Supporting Information

Verhougstraete et al. 10.1073/pnas.1415836112

SI Materials and Methods

Landscape Characteristics. The land use composition across the project area can be split into two groups (Fig. S3): The northern area has more forest and wetlands ($P < 0.01$) and the southern area has more urban and agriculture ($P < 0.01$). Southern watersheds, based on latitude, showed *E. coli* ($r > 0.345$, $P < 0.005$) and *B. theta* ($r > 0.250$, $P < 0.05$) concentrations were statistically correlated to agriculture at the reduced watershed and 60-m buffered scales.

The estimated number of on-site septic systems was highly variable across the study area, with SDs roughly twice the mean for each of the three scales. In contrast, septic system densities were similar across all three scales. Interestingly, impervious surface coverage in the 60-m buffer (average = 5.5%) and full watersheds (average = 7.5%) were correlated to septic density at the same spatial scale ($r \geq 0.370$, $P < 0.001$). The number and density of on-site septic systems was higher in the southern sites compared with northern sites at both the full watershed and the 60-m buffer scales ($P < 0.006$). The land use composition and classification of each river system including septic systems at the full watershed, reduced watershed, and 60-m riparian buffer are defined in Table S1 and summarized in Table S4.

The USEPA DMR Pollutant Loading Tool (cfpub.epa.gov/dmr/ez_search.cfm) was used to estimate the ratio of average annual WWTP effluent to measured baseflow. The total sum of WWTP discharges (million gallons per day, MGD) in each watershed was calculated in Esri ArcMap GIS software. This total WWTP discharge was compared with the measured baseflow river discharge to produce the ratio of average annual WWTP effluent to measured baseflow. The ratio of average annual WWTP effluent to measured baseflow was calculated using annual averages of WWTP discharge and field measurements; thus, values greater than 100% were possible and any watersheds exceeding 100% were removed from calculations. Although estimated levels of bacterial discharge are reported to the DMR, it was not appropriate to calculate a proportion of measured bacteria attributable to WWTP effluent because bacteria concentrations can change quickly (65) and the concentrations reported from WWTP are generally annual estimates of fecal coliforms.

The percentage of measured flow attributable to WWTP effluent was estimated to be between 0% and 52%, with a mean of 4%. The ratio of average annual WWTP effluent to measured baseflow flow was calculated using annual averages of WWTP discharge and field baseflow measurements; thus, values greater than 100% are possible. Only seven watersheds had WWTP contributions above 10% of measured flow. Our analysis also included mean population densities served by WWTP as estimated from census blocks and wastewater service boundaries. When the 28 watersheds with >80% of the population relying on WWTP service were excluded from our CART analysis, the primary split variables remained the same for *E. coli* and *B. theta* concentrations. Furthermore, sources of human bacteria could not be distinguished because no statistical difference ($P > 0.1$) of bacterial concentrations was identified between these two groups of watersheds [i.e., WWTP-reliant (>80% of the population living inside the WWTP service area, $n = 28$) or septic-reliant (>80% of the population living outside the WWTP service area, $n = 36$)]. In this case, we could not statistically differentiate the impacts of the point source WWTP effluent on receiving water bodies from the plethora of nonpoint sources measured during baseflow sample collection. Previous studies from Michigan demonstrated that *B. theta* concentrations in untreated sewage averaged 7.2 $\log_{10}$ CE/100 mL and were reduced by 3.1 logs through secondary treatment before discharge (66).

Hydrogeologic and Geochemical Properties. The watersheds included in our study were characterized under baseflow conditions to ensure precipitation was not significantly influencing stream flow. Six-hour cumulative precipitation totals were generally low with a mean of 0.14 mm. River discharge and discharge per area ranged from 0.01 to 57 $m^3 \cdot s^{-1}$ and 1.1×10^{-4} to 2.2×10^{-1} $m^3 \cdot s^{-1} \cdot km^{-2}$, respectively. Discharge for each river system is provided in Table S1. CART analysis identified the pH, total phosphorus, water temperature, potassium, and septic system numbers in the watershed as significantly related to microbial water quality. Descriptive statistics for all measured hydrogeologic variables are provided in Table S2.

