Supporting Information for

Trends in Drinking Water Nitrate Violations across the United States

Michael Pennino[1], Jana E. Compton[1], and Scott G. Leibowitz[1]

[1]US EPA, Office of Research and Development, National Health and Environmental Effects Research Laboratory, Western Ecology Division, Corvallis, OR, USA

DESCRIPTION OF SUPPORTING INFORMATION

Methods. Further details on methods for data analysis

List of Acronyms

Table S1. Metrics for nitrate violations obtained from EPA's SDWIS.

Table S2. Temporal trend statistics for proportion of systems in violation for each state with violations.

Table S3. Contingency table showing how the number of systems in violation or not in violation in 2014 and 2016 compare to the number of systems with treatment or not with treatment in 2013 and 2015, respectively.

Table S4. The number and percent of systems in violation and people served by water source, PWS Type, Owner Type, or Facility Type during period of 1994 through 2016.

Table S5. Logistic regression results for comparison of conterminous United States counties with and without violations.

Figure S1. Diagram for relationship between PWS, its facilities and population served. In this example the PWS consists of four facilities that serve three separate populations.

Figure S2. Time trends for (a) the number of systems in violation each year for the top eight contaminants and (b) percent of all violations due to nitrate each year.

Figure S3. Average concentration for samples taken by public water systems which exceed the 10 mg/L MCL.

Figure S4. Comparison of systems with and without nitrate removal treatment technologies in 2014 and 2016.

Figure S5. Maximum consecutive quarters that (a) groundwater (GW), (b) surface water (SW), or (c) all systems were in violation for each U.S. state, divided by four to convert to years.

Figure S6. Box plots of people served by systems in violation for the nitrate MCL by (a) PWS type and (b) water source.

Figure S7. Number of nitrate violations by public water system size (number of people served).

Figure S8. Total number of violations, percent of all violations, and average annual population served by systems in violation by (a & b) Water Source, (c & d) PWS Type, and (e & f) Owner Type.

Figure S9. Comparison of population served per system with number of violations per system for groundwater (a) and surface water systems (b) and comparison of population served per system with maximum years that system has been in violation for groundwater systems (c) and surface water systems (d). PWS = public water system. The red line is a LOESS curve (produced using the xyplot from the R lattice package).

Figure S10. Total number of violations by county separated by (a) groundwater and (b) surface water. These violations are based on the number of unique PWSs in violation each year per county.

Figure S11. Mean nitrate concentrations per state based on systems violating the MCL between 1994 and 2016 (thus no concentrations below 10 mg/L are included).

Figure S12. Monitoring and Reporting (MR) violations by state as (a) total number of violations and (b) percent of all violations from 1994-2016.

Figure S13. Comparison of U.S. counties with (Violator) and without violations (Non-Violator). Error bars are 95% confidence intervals for the mean.

Methods

To download maximum contaminant level violations data, within the SDWIS Federal Reports Advanced Search tool,[1] we selected both the Average (02) and Single Sample (01) Maximum Contaminant Level Violation Filter options, set Violation Rule to Nitrates, and made sure Activity Status was set to All. We determined the year and quarter each violation occurred based on the compliance period begin date.

To download monitoring and reporting violations data, within the SDWIS Federal Reports Advanced Search tool,[1] we selected both the Monitoring, Regular (03) and Monitoring, Check/Repeat/Confirmation (04) Violation Filter options, set Violation Rule to Nitrates, and made sure Activity Status was set to All.

List of Acronyms

SDWIS = Safe Drinking Water Information System
MCL = Maximum Contaminant Level
CAFOs = concentrated animal feeding operations
PWS = Public Water System
SW = Surface Water
GW = Groundwater
CWS = community water system
NTNCWS = non-transient non-community water system
TNCWS = transient non-community water system
MR = monitoring/reporting

Table S1. Study metrics for nitrate violations obtained from EPA's SDWIS.

Metric	Calculation Method
Temporal Metrics	
Number of Systems in Violation	Number of PWSs with at least one violation in a particular calendar year. No PWS in violation is counted more than once per year, even if it has a violation in multiple quarters per year.
Percent of Systems in Violation	Number of systems in violation for a calendar year divided by the inventory of active systems for that fiscal year (beginning on July 1 and ending on June 30). Inventory data is the number of systems active at least part of that fiscal year and this data (for 1994-2016) came from a request to the EPA's Office of Ground Water and Drinking Water. Using the calendar year for the number of violations and the fiscal year for the inventory may lead to some errors since they only have a 6-month period in common, however, this likely does not affect long-term trends.
Number of People Served by Systems in Violation	Sum of population served for each PWS in violation that year. Population served is provided for each PWS in SDWIS
Max Violation Duration	The maximum number of consecutive quarters in violation was calculated for each PWS between 1994 and 2016; number divided by four to convert to annual basis. Calculated by using the compliance period begin date to determine which quarter the violation occurred. Duration is not same as time between the compliance begin data and compliance end date, which would measure how long a single violation lasts.
Average Violation Duration	The average number of consecutive quarters in violation for each PWS, divided by four to convert to average consecutive years in violation. Averaged over all systems for either the 3-year moving average or for the entire 1994 to 2016 period. Also calculated using the compliance period begin date.
Number and Percent of First Time Violators and Repeat Violators	The number of systems in violation with no previous violations; repeat violators are the difference between total and first time violators. Percent repeat violators is number of repeat violators divided by total violators.
Groundwater or Surface water Violations	Number of systems in violation listed as having their source water as either groundwater or surface water.
Number of Violations by PWS Type	PWSs were filtered by PWS Type to calculate the number of systems in violations per year by PWS Type. PWS Types are: community water systems, transient non-community water systems, and non-transient non-community water systems.
Number of Violations by Owner Type	PWSs were filtered by Owner Type to calculate the number of systems in violation per year by Owner type. Owner Types are: Federal Government, Local Government, Native American,

	Private, Public/Private, and State Government. For a PWS with multiple owner types, the most common Owner Type was chosen.
Number and percent of systems with nitrate removal treatment	Facility reports were downloaded from SDWIS[1] for the years 2013 and 2015. PWSs were categorized as having nitrate removal technology if they had a Treatment Process of Reverse Osmosis, Ion Exchange, Electrodialysis, or Distillation. The percent of systems with nitrate treatment in 2014 was then calculated as the number of systems with treatment in 2013 divided the total number of active systems in 2014. 2013 treatment was used for 2014 systems to ensure that all systems had the treatment listed, in case a treatment was added in the middle or end of the year. A similar calculation was done for 2016 systems.
Geographic Metrics	
Mean Annual Number of Violations per State or County	Average annual sum of number of systems in violation by state or county, from 1994 to 2016. Within SDWIS, the county served by each PWS was obtained from the Geographic Area record, which is separate from the violations data. The county served was merged into the violations data by the PWS identifier, ensuring that each PWS was associated with only one county.
Mean Annual Population Served per State or County	Average annual number of people served by each PWS in violation per state or county from 1994 to 2016.
Mean Annual Number of GW or SW Violations per State	Average annual sum of number of systems in violation by state and by groundwater (GW) or surface water (SW), from 1994 to 2016.
Mean Annual Percent of Systems in Violation by State	The sum of violations per state and year was divided by the inventory of active systems for that state and year and multiplied by 100. Then this percent in violation per state and year was averaged over all years, 1994 to 2016.
Mean Annual Percent of People Served per State	The sum of people served per state and year was divided by the inventory of all people served by active systems for that state and year and multiplied by 100. Then this percent of people served per state and year was averaged over all years, 1994 to 2016.
Max Violation Duration per State	Maximum consecutive quarters in violation for each PWS in each state: number divided by four to convert to annual basis.
Mean and Max Violation Concentration per State	Concentration for each MCL violation provided in SDWIS was used to calculate the average and maximum concentration per state between 1994 and 2016.

Mean Annual Monitoring and Reporting Violations per state	Average number of systems per year and per state with a monitoring or reporting (MR) violation between 1994 and 2016. Listed in SDWIS as "(3) Monitoring, Regular" and "(4) Monitoring, Check/Repeat/Confirmation." A MR violation occurs when a PWS fails monitor and or report water sampling results to their primacy agency (i.e., state).

