Environmental Science & Technology

Article

Cite This: Environ. Sci. Technol. 2018, 52, 7149−7159

pubs.acs.org/est

Big Groundwater Data Sets Reveal Possible Rare Contamination Amid Otherwise Improved Water Quality for Some Analytes in a Region of Marcellus Shale Development

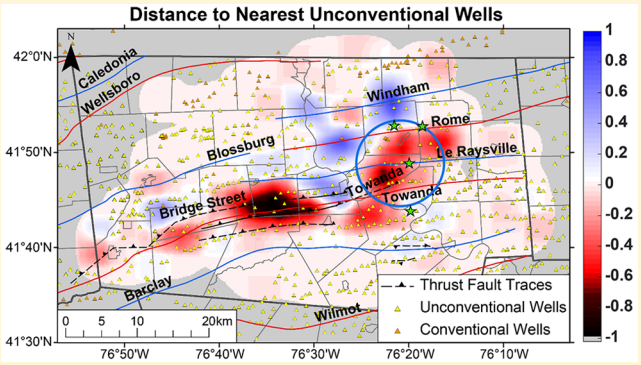
Tao Wen,*[,†] Xianzeng Niu,[†] Matthew Gonzales,[†] Guanjie Zheng,[‡] Zhenhui Li,[‡] and Susan L. Brantley[†]

[†]Earth and Environmental Systems Institute, Pennsylvania State University, University Park, Pennsylvania 16802, United States

[‡]College of Information Sciences and Technology, Pennsylvania State University, University Park, Pennsylvania 16802, United States

Supporting Information

ABSTRACT: Eleven thousand groundwater samples collected in the 2010s in an area of Marcellus shale-gas development are analyzed to assess spatial and temporal patterns of water quality. Using a new data mining technique, we confirm previous observations that methane concentrations in groundwater tend to be naturally elevated in valleys and near faults, but we also show that methane is also more concentrated near an anticline. Data mining also highlights waters with elevated methane that are not otherwise explained by geologic features. These slightly elevated concentrations occur near 7 out of the 1,385 shale-gas wells and near some conventional gas wells in the study area. For ten analytes for which uncensored data are abundant in this 3,000 km^2 rural region, concentrations are unchanged or improved as compared to samples analyzed prior to 1990. Specifically, TDS, Fe, Mn, sulfate, and pH show small but statistically significant improvement, and As, Pb, Ba, Cl, and Na show no change. Evidence from this rural area could document improved groundwater quality caused by decreased acid rain (pH, sulfate) since the imposition of the Clean Air Act or decreased steel production (Fe, Mn). Such improvements have not been reported in groundwater in more developed areas of the U.S.



INTRODUCTION

In the past decade and a half, improvements in drilling and completion of wells, including high volume hydraulic fracturing (HVHF, i.e., "fracking"), have changed the energy industry. By 2040, gas from the so-called "unconventional" formations (e.g., shale) will account for 69% of total U.S. dry natural gas production.[1] The rapid development of unconventional shale gas in the U.S. has occasionally caused environmental issues including the leakage of methane out of gas wells or spills involving other contaminants.[2−8] The debate about the causes and mechanisms of these environmental problems, including spills, casing and cementing failures, well blowouts, or out-of-zone stimulation, is still ongoing.[9−16] We address this controversy by using new techniques of data mining to look at a large number of groundwater samples in one of the PA counties with the most number of shale-gas wells (Bradford).

A key aspect of the controversy that makes it difficult to resolve in PA is that incidents appear to be rare compared to the number of shale-gas wells, and the wells are spatially distributed across thousands of hectares of heterogeneous aquifers or within complex stream networks characterized by nonuniform land use.[17] At the same time, water quality data prior to development are sparse. Furthermore, water quality data often do not follow normal univariate distributions. To be able to handle such distributions, we need new tools of statistical data analysis that also take into account spatial autocorrelation.[18]

Oil and gas companies collect groundwater samples (so-called "pre-drill" samples) in the vicinity of proposed oil/gas wells before drilling.[19] These predrill data are released to the Pennsylvania Department of Environmental Protection (PA DEP) and used to establish water quality baselines. PA DEP generally does not release all data to the public because of the cost of removing confidential information. The lack of groundwater data then hampers the ability of researchers outside of DEP to analyze the potential impacts of shale gas on water quality.[11]