Reduced Dataset Analysis. In the reduced watersheds (Fig. S1), the highest *B. theta* concentrations were driven by septic systems in the watershed, similar to the full watershed models, but with a tipping point of 3,927 septic systems in the watershed, much higher than the full watershed CART models. The highest concentrations of *E. coli* in the reduced watersheds (Fig. S1) were associated with potassium levels greater than 0.91 $mg \cdot L^{-1}$.

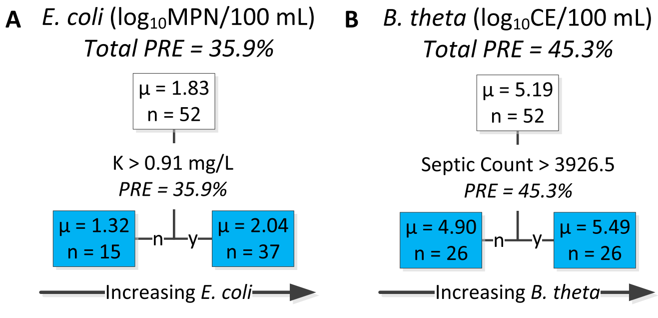

Fig. S1. CART analyses for log-transformed (*A*) *E. coli* and (*B*) *B. theta* concentrations as dependent variables and land use, nutrient, chemical, hydrologic, and environmental parameters as independent variables in reduced watersheds ($n = 52$).

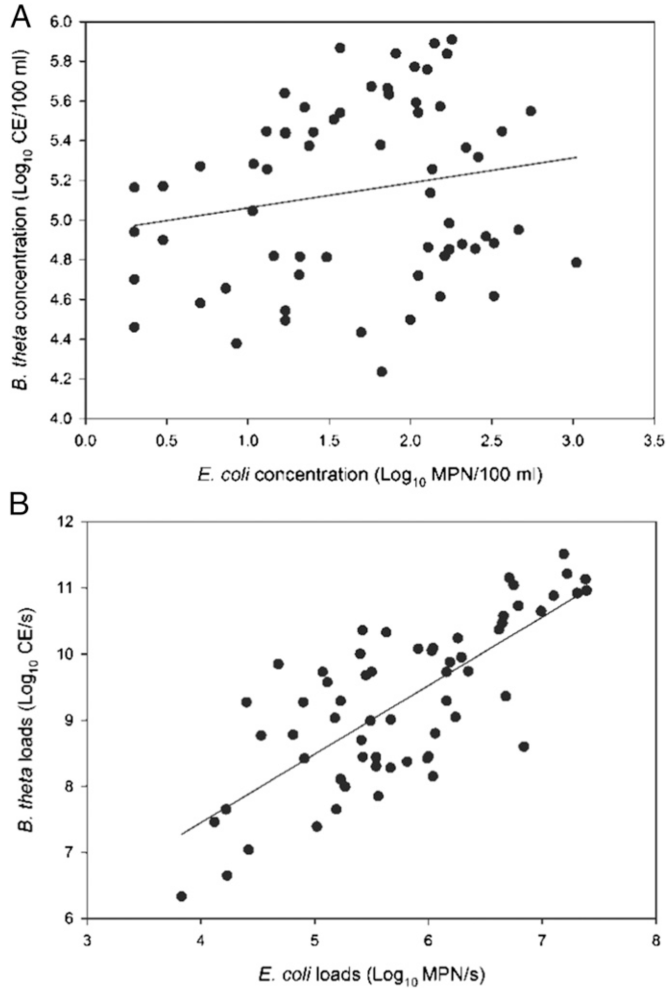

Fig. S2. Scatter plots of *B. theta* versus *E. coli* (*A*) concentrations ($n = 64$) and (*B*) loads ($n = 63$).