* PWS = public water system, MCL = maximum contaminant level, SDWIS = Safe Drinking Water Information System.

Table S2. Temporal trend statistics for proportion of systems in violation for each state with violations.

State	Slope	p-Value	Range for Percent in Violation
Arizona	0.03	0.16	0.06-1.62
Arkansas	-0.02	0.61	0.08-0.15
California	0.07	0.00	0.13-1.63
Colorado	0.03	0.00	0.05-0.78
Connecticut	0.01	0.02	0.02-0.33
Delaware	-0.03	0.37	1.39-4.53
Florida	0.00	0.02	0.02-0.17
Georgia	0.00	0.04	0.04-0.08
Idaho	-0.01	0.06	0.2-0.6
Illinois	0.00	0.05	0.02-1.26
Indiana	0.02	0.17	0.04-0.93
Iowa	-0.06	0.00	0.31-1.97
Kansas	-0.05	0.02	1.1-3.69
Kentucky	-0.03	1.00	0.29-0.43
Maine	-0.01	0.13	0.05-0.29
Maryland	0.00	0.73	0.12-0.6
Massachusetts	0.00	0.13	0.06-0.27
Michigan	0.00	0.01	0.03-0.15
Minnesota	-0.01	0.01	0.06-0.43
Missouri	0.00	0.80	0.04-0.11
Montana	0.01	0.44	0.05-0.73
Nebraska	0.02	0.96	0.14-4.1
Nevada	0.00	0.15	0.16-0.68
New Hampshire	0.00	0.70	0.04-0.25
New Jersey	0.02	0.04	0.02-0.85
New Mexico	0.01	0.00	0.08-0.58
New York	0.00	0.60	0.01-0.18
North Carolina	0.00	0.63	0.01-0.23
North Dakota	-0.01	0.70	0.15-0.7
Ohio	-0.01	0.03	0.02-0.34
Oklahoma	-0.05	0.02	1.1-3.36
Oregon	0.01	0.00	0.04-0.39
Pennsylvania	-0.03	0.00	0.21-1
Rhode Island	0.01	0.81	0.2-0.84
South Carolina	0.00	0.11	0.06-0.14
South Dakota	0.01	0.13	0.13-0.76
Tennessee	0.00	NA	0.09-0.12
Texas	0.03	0.00	0.25-1.2
Utah	0.00	NA	0.1-0.11
Vermont	0.00	0.73	0.07-0.14
Virginia	0.00	0.74	0.02-0.12
Washington	0.00	0.09	0.02-1.19
West Virginia	-0.01	0.95	0.07-0.46
Wisconsin	-0.02	0.00	0.04-0.83
Wyoming	0.00	0.11	0.12-0.44

*Regression statistics are based on non-parametric Mann-Kendall and Sen's slope test, except for Tennessee and Utah which had only 2 years of data and so a linear model was used.

Table S3. Contingency table showing how the number of systems in violation or not in violation in 2014 and 2016 compare to the number of systems with treatment or not with treatment in 2013 and 2015, respectively. A similar contingency table for 2016 violations is also included that is restricted to those systems that were in violation in 2014. The values in the parenthesis are percentages of row sums except for the final column; these parenthetical values are the percent of systems with or without nitrate treatment, respectively.

2014 violators and non-violators, with and without treatment			
	In Violation	Not in Violation	Row Sum
Nitrate Treatment	85	20718	20803
	(0.41%)	(99.59%)	(13.9)
No Nitrate Treatment	439	128278	128717
	(0.34%)	(99.66%)	(86.1%)
Column Sum	524	148996	149509
	(0.35%)	(99.66%)	(Total Systems)
2016 violators and non-violators, with and without treatment			
Nitrate Treatment	72	22050	22122
	(0.325%)	(99.67%)	(15.0%)
No Nitrate Treatment	411	125231	125642
	(0.327%)	(99.67%)	(85.0%)
Column Sum	483	147281	147755
	(0.327%)	(99.68%)	(Total Systems)
2016 violators and non-violators based only on systems that were 2014 violators			
Nitrate Treatment	21	75	96
	(21.9%)	(78.1%)	(18.3%)
No Nitrate Treatment	184	244	428
	(43.0%)	(57.0%)	(81.7%)
Column Sum	205	319	524
	(39.1%)	(60.9%)	(Total Systems)

Table S4. The number and percent of systems in violation and people served by water source, PWS Type, Owner Type, or Facility Type during period of 1994 through 2016.

Category	Type	Number of Violations	% of Violations	% of Systems in Violation	Mean Population Served per Year	% of Population Served by Systems in Violation
Water Source	Ground water	11,736	95	3.6	270,660	35
	Surface water	560	5	0.2	491,943	65
PWS Type	CWS	4,629	38	1.4	711,095	93
	NTNCWS	1,743	14	0.5	20,472	2.7
	TNCWS	5,924	48	1.8	31,036	4.1
Owner Type	Federal government	56	0.5	NA	1,411	0.2
	Local government	3,087	25	NA	652,088	86
	Native American	21	0.2	NA	1,510	0.2
	Private	8,476	69	NA	99,070	13
	Public/Private	454	3.7	NA	4,243	0.6
	State government	202	1.6	NA	4,282	0.6

*CWS = Community water system, NTNCWS = Non-Transient non-community system, TNCWS = Transient non-community system

Table S5. Logistic regression results for comparison of conterminous United States counties with and without violations.

	Coefficient	p-value
Percent Cultivated	0.030	<0.001
Water table depth	0.108	<0.001
Soil permeability (m/h)	0.100	<0.001
Soil Organic matter (% by weight)	0.088	<0.001
Precipitation (30-year normal (1981-2010)	-0.080	<0.001
Percent agricultural drainage	-0.029	<0.001
Hortonian overland flow	-0.013	0.12
Nitrogen from manure (kg/hectare)	0.038	<0.001
Population Density (× 2)	-0.003	0.015
Percent semi-consolidated aquifer	-0.011	<0.001
Nitrogen from farm fertilizer (kg/hectare)	0.012	0.037
Percent developed land	0.058	<0.001
County area (sq. km × 1000)	0.086	<0.001