Many studies of groundwater quality have nonetheless been conducted that have focused on methane,[2,8,20,21] the most common contaminant related to oil and gas activity as reported by the PA DEP.[11] However, due to the vastly different volumes of data (∼100 to ∼10,000 sites) used in these studies, researchers have reached different conclusions.[9,10] For example,

Received: February 28, 2018
Revised: May 9, 2018
Accepted: May 21, 2018
Published: May 22, 2018

DOI: 10.1021/acs.est.8b01123
Environ. Sci. Technol. 2018, 52, 7149−7159

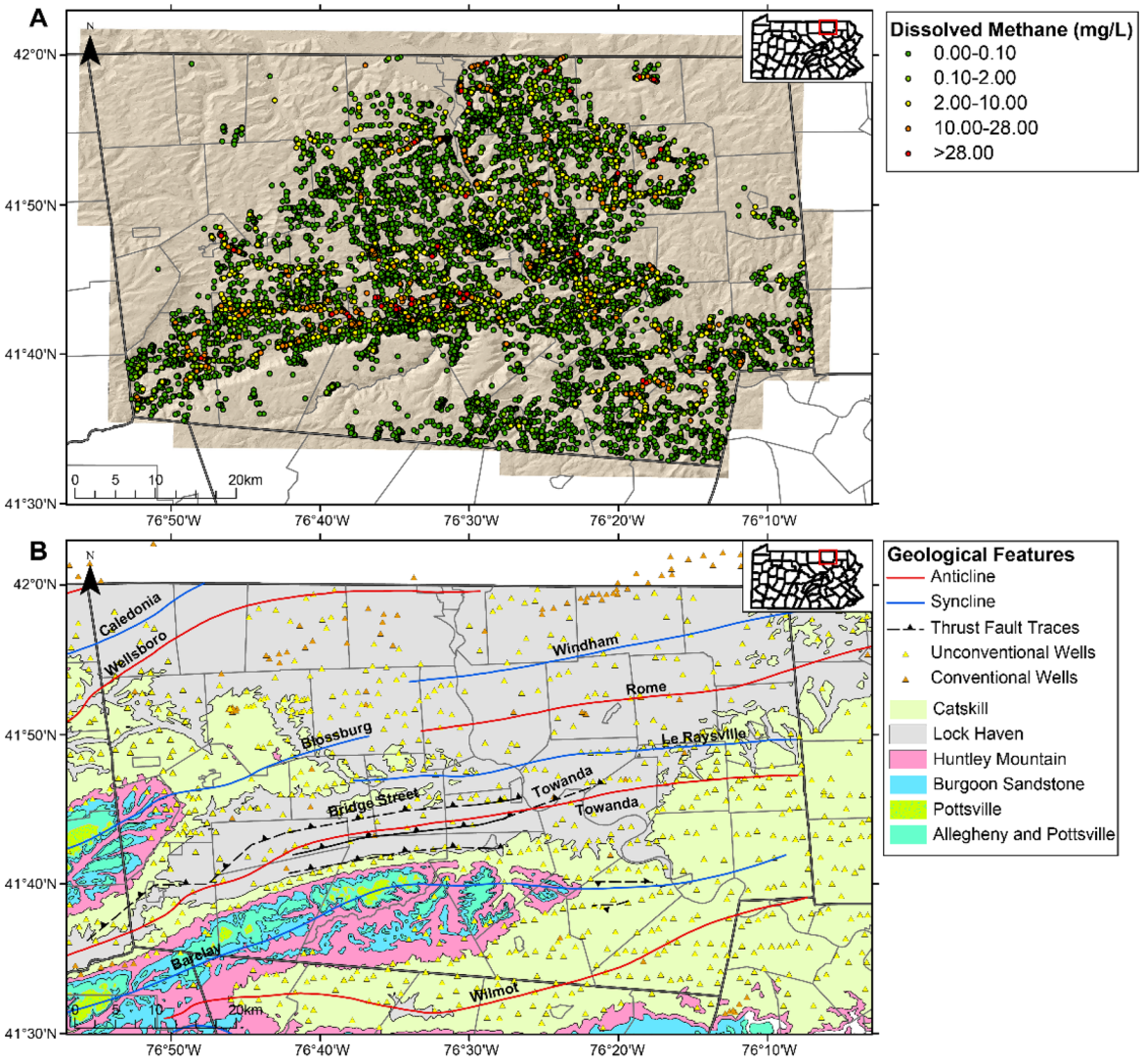


Figure 1. Location of water samples and geologic setting in Bradford County, PA. (A) Locations of 11,156 groundwater samples in BO, BN, and CHK data sets. The thick outline indicates the extent of Bradford County, PA. Symbol color represents the concentration of dissolved methane in each groundwater sample. A small subset of these (n = 1,690) was previously discussed by Li et al.[9,10] (compare to their Figure 1). Analyses are not available for some parts of the county because in those regions either no shale-gas drilling has occurred since 2004 or data that were released to the regulator have not yet been formatted and checked for inspection. (B) Locations of conventional and unconventional wells in Bradford County, PA and neighboring areas including some townships in NY. Locations of anticlines, synclines, and faults are provided based on published information.[7,32−34] Bedrock formations are also indicated in the study area based on Berg et al.[32]

methane concentrations correlated with proximity to shale-gas wells for a small data set collected by one research team[8] but did not correlate in an analysis of >11,000 water samples from an industry-controlled data set that was not released in entirety.[15] In contrast to those studies, Li et al.[9,10] published complete data sets with intermediate numbers of analyses from predrill data and used a new data mining technique to show that dissolved methane concentrations correlate with proximity to shale-gas wells in subregions of larger areas.

In addition to methane, other contaminants such as salts and heavy metals have also been identified in groundwater and occasionally attributed to gas development in shale-gas regions.[11] In PA, the DEP issues determination letters to properties that register a complaint about a water supply after an investigation into whether the oil/gas activity is responsible for the degradation. In letters where positive determinations of presumed responsibility were made from 2008 to 2012, Fe and Mn were the second most common contaminants of groundwater after contamination by natural gas.[11]

Salt species such as Cl and Ba that are enriched in the gas production waters have also been identified as transient contaminants in some groundwaters near oil and gas activities. Furthermore, the public also worries about other contaminants in groundwater near oil and gas development such as arsenic (As), an element that has well-known adverse impacts on human health.[22] Elevated levels of other heavy metals have also been found in groundwater overlying shale energy extraction.[3,4]

In this study, we showcase a new data set of predrill groundwater chemistry for 11,156 samples analyzed by commercial laboratories and released to the PA DEP. They were collected from 2010 to 2016 from Bradford County, the county that hosts the second-highest number of shale-gas wells among PA counties[23] (Figure 1a; Figure S1). Development of the Marcellus unconventional gas reservoir started in Bradford County in 2005, with a marked increase after 2008. From 2005