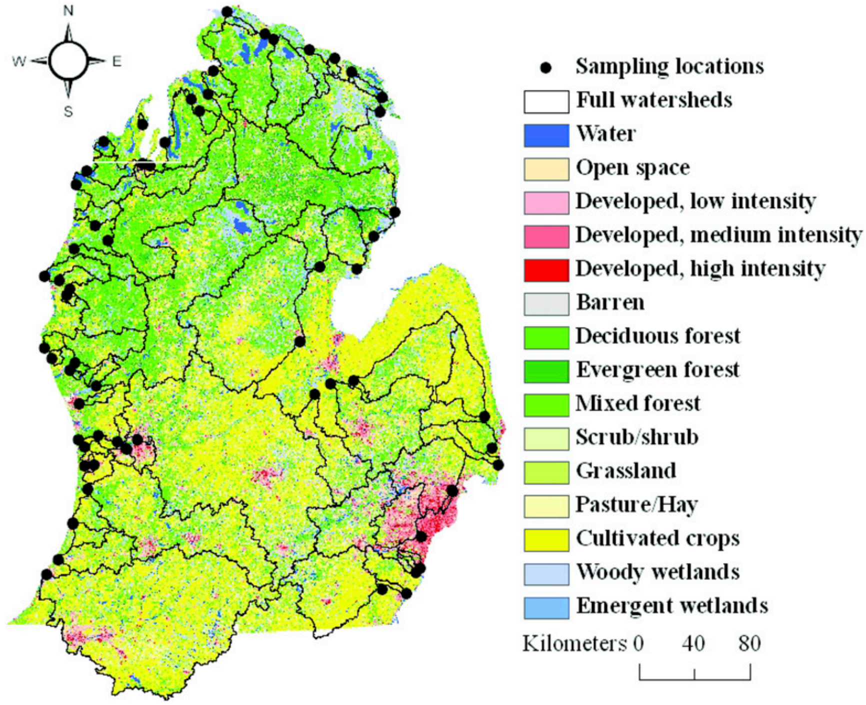

Fig. S3. Watersheds of sampled river systems that drain Michigan's Lower Peninsula and states to the south, colored by 2006 NLCD land use classes.

Table S1. Description of land, discharge, and *E. coli* and *B. theta* concentrations at 64 Michigan rivers under baseflow