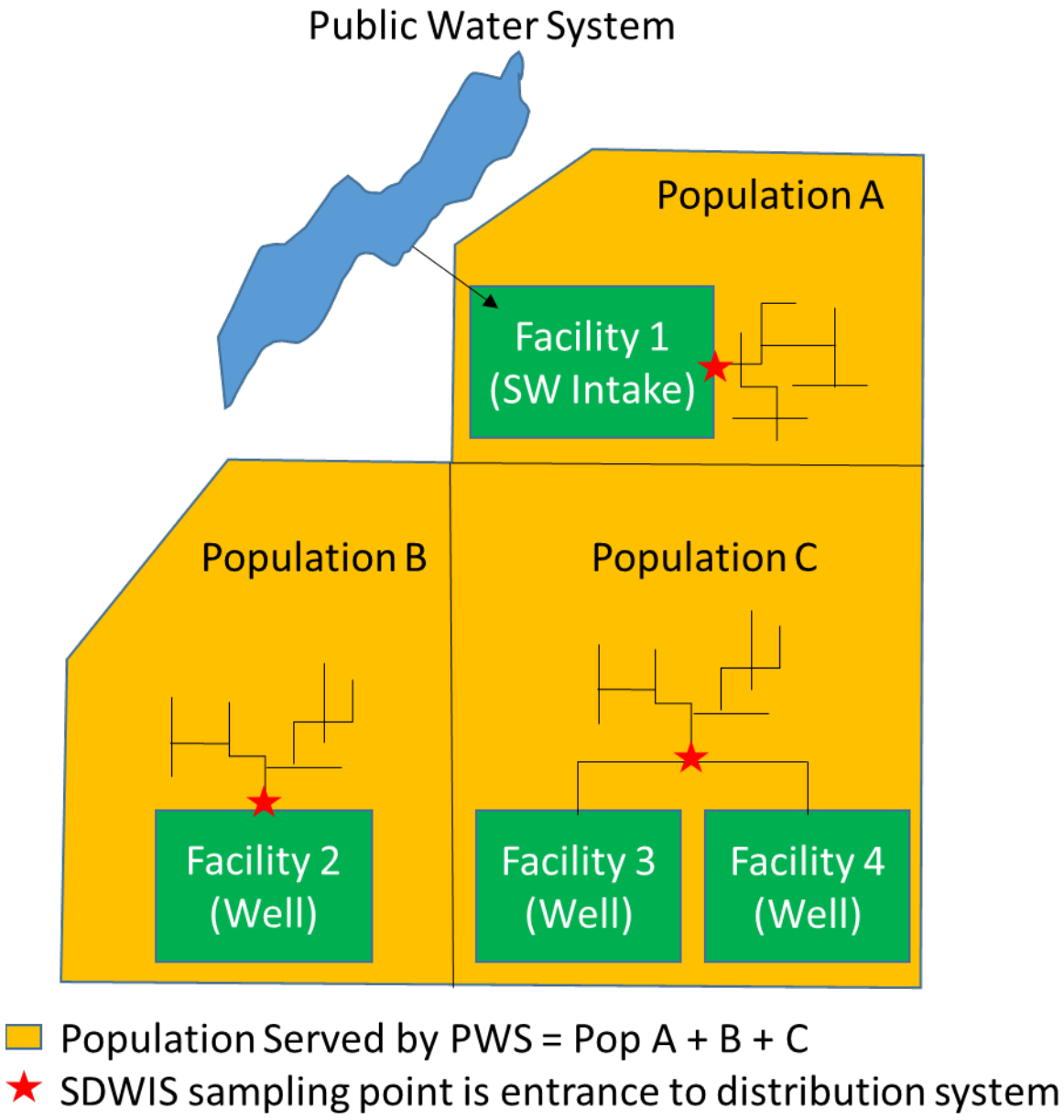


Figure S1. Diagram for relationship between a PWS, its facilities and population served. In this example the PWS consists of four facilities that serve three separate populations.

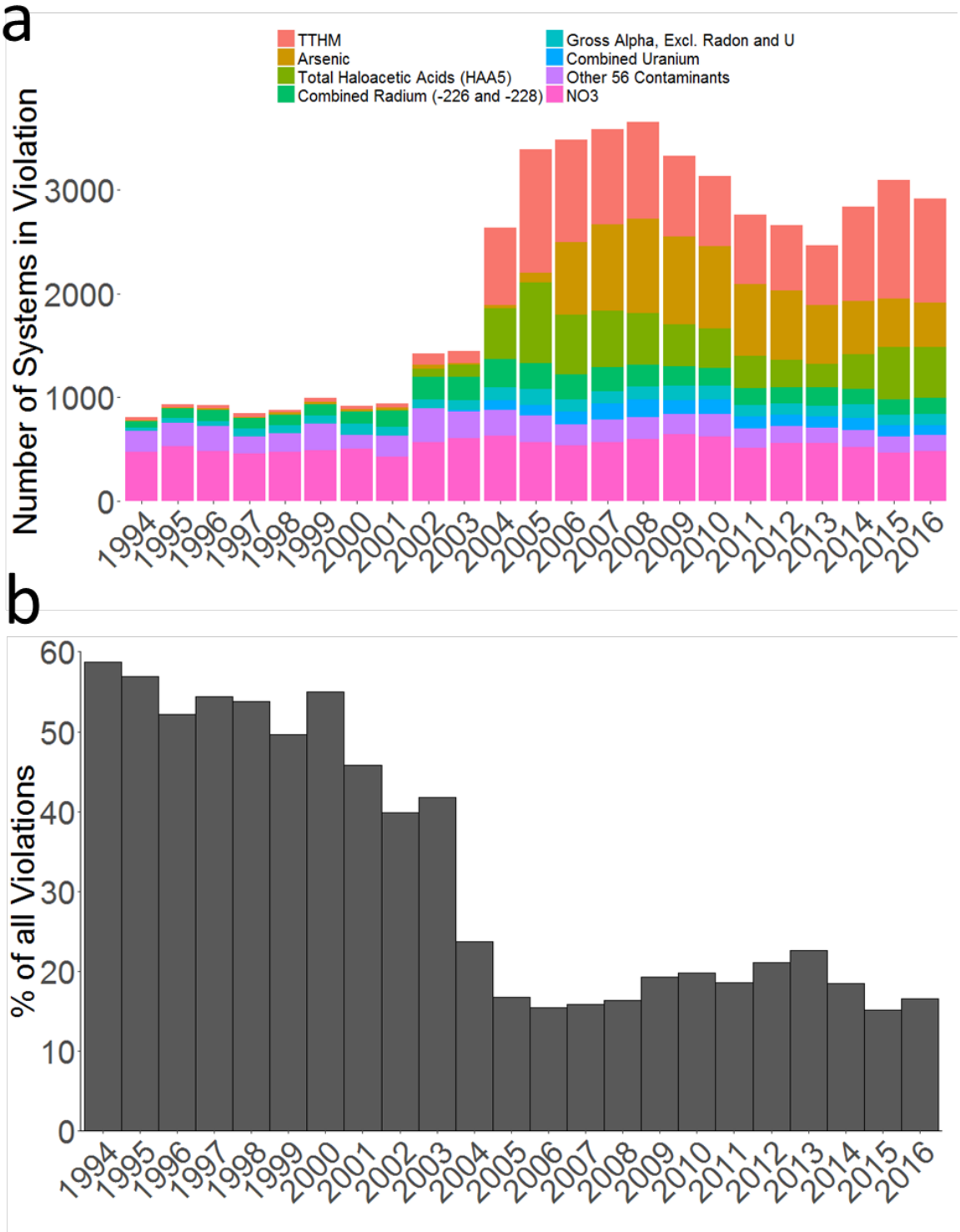


Figure S2. Time trends for (a) the number of systems in violation each year for the top eight contaminants and (b) percent of all violations due to nitrate each year. TTHM = total trihalomethanes.

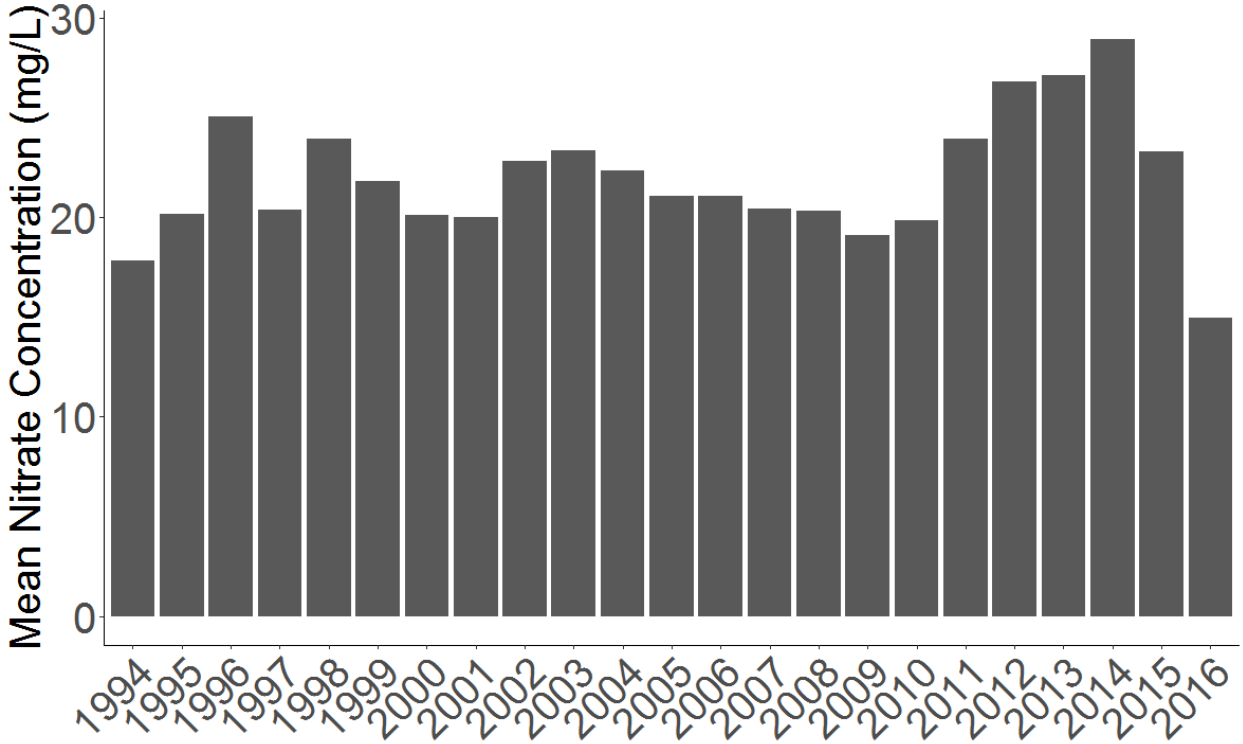


Figure S3. Average nitrate concentration for samples taken by public water systems which exceed the 10 mg/L MCL.