DOI: 10.1021/acs.est.8b01123
Environ. Sci. Technol. 2018, 52, 7149−7159

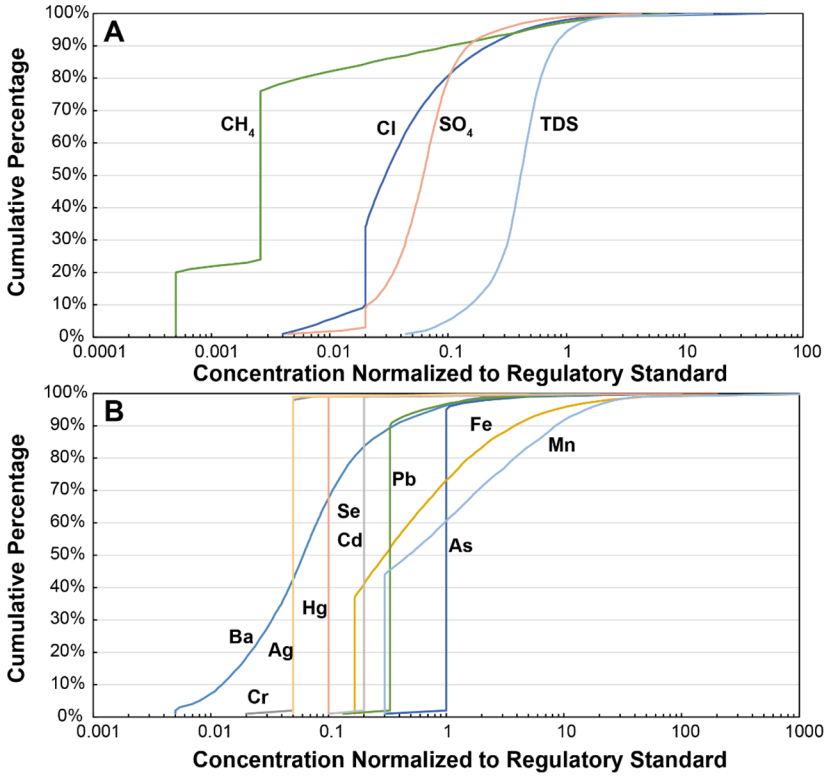


Figure 2. Cumulative probability diagram for (A) CH_4, Cl, SO_4, TDS, and (B) 10 trace elements with over 1,000 measurements in the Bradford-2010s groundwater data set and with established EPA standards. Concentrations are normalized to the standards so that at a value of unity, concentration is equal to regulatory standard. Analyses reported to be below the reporting limit (RL) are plotted at the RL (i.e., vertical sections of the curves). Different RLs are established by the various analytical laboratories, explaining data plotting below the predominant RL. Most of the reported concentrations for trace metals, except Ba, are below the RLs. For methane, 26.9% (2948/10972) measurements are above the three RLs for the relevant analytical laboratories, i.e., 0.001, 0.005, and 0.026 mg/L. Regulatory thresholds (* denotes SMCL) for the analytes (in mg/L) are CH_4 (10), Cl (250*), SO_4 (250*), TDS (500*), Ba (2), Ag (0.1*), Hg (0.002), Cr (0.1), Cd (0.005), Se (0.05), Pb (0.015), As (0.01), Fe (0.3*), and Mn (0.05*). These standards are derived from the US EPA except for CH_4 which is from US Department of Interior. TDS is rarely below the RL. RLs for Cl, SO_4, and TDS are well below the EPA standard. RL for As is equal to the EPA MCL of 0.01 mg/L. Therefore, when counting the number of As values failing EPA MCL, we excluded those censored data and only counted values above the MCL.

to 2016 in the county, 40 out of the 53 positive-determination letters issued by the PA DEP (where oil and gas development activity was presumed responsible) reported elevated methane concentrations, and 21 reported Fe and Mn contamination.

Only by using such large data sets and novel data mining techniques can we simultaneously explore large areas while looking for rare incidents. Such techniques thus hold the promise of helping stakeholders understand the frequency of problems as well as which areas are more prone to problems and why, but the real test demanded by the public is to determine if water quality is deteriorating with time. We therefore also compare data sets to understand water quality before and after onset of shale-gas drilling in Bradford.

MATERIALS AND METHODS

PA DEP shared three batches of "pre-drill" water chemistry data with us, referred to here as BO, BN, and CHK. The original data set (BO, 1690 data values) was analyzed by Li et al.[9,10] These data sets were digitized, cleaned, and merged for quality assurance (QA) (SI). The data set that merged all 3 subsets (hereafter noted as "Bradford-2010s study") contains 11,156 groundwater samples − 10,213 water samples from water wells and 943 from springs (Figure 1a). For water wells, 6,123 well depths are reported that range from 1 to 488 m with an average depth of 54 m, while the elevations of all sampling sites vary from 198 to 721 masl (meters above sea level). Since most groundwater samples are collected from shallow water wells, the observed spatial and temporal trends we discuss for groundwater quality in this paper are mostly indicative of shallow groundwaters that migrate to domestic water wells (see Figure 1b and the SI for geological and hydrogeological settings in the region). To analyze the data, distances were calculated between sampling sites and features such as oil and gas wells, geologic faults, the cores of anticlines, and streams (see the SI).