River system	*E. coli*, MPN·100 mL^{-1}	*B. theta*, CE·100 mL^{-1}	Discharge, $m^3 \cdot s^{-1}$	Area, km^2	Urban, %	Agriculture, %	Rangeland, %	Forest, %	Water, %	Wetland, %	Barren, %	No. of septic systems	Average annual WWTP discharge, MGD
St. Joseph River	1.0	5.3	57.3	11,061	14.3	59.5	1.1	10.4	2.4	12.2	0.2	93,643	24.2
Paw Paw River	1.8	5.7	8.0	1,027	11.5	47.5	2.6	21.1	1.4	15.6	0.3	18,106	2.3
Kalamazoo River	1.6	5.9	43.6	5,002	14.1	47.8	1.8	21.5	2.1	12.4	0.4	102,948	49.2
Grand River	1.5	5.5	50.7	12,854	12.7	55.3	1.0	16.6	1.5	12.7	0.3	246,033	88.1
Muskegon River	1.8	5.4	37.8	6,418	7.6	19.6	9.6	40.5	3.9	18.6	0.1	47,115	6.7
White River	2.1	5.8	7.7	1,049	5.2	20.3	9.7	49.7	0.7	14.3	0.1	6,224	0.0
Pere Marquette River	2.2	5.8	12.2	1,790	5.0	9.3	8.3	61.6	1.2	14.5	0.1	6,094	0.0
Big Sable River	0.7	5.3	2.9	476	5.3	11.5	8.0	52.4	5.0	17.0	0.8	1,975	0.0
Little Manistee River	1.4	5.4	4.5	526	4.7	3.9	12.5	68.6	0.7	9.6	0.1	1,228	0.0
Manistee River	0.2	4.7	42.8	3,559	5.7	9.6	15.9	56.4	1.4	10.9	0.1	11,951	0.0
Bear Creek	2.1	5.9	3.0	350	6.3	13.8	20.1	37.6	2.3	19.7	0.1	1,656	0.0
Betsie River	2.0	5.8	22.8	618	8.1	7.6	13.1	46.2	9.9	15.0	0.1	6,140	0.0
Platte River	0.2	5.2	4.8	471	6.6	9.9	13.4	56.2	7.5	6.1	0.2	4,697	0.0
Boardman River	1.1	5.3	6.7	716	10.8	10.4	18.8	46.7	2.1	10.9	0.2	9,824	4.0
Elk-Torch River	0.2	4.5	12.8	1,308	7.6	14.4	13.6	45.4	11.3	7.4	0.2	9,595	0.0
Cheboygan River	0.2	4.9	26.1	2,317	6.4	8.2	11.6	51.0	8.1	14.5	0.2	11,660	0.0
Black River	1.2	4.8	11.6	1,509	5.5	4.4	12.1	47.1	3.9	27.0	0.1	2,747	0.3
Thunder Bay River	0.5	4.9	12.6	2,241	6.3	11.0	8.8	40.2	2.7	31.0	0.1	7,283	1.6
Au Sable River	1.2	5.6	32.5	5,287	8.4	3.2	14.5	58.9	2.0	12.8	0.1	23,156	0.6
Au Gres River	2.1	5.3	1.1	987	6.4	23.3	7.6	37.7	2.2	22.0	0.7	4,573	0.4
Rifle River	1.6	5.5	5.0	858	9.3	16.5	8.9	44.2	1.6	19.4	0.1	6,322	1.6
Black River	1.1	5.4	0.66	1,250	6.2	74.2	1.2	10.6	0.1	7.5	0.1	6,549	4.1
Pine River	2.6	5.4	0.05	440	9.0	46.5	3.2	33.3	0.3	7.5	0.1	6,430	0.0
Belle River	2.0	5.6	1.4	512	9.5	59.7	1.7	19.0	0.3	9.7	0.2	6,416	3.2
Clinton River	2.0	5.5	n/a	1,880	51.5	20.2	1.3	14.9	2.8	8.6	0.7	95,581	152.9
River Rouge	2.7	5.5	0.32	1,033	82.9	5.4	0.5	7.2	0.7	2.9	0.3	41,345	10.9
Huron River	1.9	5.8	7.7	2,298	32.5	24.5	1.2	21.8	4.2	15.1	0.6	117,122	43.9
Raisin River	1.4	5.4	4.7	2,683	10.8	67.4	0.8	11.1	1.4	8.3	0.2	41,999	14.0
South Branch Black River	2.3	5.9	1.1	313	9.1	45.8	4.4	22.8	1.2	16.5	0.2	4,099	1.7
North Branch Black River	2.2	5.6	1.5	398	7.0	43.6	5.6	24.8	1.7	17.1	0.2	4,073	0.6
Macatawa River	1.3	5.6	0.16	292	23.5	67.8	0.7	4.0	0.2	3.1	0.9	5,960	1.1
Pine Creek	2.1	4.9	0.27	48	48.4	30.9	1.1	12.1	0.3	6.1	1.1	5,487	0.0
Pigeon River	2.7	5.0	0.08	102	11.0	66.0	2.0	15.3	0.1	5.1	0.5	1,761	0.0
Rush Creek	2.3	5.4	0.12	152	56.4	31.5	0.4	7.6	1.1	2.3	0.6	12,607	0.0
Buck Creek	2.2	4.6	0.64	3	91.3	0.0	0.6	1.4	0.9	5.8	0.0	0	0.0
Sand Creek	2.3	4.9	0.31	142	19.1	60.8	0.8	11.3	0.2	7.6	0.3	3,780	0.0
Bass River	2.1	5.1	0.20	127	11.1	63.6	2.1	16.0	0.2	6.6	0.5	2,594	0.0
Little Pigeon Creek	2.5	4.9	0.03	14	18.9	16.4	6.2	41.9	0.0	16.3	0.3	167	0.0
Black Creek	3.0	4.8	0.66	136	14.9	34.8	5.3	29.9	4.8	10.1	0.2	4,197	0.0
Silver Creek	1.7	4.4	0.37	41	11.7	0.6	15.2	63.7	4.2	4.4	0.2	643	0.0
Flower Creek	2.5	4.6	0.34	79	10.2	45.6	10.7	27.7	0.6	3.4	1.8	580	0.0
Stony Lake Outlet	0.5	5.2	1.3	160	10.1	37.7	11.7	35.1	1.0	4.1	0.3	1,635	0.0
Swan Creek	2.5	4.9	0.34	54	5.5	57.9	8.2	15.5	1.3	11.5	0.0	248	0.0
Lincoln River	2.2	5.0	0.66	215	5.6	33.2	11.8	30.6	2.1	16.5	0.2	1,087	0.0
Crystal River	0.7	4.6	1.6	110	4.7	3.4	8.7	53.7	23.7	3.3	2.4	518	0.0

Table S1. Cont.