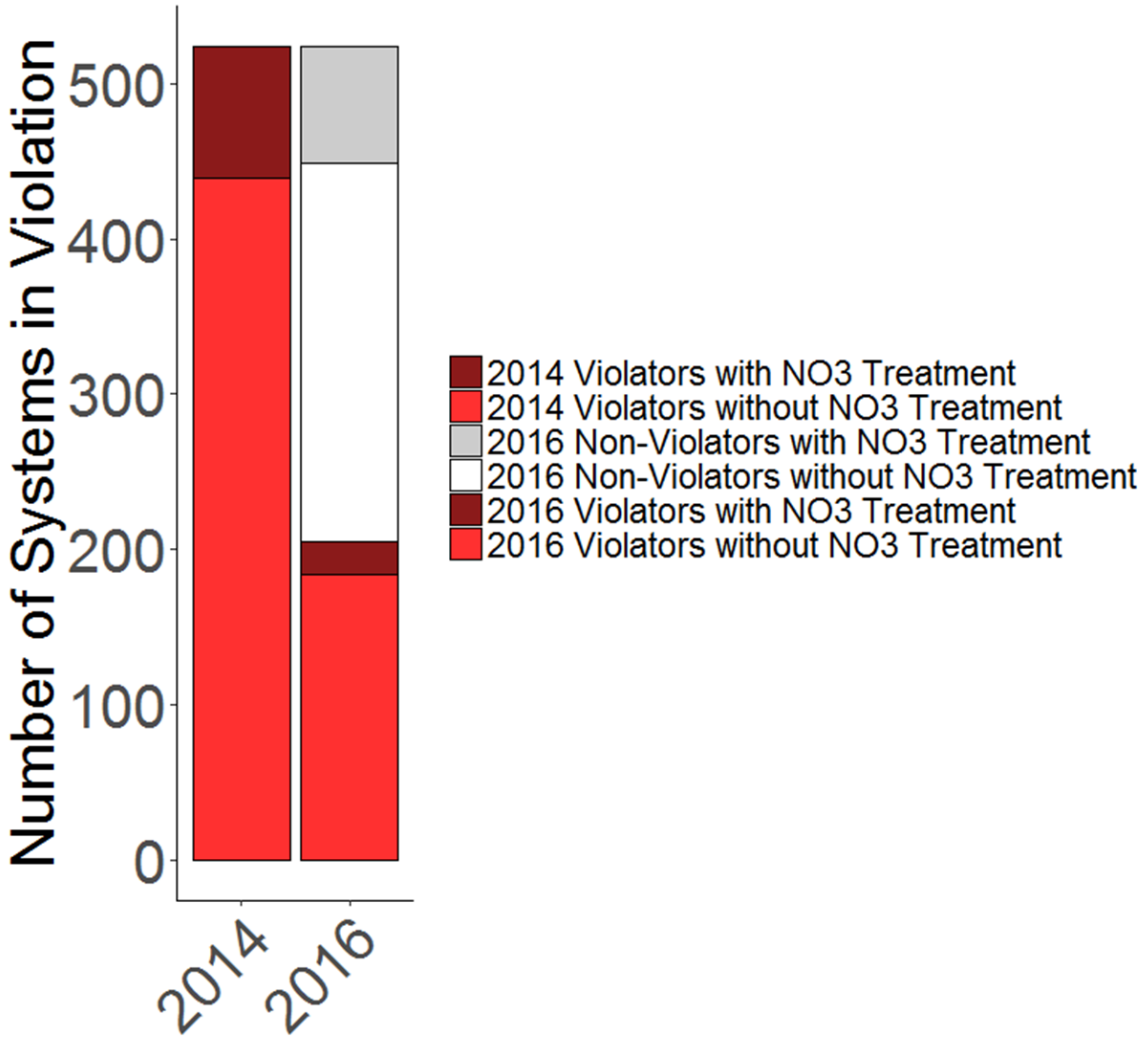


Figure S4. Comparison of systems with and without nitrate removal treatment technologies in 2014 and 2016. Note that the systems in violation in 2016 only include systems that were previously in violation in 2014, not new systems in violation, thus this number is less than the actual number of systems in violation in 2016.

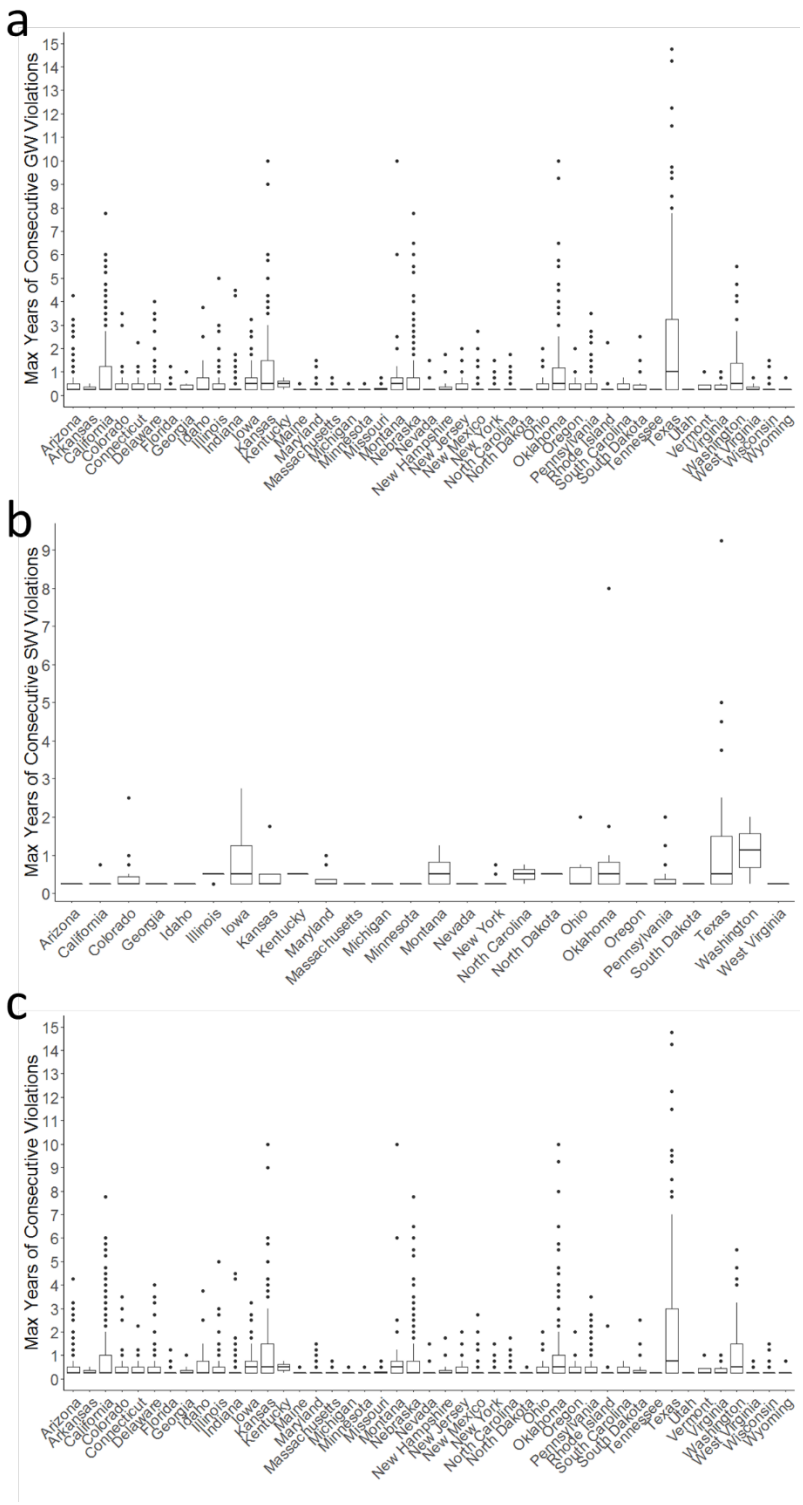


Figure S5. Maximum consecutive quarters that (a) groundwater (GW), (b) surface water (SW), or (c) all systems were in violation for each U.S. state, divided by four to convert to years.

