+02
44	5	3	0.00	7.0E+01	3	0	7.0E+01
45	5	4	0.00	4.6E+02	4	0	4.6E+02
46	3	30	4.00	1.7E+01	30	0	1.7E+01
47	6	0	1.00	2.3E+01	0	0	2.3E+01
49	5	140	1.00	1.7E+01	140	0	1.7E+01
50	3	0	0.00	3.3E+01	0	0	3.3E+01
51	5	0	0.00	4.5E+00	0	0	4.5E+00

Symbology:

Land use: 1 Urban area 2 Suburban area, pastures and crops 3 Pastures
 4 Bare soil 5 Native vegetation 6 Forest vegetation


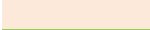


 Fulfilment of regulation
 Non-fulfilment of regulation
 Place where the first rule (branch) of the decision tree is applicable
 Place where the second rule(branch) of the decision tree is applicable

Table S2: Verification of the fulfilment of the agriculture and livestock water use regulations associated with fecal coliforms according the decision tree models (DTMs), before the optimization process.

Sampled Points	Land Use	Agriculture fecal coliform regulation			Threshold = 1.00E+03 MPN/100 mL			
		Model: 2ap3			Models: 2a1, 2a2, 2a3, 2a4, 2a5, 2a6.			
		Presence/Absence			Fecal Coliforms	Abundance		
	Perlidiae	Baetidae	Leptophlebiidae		Baetidae	Perlidiae	Fecal Coliforms	
3	5	-	P	P	7.0E+01	3	0	7.0E+01
4	5	P	P	-	4.5E+00	13	1	4.5E+00
13	5	P	P	P	3.3E+02	37	7	3.3E+02
15	3	-	-	-	3.3E+02	0	0	3.3E+02
19	5	-	-	-	7.9E+01	0	0	7.9E+01
20	4	-	P	-	4.9E+01	2	0	4.9E+01
24	3	P	P	-	2.4E+02	43	6	2.4E+02
25	5	P	P	P	4.9E+01	16	1	4.9E+01
26	6	-	P	P	4.9E+01	2	0	4.9E+01
27	3	P	P	P	4.6E+02	38	4	4.6E+02
28	6	P	P	P	4.9E+01	3	1	4.9E+01
29	6	P	P	-	3.3E+01	2	1	3.3E+01
30	3	P	P	-	7.9E+01	1	1	7.9E+01
31	6	-	P	-	3.3E+01	1	0	3.3E+01
32	3	-	-	-	4.9E+01	0	0	4.9E+01
33	6	P	P	-	3.3E+01	3	1	3.3E+01
34	2	-	P	-	1.7E+05	53	0	1.7E+05
35	2	-	P	-	1.6E+04	7	0	1.6E+04
36	2	-	P	-	5.4E+05	189	0	5.4E+05
37	2	-	P	-	6.0E+03	45	0	6.0E+03
38	2	-	P	-	1.1E+05	108	0	1.1E+05
39	1	-	P	-	1.7E+04	151	0	1.7E+04
40	1	-	P	-	1.7E+05	31	0	1.7E+05
41	1	-	P	-	1.1E+05	17	0	1.1E+05
42	1	-	P	-	4.6E+04	19	0	4.6E+04
43	5	P	P	-	1.3E+02	1	1	1.3E+02
44	5	-	P	P	7.0E+01	3	0	7.0E+01
45	5	-	P	-	4.6E+02	4	0	4.6E+02
46	3	P	P	P	1.7E+01	30	1	1.7E+01
47	6	-	-	-	2.3E+01	0	0	2.3E+01
49	5	P	P	P	1.7E+01	140	9	1.7E+01
50	3	-	-	-	3.3E+01	0	0	3.3E+01
51	5	-	-	-	4.5E+00	0	0	4.5E+00

Symbology:

Land use: 1 Urban area 2 Suburban area, pastures and crops 3 Pastures
4 Bare soil 5 Native vegetation 6 Forest vegetation





 Fulfilment of regulation
 Non-fulfilment of regulation
 Place where the first rule (branch) of the decision tree is applicable
 Place where the second rule(branch) of the decision tree is applicable

Table S3: Verification of the fulfilment of the recreational with primary contact Ecuadorian water use regulations associated with fecal coliforms according the decision tree models (DTMs), with new dataset taken in July of 2015 (dry season) and March of 2016 (rainy season).