River system	*E. coli*, MPN·100 mL^{-1}	*B. theta*, CE·100 mL^{-1}	Discharge, $m^3 \cdot s^{-1}$	Area, km^2	Urban, %	Agriculture, %	Rangeland, %	Forest, %	Water, %	Wetland, %	Barren, %	No. of septic systems	Average annual WWTP discharge, MGD
Belangers Creek	1.3	4.7	0.08	25	6.7	38.4	12.7	30.8	1.5	9.9	0.0	188	0.0
Mitchell Creek	2.2	4.8	0.29	38	28.3	22.8	16.3	19.4	0.2	13.0	0.1	1,608	0.0
Jordan River	0.9	4.4	4.1	174	3.2	7.8	6.5	70.7	0.0	11.8	0.1	367	0.0
Monroe Creek	1.2	4.5	0.08	27	4.2	22.3	8.8	44.5	2.2	18.1	0.0	113	0.0
Boyne River	1.2	5.4	1.7	199	8.3	16.1	10.8	54.5	0.6	9.4	0.2	1,647	0.0
Bear River	1.5	4.8	1.6	293	6.4	13.3	7.0	48.5	6.6	18.1	0.2	2,276	0.0
Carp River	1.0	5.0	1.8	119	6.2	8.6	7.7	22.0	7.0	48.3	0.1	375	0.0
Ocqueoc River	0.9	4.7	2.4	369	4.7	6.5	11.5	43.4	2.2	31.4	0.3	576	0.0
Trout River	1.3	4.8	0.41	82	4.6	13.5	9.5	28.8	0.1	43.1	0.3	291	0.0
Little Trout River	2.4	4.9	0.06	28	5.4	27.8	7.5	14.3	0.1	44.6	0.3	95	0.0
Long Lake Creek	1.8	4.2	0.03	162	5.7	11.7	7.1	20.7	15.7	39.1	0.1	1,142	0.0
Tawas River	1.2	4.5	1.6	403	8.4	7.1	6.9	51.6	2.0	24.0	0.0	1,597	0.9
Harrington Drain	2.0	4.5	0.01	53	99.7	0.0	0.0	0.2	0.0	0.1	0.0	1,714	0.0
Marsh Creek	2.4	5.3	0.13	78	72.0	4.7	1.7	15.4	0.0	5.9	0.2	1,838	0.0
Sandy Creek	2.2	4.9	0.02	82	26.2	58.7	1.3	10.6	0.0	2.7	0.4	1,506	0.5
Cass River	1.2	5.4	2.0	2,174	6.9	57.4	2.2	19.7	0.2	13.5	0.1	15,999	2.8
Flint River	1.9	5.7	6.3	3,206	21.0	40.6	2.0	24.1	1.6	10.4	0.3	141,168	75.6
Shiawassee River	2.0	4.7	4.4	1,517	15.7	52.5	0.7	17.0	2.2	11.4	0.4	37,043	13.5
Tittabawassee River	1.9	5.6	17.5	6,211	8.6	32.8	7.3	30.6	1.5	19.1	0.2	55,635	16.3

n/a, not applicable.

Table S2. Descriptive statistics of physical, chemical, and hydrologic variables measured during baseflow conditions at 64 rivers