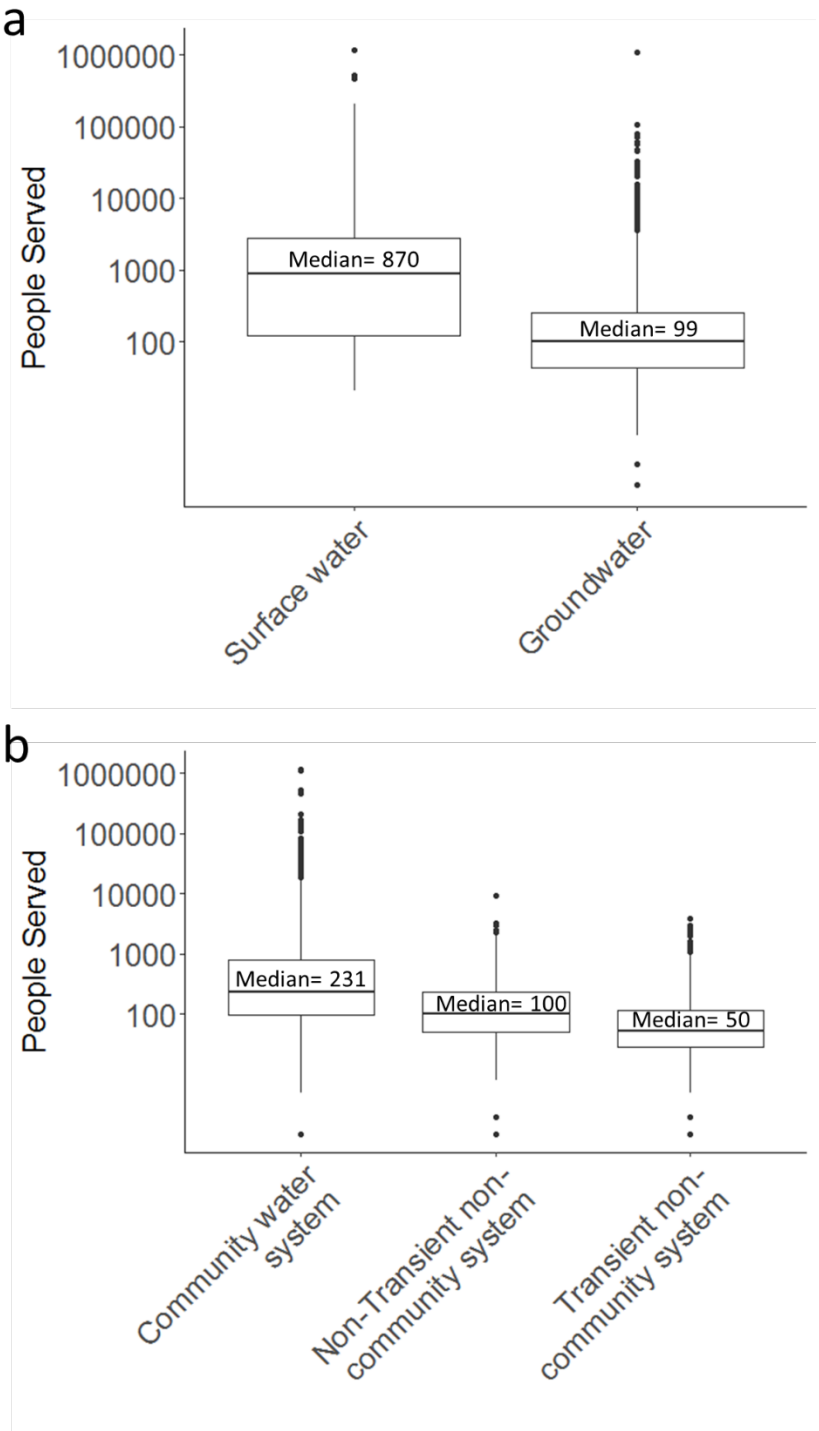


Figure S6. Box plots of people served by systems in violation for the nitrate MCL by (a) PWS type and (b) water source.

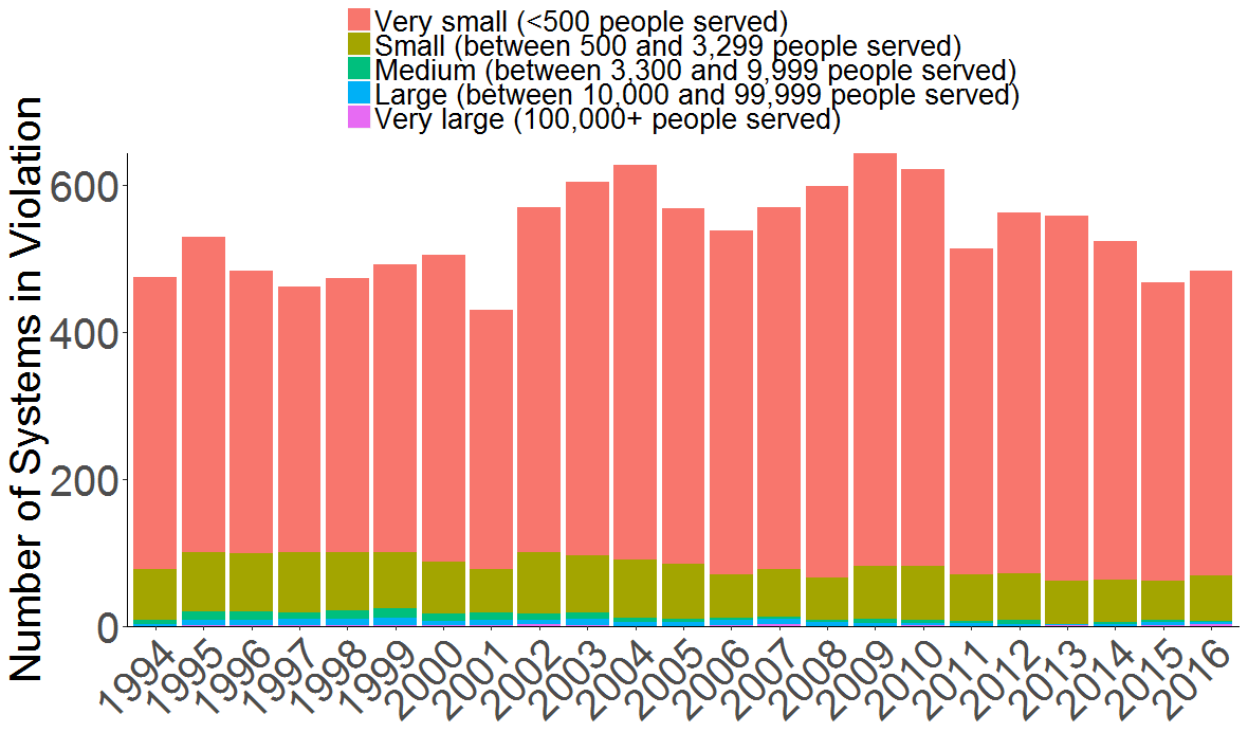


Figure S7. Number of nitrate violations by public water system size (number of people served).