We use a "sliding window" data mining technique (source code released to GitHub (https://git.io/vNdsd)). The sliding window technique was used to explore the correlation between methane concentrations in water and the distance to factors of interest (e.g., distance between water well and nearest shale-gas well) across a large region. More details on the sliding window technique are provided in the Discussion section.

RESULTS

Forty-five of the analytes with at least 100 measurements were grouped into four categories (Table S1): (1) major inorganic (median ≥1 mg/L); (2) trace inorganic (median <1 mg/L); (3) organic; and (4) other. Selected analytes of each category are discussed below. Of these, 11 analytes (Cl, Na, SO_4, TDS, Mn, Fe, As, Ba, Pb, CH_4, and pH) were further interpreted because of their higher prevalence of reporting and rate of

DOI: 10.1021/acs.est.8b01123
Environ. Sci. Technol. 2018, 52, 7149−7159

failure of regulatory thresholds. To ascertain whether the onset of shale-gas development impacted groundwater chemistry, we also evaluated temporal trends by comparison with two previous studies. The first study[24] (hereafter noted as "Bradford-1980s study") included complete chemistry reported for samples from 108 wells covering all Bradford County. The analyses were measured on waters collected from 1935 to 1986 with the vast majority (103/108) in the 1980s (Figure S1). In addition, Boyer et al.[25] (hereafter noted as "Rural PA-2010s study") collected predrill water from 233 private water wells in 2010 and 2011 and reported the data as a statistical summary for all of rural Pennsylvania. Most of these data derived from the two hotspots of shale-gas drilling, i.e., southwestern or northeastern PA.

By comparing the Bradford-1980s study to that of Bradford-2010s and to that of Rural PA-2010s, we made two temporal comparisons for selected analytes. The first comparison focuses on mostly the same region in Bradford but with vastly different volumes of data, while the second comparison includes roughly the same data volume for each time period but treats different regions (i.e., Bradford-1980s treats Bradford (i.e., northeastern) whereas Rural PA-2010s treats rural areas of PA, with an emphasis on southwestern and northeastern near Bradford). Inclusion of southwestern water in this data set is expected to bring in water chemistry that is generally worse than that of the northeastern PA because southwestern PA is more urbanized (i.e., Pittsburgh), has more roads, and is the center of much of the PA coal mining industry.[26]

For these temporal comparisons, the means for metals, arsenic, sulfate, or chloride are not compared because their distributions are skewed: skewness for the distributions varies from 20.1 to 49.9 in Bradford-2010s and is thus much higher than the threshold of 2 for normal univariate distributions.[27] Instead, we use Kruskal−Wallis rank sum tests (K−W tests) to compare medians for Bradford-1980s and 2010s. A *t*-test is used for comparisons of pH mean values.

Major Inorganic Analytes. The total dissolved solids (TDS) in the shallow groundwaters in the Bradford-2010s data range from <500 mg/L to 8560 mg/L, with a median value of 205 mg/L (Table S1). Two major types of waters are observed, characterized as either Na-Ca-Cl or Ca-Na-HCO_3 waters. These were previously identified in PA as deep saline water and shallow modern waters, respectively.[20,24]

Using ternary diagrams (Figure S2) and Br/Cl ratios,[20] water types can be further classified. (All concentrations and ratios are on a mg/L basis.) Waters in so-called types A and B (n = 8148) correspond to low salinity water (Cl < 20 mg/L) and are dominated by Ca-HCO_3 or Na-HCO_3, respectively. For these waters, the major ions generally come from rock, soil, or precipitation. In contrast, types C (n = 10) and D (n = 513) waters are characterized by higher salinity (Cl > 20 mg/L) and varying Br/Cl ratios (type C: Br/Cl < 0.001; type D: Br/Cl > 0.001). Type C waters have received inputs from wastewater or road salt, while type D are impacted by Appalachian Basin brine. These brines are found throughout the Marcellus region at depth and are a likely source for the briny flowback and production waters that accompany gas production in the Appalachian basin.[28]

The U.S. Environmental Protection Agency (US EPA) does not establish a Maximum Contaminant Level (MCL) in national drinking water standards for any of these major inorganic analytes. However, Secondary MCLs (SMCLs) are listed for chloride, sulfate, and TDS in drinking water (Table S1). In the Bradford-2010s data, 2% (214/10931), 1% (109/10933), and 5.8% (635/10936) of chloride, sulfate, and TDS measurements, respectively, fail these corresponding SMCLs (Table S1) as shown in the cumulative probability diagrams normalized to the corresponding SMCLs (Figure 2).