Sampled Points	Season	Land Use	Recreational fecal coliform regulation			Threshold = 2.00E+02 MPN/100 mL		
			Models: 1a8			Models: 1a5, 1a6 and 1a7.		
			Abundance		Fecal	Abundance		Fecal
		Baetidae	Elminthidae	Coliforms	Baetidae	Scirtidae	Coliforms	
01V	Rainy	5	4	83.00	6.8E+00	4	0	6.8E+00
03V	Rainy	6	1	13.00	4.5E+00	1	0	4.5E+00
05V	Rainy	6	2	4.00	4.5E+00	2	0	4.5E+00
06V	Rainy	6	16	0.00	4.0E+01	16	0	4.0E+01
07V	Rainy	3	5	3.00	2.3E+02	5	0	2.3E+02
09V	Rainy	2	10	0.00	1.7E+04	10	0	1.7E+04
11V	Rainy	1	48	4.00	2.7E+04	48	0	2.7E+04
13V	Rainy	1	38	2.00	9.3E+04	38	0	9.3E+04
15V	Rainy	5	2	0.00	1.8E+00	2	0	1.8E+00
16V	Rainy	5	0	37.00	1.8E+00	0	0	1.8E+00
17V	Rainy	5	0	25.00	4.5E+00	0	0	4.5E+00
01V	Dry	5	2	16	2.0E+00	2	1	2.0E+00
03V	Dry	6	0	2	4.0E+00	0	0	4.0E+00
05V	Dry	6	3	0	1.8E+00	3	0	1.8E+00
06V	Dry	6	7	8	1.3E+01	7	0	1.3E+01
07V	Dry	3	58	1	1.7E+02	58	0	1.7E+02
08V	Dry	2	67	2	5.4E+03	67	0	5.4E+03
09V	Dry	2	376	1	3.3E+03	376	0	3.3E+03
10V	Dry	2	463	2	2.2E+04	463	5	2.2E+04
11V	Dry	1	91	1	1.7E+05	91	0	1.7E+05
12V	Dry	1	121	0	1.1E+06	121	0	1.1E+06
13V	Dry	1	8	0	7.0E+05	8	0	7.0E+05
15V	Dry	5	0	0	1.8E+00	0	0	1.8E+00
16V	Dry	5	2	17	1.8E+00	2	0	1.8E+00
17V	Dry	5	1	11	2.0E+00	1	2	2.0E+00

Symbology:

Land use:

- | | | | | | |
|---|------------|---|-----------------------------------|---|-------------------|
| 1 | Urban area | 2 | Suburban area, pastures and crops | 3 | Pastures |
| 4 | Bare soil | 5 | Native vegetation | 6 | Forest vegetation |

	Fulfilment of regulation
	Non-fulfilment of regulation
	Place where the first rule (branch) of the decision tree is applicable
	Place where the second rule(branch) of the decision tree is applicable

Table S4: Verification of the fulfilment of the agriculture and livestock Ecuadorian water use regulations associated with fecal coliforms according the decision tree models (DTMs), with new dataset taken in July of 2015 and March of 2016.

Sampled Points	Season	Land Use	Agriculture fecal coliform regulation				Threshold = 1.00E+03 MPN/100 mL			
			Model: 2ap3			Fecal Coliforms	Abundance		Fecal Coliforms	
			Presence/Absence	Perlidiae	Baetidae		Leptophlebiidae	Baetidae	Perlidiae	Perlidiae
01V	Rainy	5	-	P	P	6.8E+00	4	-	6.8E+00	
03V	Rainy	6	-	P	-	4.5E+00	1	-	4.5E+00	
05V	Rainy	6	-	P	-	4.5E+00	2	-	4.5E+00	
06V	Rainy	6	-	P	-	4.0E+01	16	-	4.0E+01	
07V	Rainy	3	-	P	-	2.3E+02	5	-	2.3E+02	
09V	Rainy	2	-	P	-	1.7E+04	10	-	1.7E+04	
11V	Rainy	1	-	P	-	2.7E+04	48	-	2.7E+04	
13V	Rainy	1	-	P	-	9.3E+04	38	-	9.3E+04	
15V	Rainy	5	-	P	-	1.8E+00	2	-	1.8E+00	
16V	Rainy	5	-	-	P	1.8E+00	0	-	1.8E+00	
17V	Rainy	5	-	-	P	4.5E+00	0	-	4.5E+00	
01V	Dry	5	-	P	P	4.0E+00	2	-	4.0E+00	
03V	Dry	6	-	-	-	1.8E+00	0	-	1.8E+00	
05V	Dry	6	-	P	-	1.3E+01	3	-	1.3E+01	
06V	Dry	6	-	P	-	1.7E+02	7	-	1.7E+02	
07V	Dry	3	-	P	-	5.4E+03	58	-	5.4E+03	
08V	Dry	2	-	P	-	3.3E+03	67	-	3.3E+03	
09V	Dry	2	-	P	-	2.2E+04	376	-	2.2E+04	
10V	Dry	2	-	P	-	1.7E+05	463	-	1.7E+05	
11V	Dry	1	-	P	-	1.1E+06	91	-	1.1E+06	
12V	Dry	1	-	P	-	7.0E+05	121	-	7.0E+05	
13V	Dry	1	-	P	-	1.8E+00	8	-	1.8E+00	
15V	Dry	5	-	-	-	1.8E+00	0	-	1.8E+00	
16V	Dry	5	-	P	P	2.0E+00	2	-	2.0E+00	
17V	Dry	5	-	P	P	0.0E+00	1	-	0.0E+00	