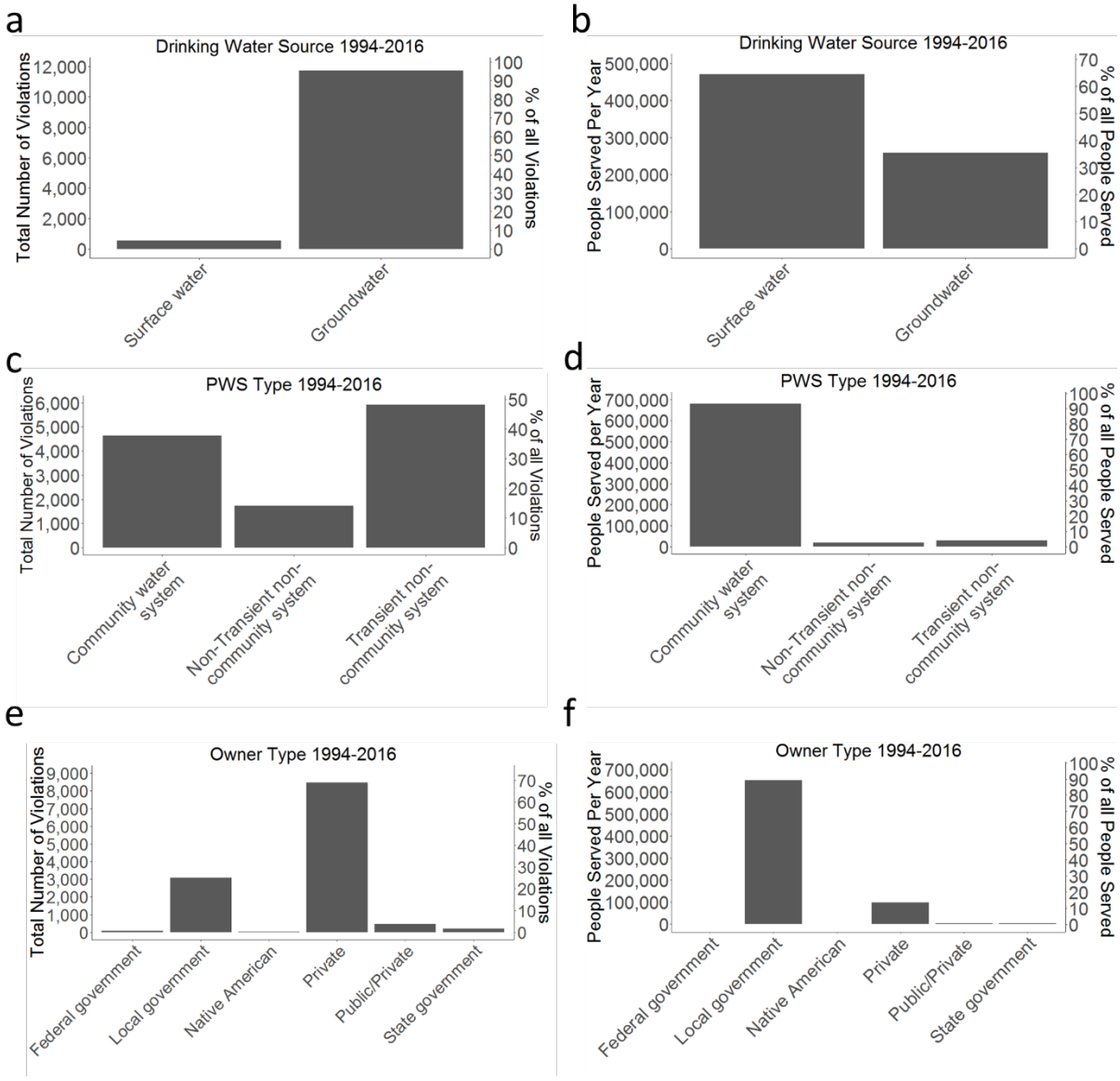


Figure S8. Total number of violations, percent of all violations, and average annual population served by systems in violation by (a & b) Water Source, (c & d) PWS Type, and (e & f) Owner Type.

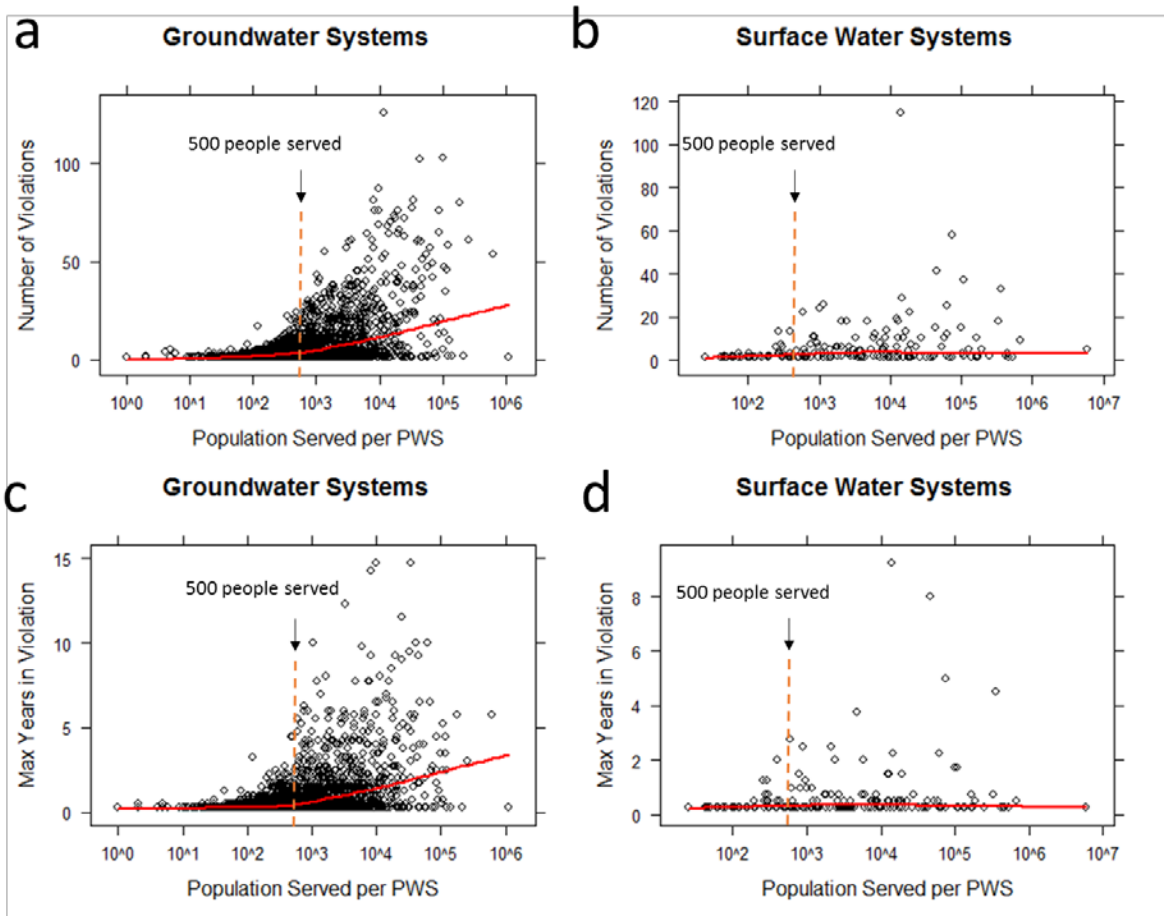


Figure S9. Comparison of population served per system with number of violations per system for groundwater (a) and surface water systems (b) and comparison of population served per system with maximum years that system has been in violation for groundwater systems (c) and surface water systems (d). PWS = public water system. The red line is a LOESS curve (produced using the xyplot from the R lattice package).