Parameter	Count	Minimum	Mean	Maximum	SD	5th percentile	95th percentile
Ammonia, μg·L^{-1}	63	0.0	23.6	280.0	45.6	0.0	98.5
Calcium, mg·L^{-1}	63	30.0	62.4	160.6	21.6	33.8	98.2
Chlorine (Cl^-), mg·L^{-1}	63	3.4	42.3	291.8	54.4	5.9	174.8
Dissolved oxygen, mg·L^{-1}	64	5.9	9.8	13.3	1.7	7.2	12.2
Dissolved organic carbon (NPOC), mg·L^{-1}	63	1.6	6.1	26.8	4.2	2.1	15.6
Magnesium, mg·L^{-1}	63	7.0	18.4	34.2	6.3	10.3	29.1
Nitrate/nitrite (NO_x), μg·L^{-1}	64	0.0	858.3	5,638.9	1,310.3	0.0	4,095.6
Pheophytin corrected chlorophyll a, μg·L^{-1}	59	0.0	0.8	4.4	1.0	0.1	3.4
pH	63	7.9	8.2	8.4	0.1	8.0	8.4
Potassium, mg·L^{-1}	63	0.4	2.2	9.8	1.9	0.5	6.0
Sodium, mg·L^{-1}	63	3.0	27.0	199.3	36.9	3.4	113.0
Soil hydraulic conductivity (K_{sat}), m·d^{-1}	64	0.5	2.2	4.7	1.1	0.6	4.2
Specific conductance, μS·cm^{-1}	63	257.0	527.0	1,589.0	264.2	265.2	1,039.8
Soluble reactive P, μg·L^{-1}	64	0.9	23.3	266.0	45.0	2.1	87.0
Sulfate, μg·L^{-1}	63	2.4	32.1	169.8	30.5	5.6	89.6
Total dissolved N, μg·L^{-1}	64	0.0	1,423.3	6,033.7	1,346.5	337.6	5,414.1
Total dissolved P, μg·L^{-1}	64	3.1	25.2	292.3	38.6	3.9	58.0
Total N, μg·L^{-1}	64	81.8	1,082.1	5,583.1	1,129.3	110.8	3,610.6
Total P, μg·L^{-1}	64	7.7	37.8	395.5	52.4	8.9	102.5
Total chlorophyll a, μg·L^{-1}	59	0.1	1.6	7.8	1.9	0.2	7.4
Precipitation,† mm							
6 h	64	0.0	0.1	9.2	1.2	—	—
12 h	64	0.0	1.9	77.9	10.2	0.0	7.9
18 h	64	0.0	3.4	78.6	11.6	0.0	30.8
24 h	64	0.0	4.4	78.6	11.9	0.0	31.0
2 d	64	0.0	6.0	78.6	11.9	0.0	31.0
3 d	64	0.0	7.7	80.1	13.7	0.0	34.2
4 d	64	0.0	8.3	80.5	13.5	0.0	34.2
6 d	64	0.0	9.0	87.3	14.1	0.0	34.2
8 d	64	0.0	11.6	92.6	16.5	0.0	57.2
Discharge, m^3·s^{-1}	63	0.0	6.7	57.3	12.5	0.0	43.4
Water temperature, °C	64	7.0	13.1	17.5	2.6	8.2	16.6

†Precipitation measured at hourly averages from 16-km^2 NEXRAD cells and reported in cumulative millimeters per time.

Table S3. Summary of chemical and nutrient methods

Assay	Method description	Refs.
Ammonia, μg·L^{-1}	Phenate method	Standard methods 4500-NH_3-G (68)
Calcium, mg·L^{-1}	Flame atomic absorption spectrophotometry	65
Chlorine (Cl^-), mg·L^{-1}	Dionex membrane-suppression ion chromatography	65, 66
Magnesium, mg·L^{-1}	Flame atomic absorption spectrophotometry	65
Nitrate/nitrite, μg·L^{-1}	Cadmium reduction	Standard methods 4500-NO_3-E (68)
Pheophytin corrected chlorophyll a, μg·L^{-1}	Fluorometry with pheophytin correction following ethanol extraction	Standard methods 10200.H (68)
pH	Hydrolab multisonde	66
Potassium, mg·L^{-1}	Flame atomic absorption spectrophotometry (0.5% HNO_3 preservative)	66
Sodium, mg·L^{-1}	Flame atomic absorption spectrophotometry (0.5% HNO_3 preservative)	66
Soluble reactive phosphorus, μg·L^{-1}	Ascorbic acid method	Standard methods 4500-P.E. (68)
Sulfate (SO_4), μg·L^{-1}	Dionex membrane- suppression ion chromatography	66
Total dissolved nitrogen, μg·L^{-1}	Second derivative spectroscopy following persulfate digestion	67
Total dissolved phosphorus, μg·L^{-1}	Ascorbic acid method following persulfate digestion	Standard methods 4500-P.E and 4500-N.C (68)
Total nitrogen, μg·L^{-1}	Second derivative spectroscopy following persulfate digestion	67
Total phosphorus, μg·L^{-1}	Ascorbic acid method following persulfate digestion	Standard methods 4500-P.E and 4500-N.C (68)
Total chlorophyll a, μg·L^{-1}	Fluorometry following ethanol extraction	Standard methods 10200.H (68)