Symbology:

Land use: 1 Urban area 2 Suburban area, pastures and crops 3 Pastures
4 Bare soil 5 Native vegetation 6 Forest vegetation

	Fulfilment of regulation
	Non-fulfilment of regulation
	Place where the first rule (branch) of the decision tree is applicable
	Place where the second rule(branch) of the decision tree is applicable

Table S5: Calculation of the variation of correctly classified instances (CCI) and Kappa statistics in the recreational fecal regulation models, in which three cross validations were manually applied

Dataset	CROSS VALIDATION		Cost Matrix Weights (CMW)				CCI ^e (%)			Kappa statistics		
	FOLDS	Pruning confidence factor (PCF)	TP ^a	FN ^b	FP ^c	TN ^d	mean	±	sd	mean	±	sd
Abundance	3	0.25					84.8	±	10.5	0.70	±	0.21
Abundance	3	0.10					81.8	±	9.1	0.63	±	0.18
Abundance	3	0.25	0	1	2	0	78.8	±	18.9	0.58	±	0.37
Abundance	3	0.10	0	1	2	0	78.8	±	18.9	0.58	±	0.38
Abundance	3	0.25	0	1	3	0	69.7	±	22.9	0.42	±	0.43
Abundance	3	0.10	0	1	3	0	75.8	±	13.9	0.52	±	0.28
Abundance	3	0.25	0	1	5	0	69.7	±	22.9	0.42	±	0.43
Abundance	3	0.10	0	1	5	0	69.7	±	22.9	0.42	±	0.43
Abundance	3	0.25	0	1	7	0	54.5	±	15.7	0.14	±	0.27
Abundance	3	0.10	0	1	7	0	54.5	±	15.7	0.15	±	0.27
Presence-absence	3	0.25					72.7	±	9.1	0.44	±	0.20
Presence-absence	3	0.10					63.6	±	9.1	0.27	±	0.16
Presence-absence	3	0.25	0	1	2	0	54.5	±	9.1	0.13	±	0.16
Presence-absence	3	0.25	0	1	3	0	54.5	±	9.1	0.14	±	0.16
Presence-absence	3	0.25	0	1	5	0	51.5	±	10.5	0.09	±	0.18
Presence-absence	3	0.25	0	1	7	0	45.5	±	0.0	-0.01	±	0.02
Presence-absence	3	0.25	0	1	10	0	45.5	±	0.0	0.00	±	0

^a TP = True positives

^b FN = False negative

^c FP = False positive

^d TN = True negative

^e CCI = Correctly classified instances

Table S6: Calculation of the variation of correctly classified instances (CCI) and Kappa statistics in the agriculture fecal regulation models, in which three cross validations were manually applied

Dataset	CROSS VALIDATION		Cost Matrix Weights (CMW)				CCI ^e (%)			Kappa statistics		
	FOLDS	Pruning confidence factor (PCF)	TP ^a	FN ^b	FP ^c	TN ^d	mean	±	sd	mean	±	sd
Abundance	3	0.25					78.8	±	13.9	0.53	±	0.28
Abundance	3	0.10					78.8	±	13.9	0.53	±	0.28
Abundance	3	0.25	0	1	2	0	84.8	±	18.9	0.34	±	0.41
Abundance	3	0.25	0	1	3	0	84.8	±	18.9	0.34	±	0.41
Abundance	3	0.25	0	1	5	0	84.8	±	18.9	0.34	±	0.41
Abundance	3	0.25	0	1	7	0	84.8	±	18.9	0.34	±	0.41
Abundance	3	0.25	0	1	10	0	84.8	±	18.9	0.34	±	0.41
Presence-absence	3	0.25					75.8	±	5.2	0.38	±	0.15
Presence-absence	3	0.10					72.7	±	9.1	0.17	±	0.22
Presence-absence	3	0.25	0	1	2	0	72.7	±	9.1	0.39	±	0.28
Presence-absence	3	0.25	0	1	3	0	81.8	±	9.1	0.63	±	0.16
Presence-absence	3	0.25	0	1	5	0	57.6	±	5.2	0.19	±	0.20
Presence-absence	3	0.25	0	1	7	0	57.6	±	5.2	0.19	±	0.20
Presence-absence	3	0.25	0	1	10	0	57.6	±	5.2	0.19	±	0.20

^a TP = True positives

^b FN = False negative

^c FP = False positive

^d TN = True negative

^e CCI = Correctly classified instances