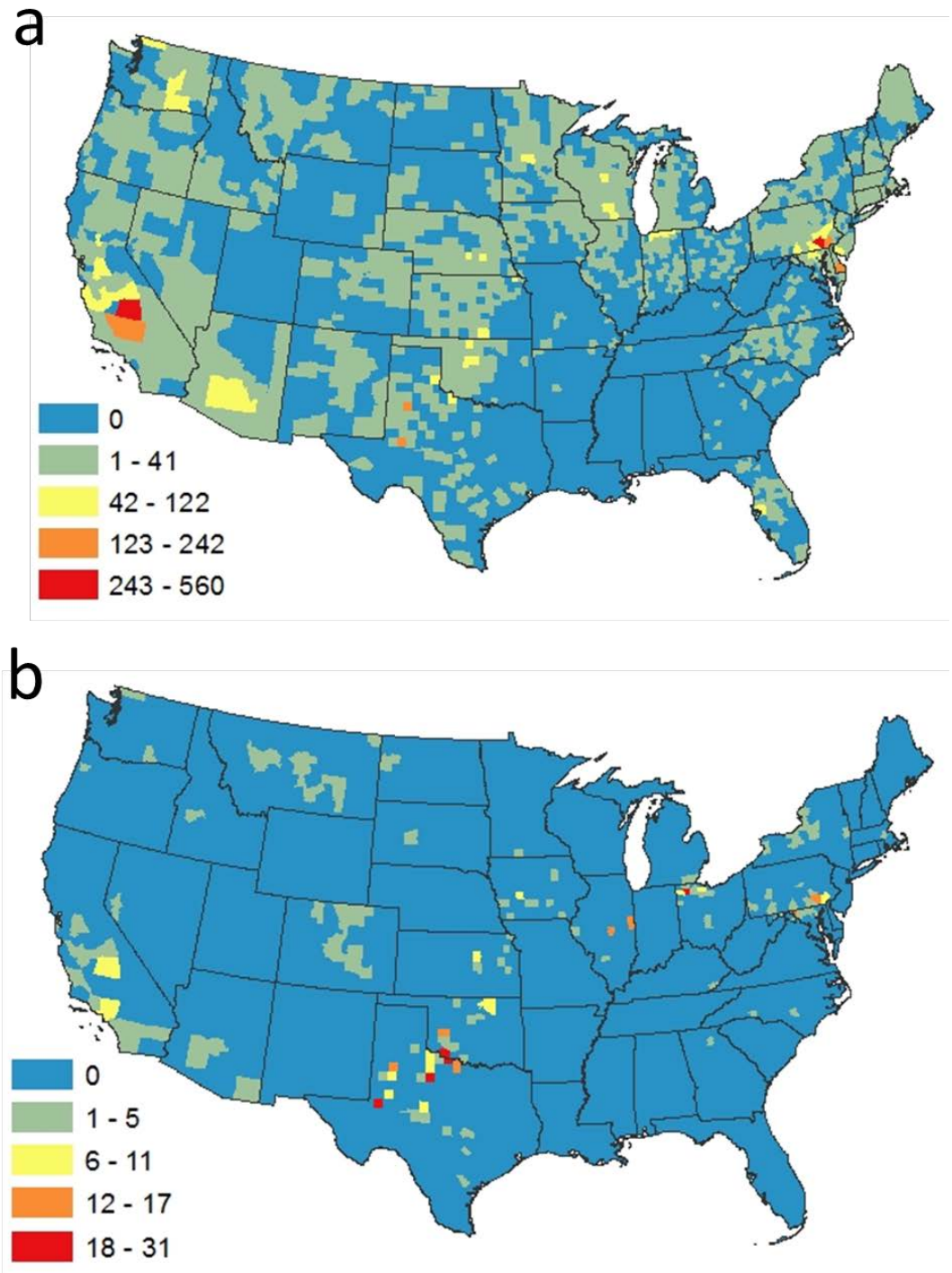


Figure S10. Total number of violations by county separated by (a) groundwater and (b) surface water. These violations are based on the number of unique PWSs in violation each year per county. Note that the scales are different for each panel.

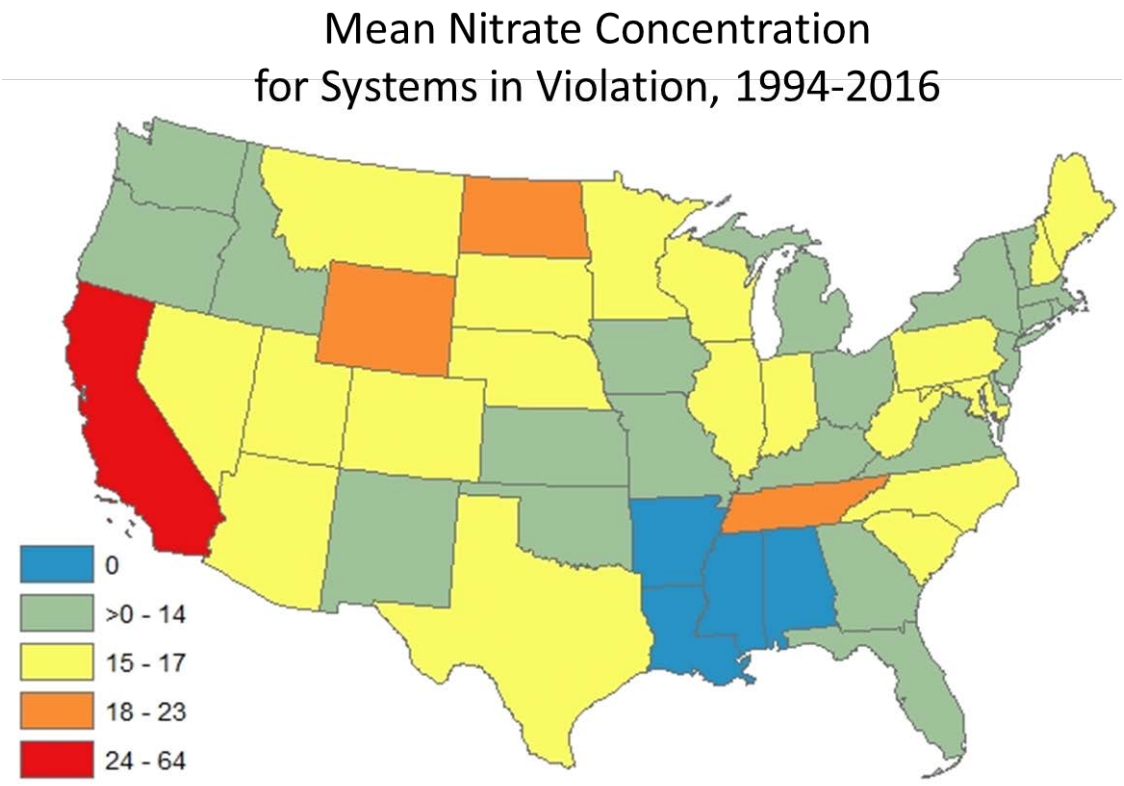
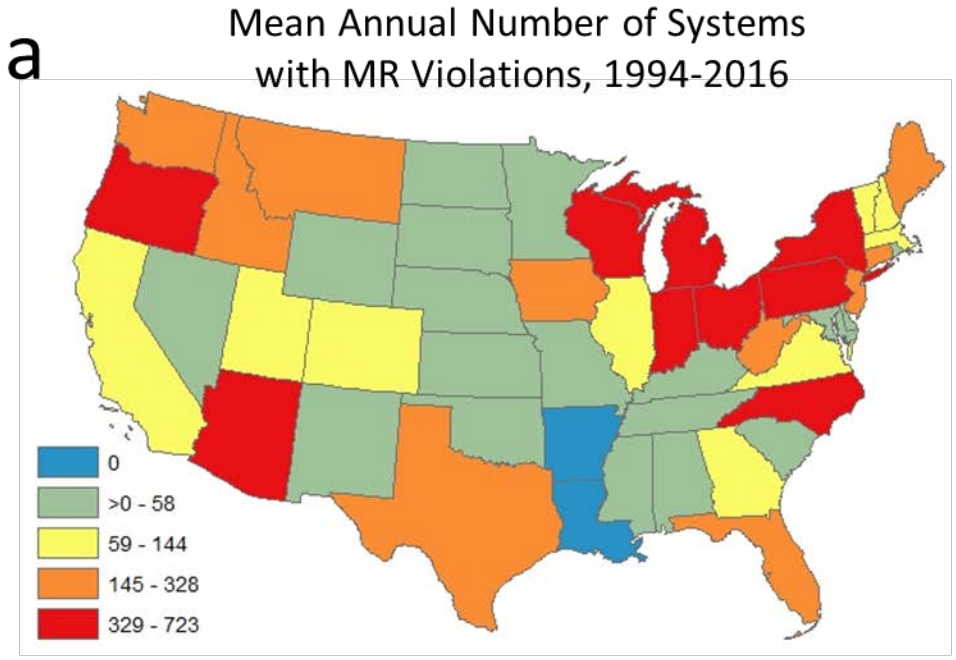


Figure S11. Mean nitrate concentrations per state based on systems violating the MCL between 1994 and 2016 (thus no concentrations below 10 mg/L are included). Note that the scales are different for each panel.