Table S4. Land use summary for full watersheds, reduced watersheds, and 60-m buffers

Scale parameter	Minimum	Mean	Maximum	SD
Full watershed[†]				
Area, km^2	2.88	1,377	12,854	2,431
Estimated septic systems	0.0	19,579	246,033	41,902
Septic density, no. per km^2	0.0	15.7	113.7	19.5
Population density, persons per km^2	7	131	1,597	281
Population density on WWTP	0	98	1,589	279
Population density on septic	7	114	1,567	236
Impervious surface, km^2	0.41	5.13	56.9	9.8
Urban, %	3.16	16.7	99.7	0.21
Agriculture, %	0.0	28	74.2	0.22
Open, %	0.0	6.97	20.1	0.05
Forest, %	0.19	31.4	70.7	0.18
Water, %	0.0	2.68	23.7	0.04
Wetland, %	0.07	14	48.3	0.1
Barren, %	0.0	0.31	2.45	0.0
Reduced watershed[‡]				
Area, km^2	0.15	366	4,065	630
Estimated septic systems	0.0	22,299	246,033	45,592
Septic density, no. per km^2	0.0	16.1	114	18.8
Population density, persons per km^2	7	147	1,597	306
Population density on WWTP	0	115	1,589	307
Population density on septic	7	124	1,567	259
Impervious surface, km^2	0.4	7.5	55.9	13.6
Urban, %	3.1	21.3	99.7	26.2
Agriculture, %	0.0	27.2	77.4	24
Open, %	0.0	6.16	18.8	5.27
Forest, %	0.0	29	71.2	19.5
Water, %	0.0	1.61	15.4	3.32
Wetland, %	0.0	13.9	47.9	12.1
Barren, %	0.0	0.77	31.1	3.87
60-m riparian buffer[‡]				
Area, km^2	0.06	46	497	78.3
Estimated septic systems	0.0	2,672	28,256	5,596
Septic density, no. per km^2	0.0	15	104	21
Population density, persons per km^2	7	124	1,567	259
Population density on WWTP	0	0	0	0
Population density on septic	0	32	105	24
Impervious surface, km^2	0.0	5.5	42.7	9.64
Urban, %	0.0	18.9	98.3	23
Agriculture, %	0.0	21.4	72.1	21.7
Open, %	0.0	3.64	19.4	3.8
Forest, %	0.0	22.1	62.6	14.9
Water, %	0.0	6.09	63.2	12
Wetland, %	0.0	27.3	76.3	17.9
Barren, %	0.0	0.59	24.9	3.12

[†]Entire upstream drainage area including lakes (n = 64).
[‡]Watersheds were defined as the total upstream area to the nearest lake draining to each respective river sampling point (n = 52).

Table S5. Anderson level 1 land use classifications and descriptions

Classification	Description	Examples	Associated NLCD classifications (code)
Urban	Intensive use with structures covering the majority of land	Cities, shopping, industrial, and commercial centers	Developed open space (21) Developed low intensity (22) Developed medium intensity (23) Developed high intensity (24)
Agricultural	Land used for food production	Pasture, row crop, orchards, confined feeding operations	Pasture and hay (81) Cultivated crops (82)
Open	Predominant natural vegetation is grass or shrubs	Herbaceous, shrub, brush	Shrub and scrub (52) Grassland and herbaceous (71)
Forest	Closed canopy at least 10% from timber quality trees	Deciduous, coniferous, and mixed forested	Deciduous forest (41) Evergreen forest (42) Mixed forest (43)
Water	Area predominantly covered by water throughout year	Streams, lakes, bays, and reservoirs	Water (11)
Wetland	Land with water table near land surface for significant portion of year	Marshes, swamps, perched bogs	Woody wetland (90) Emergent herbaceous wetland (95)
Barren	Land that has less than one-third vegetative cover	Beaches, exposed rock, gravel pits	Barren (31)