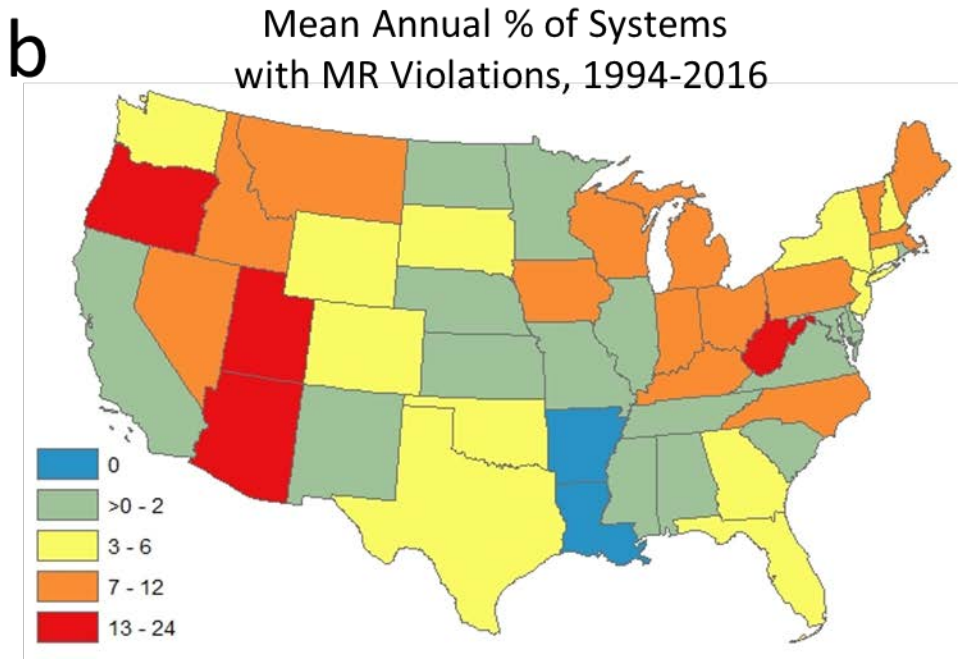


Figure S12. Monitoring and Reporting (MR) violations by state as (a) total number of violations and (b) percent of all violations from 1994-2016. Note that the scales are different for each panel.

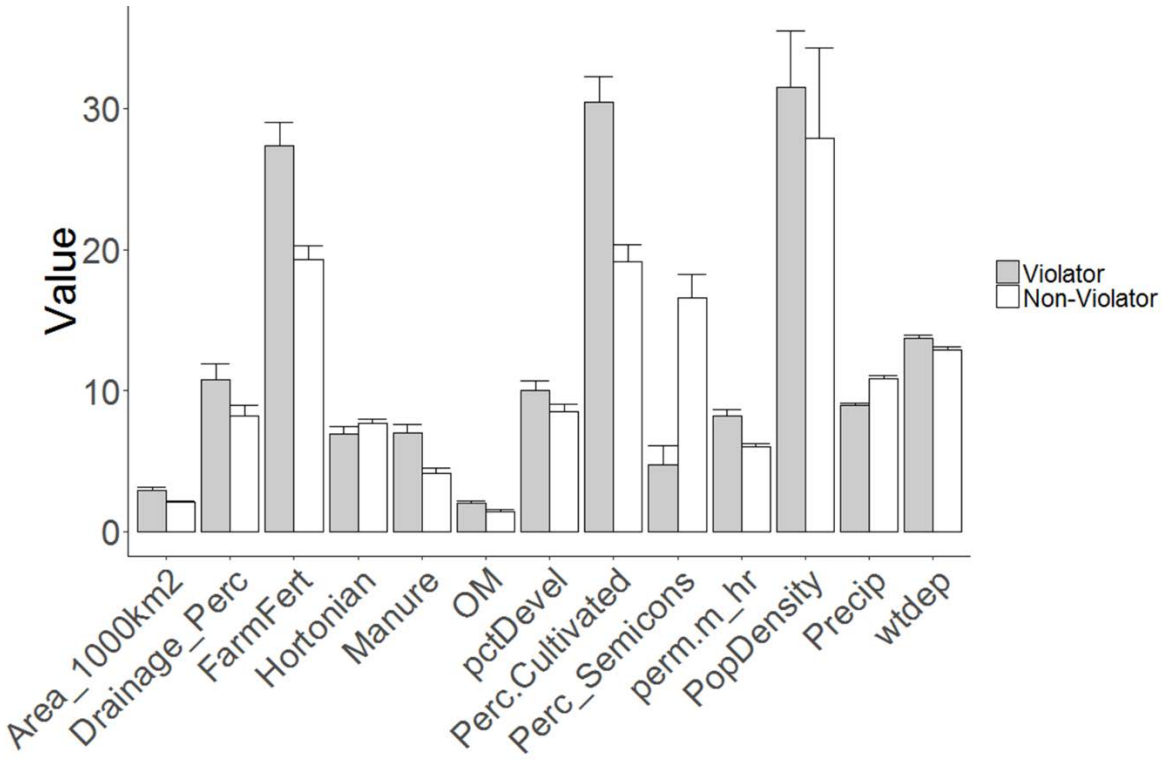


Figure S13. Comparison of U.S. counties with (Violator) and without violations (Non-Violator). Error bars are 95% confidence intervals for the mean. "Area_1000km2" is the county area in square km × 1000. "Drainage_Perc" is the percent of land within the county with man-made agricultural drainage;[2] "FarmFert" is the average annual nitrogen input from commercial fertilizer applied to agricultural lands, 1992-2001, in kilograms per hectare;[3] "Hortonian overland flow" is infiltration excess overland flow estimated by TOPMODEL, in percent of streamflow;[3] "Manure" is the average annual nitrogen input from confined animal manure, 1992 and 1997, in kilograms per hectare;[3] "OM" is soil organic matter content (% by weight);[3] "pctDevel" is the percent of county land that is classified as developed land by the National Land Cover Dataset (NLCD);[4] "Perc.Cultivated" is the percent of county land that is classified as cultivated land by the NLCD;[4] "Perc_Semicons" is the presence or absence of semiconsolidated sand aquifers;[3] "perm.m_hr" is the mean permeability (m/hour) of soils (STATSGO) within the county;[5] "PopDensity" is the population density per county × 2;[3] "Precip" is the 30-year normal mean precipitation (mm): Annual period: 1981-2010;[6] and "wtdep" is the water table depth (cm x 1000).[5]

References

1. SDWIS. U.S. Environmental Protection Agency Safe Drinking Water Information System. https://ofmpub.epa.gov/apex/sfdw/f?p=108:1:::NO:1 (accessed September 15, 2016).
2. Community for Data Integration, County-level "Land drained by tile" (and other land use practices) from the 2012 Census of Agriculture U.S. Geological Survey. https://my.usgs.gov/confluence/display/cdi/ETWG+Integration+-+Tile+drainage: 2016.
3. USGS. Spatial Data Sets Available on the WRD NSDI Node. U.S. Geological Survey. https://water.usgs.gov/lookup/getgislist (accessed October 4, 2017).
4. Homer, C.; Dewitz, J.; Yang, L.; Jin, S.; Danielson, P.; Xian, G.; Coulston, J.; Herold, N.; Wickham, J.; Megown, K., Completion of the 2011 National Land Cover Database for the conterminous United States–representing a decade of land cover change information. *Photogrammetric Engineering & Remote Sensing* **2015,** *81* (5), 345-354.
5. Soil Survey Staff. Natural Resources Conservation Service. United States Department of Agriculture. Web Soil Survey. Available online at http://websoilsurvey.nrcs.usda.gov/ (accessed October 4, 2017).
6. Daly, C.; Halbleib, M.; Smith, J. I.; Gibson, W. P.; Doggett, M. K.; Taylor, G. H.; Curtis, J.; Pasteris, P. P., Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International journal of climatology* **2008,** *28* (15), 2031-2064.