In a comparison between the Bradford-2010s data and the earlier data sets, statistical evidence suggests that Cl concentrations are the same or improving. For example, the medians of the Bradford-1980s and -2010s studies are not statistically distinguishable (12 vs 7.38 mg/L; p = 0.1370), while the median of the Rural PA-2010s data is lower, 5.9 mg/L (Table S1 and S2). Furthermore, the rate of failure of the EPA standard declined significantly from 8.0% (8/100 in the Bradford-1980s) to 2.0% (214/10931 in Bradford-2010s) or <1% (<3/226 in the Rural PA-2010s). Similar to Cl, the medians of Na concentration in Bradford-1980s (17.5 mg/L) and -2010s (16.3 mg/L) data sets are not statistically distinct (p = 0.3138), but the Rural PA-2010s median is lower, 10.7 mg/L. USEPA[29] recommended sodium concentration in water not exceed 20 mg/L for health benefits. 46.1% and 43.2% of samples are above this EPA guidance level of sodium in Bradford-1980s and -2010s data sets, respectively. In addition, 65.4% and 66.1% of Bradford-1980s and -2010s data sets are above 10 mg/L. Such high levels of sodium being consistently present in some water from 1980s through 2010s might suggest the natural migration of deep brine to shallow formations in parts of Bradford. Over the same time period of 1980s to 2010s, the median of TDS and the rate of failing EPA standard declined from 248 to 205 mg/L (p < 0.01) and from 13.6% to 5.8%, respectively. In comparison, TDS median is 190 mg/L in the Rural PA-2010s and its rate of failure of the EPA standard is 3%.

The comparison of sulfate in the Bradford-1980s and Bradford-2010s data sets also indicates improvement: a statistically significant decline in median from 20 to 15.5 mg/L (p < 0.01). Likewise, the median of the Rural PA-2010s data (14.0 mg/L) is also lower than that of Bradford-1980s, even though this rural PA data set includes waters from the coal-mining area where acid mine drainage is a significant problem.[26]

Trace Inorganic Analytes. Over 10,000 concentrations are reported in the Bradford-2010s data for each of the trace metals except lithium (Li) and strontium (Sr). Over 90% of these are below the reporting limit (RL). In contrast, over half of the measurements of iron (Fe), barium (Ba), and manganese (Mn) are above RLs. All trace elements (e.g., As, Ba, Pb) have a < 5% incidence rate of failing EPA standards except Fe (26.9%) and Mn (39.3%) (Figure 2, Table S1).

The median Mn concentration in Bradford decreases from 0.08 in Bradford-1980s to 0.02 mg/L in Bradford-2010s (p < 0.01). In comparison, the median is 0.01 mg/L in the Rural PA-2010s data. The rates that samples failed the EPA SMCL for Mn also declined from 60.6% in Bradford-1980s (60/99) to 39.3% in Bradford-2010s (4300/10934) (Figure S3) and 27% in Rural PA-2010s (55/203). Based on these data, then, Mn concentrations in Bradford county groundwater have thus been decreasing from the 1980s to the 2010s.

Like Mn, the median value for Fe decreases from 0.27 to 0.09 mg/L from the Bradford-1980s to the Bradford-2010s study (p < 0.01). Furthermore, the rates of samples failing the EPA SMCL declined from 49.5% (52/105) to 26.9% (2936/10934), respectively (Figure S3). In comparison, for the Rural PA-2010s data, the median is 0.05 mg/L and a 20% failing rate.

DOI: 10.1021/acs.est.8b01123
Environ. Sci. Technol. 2018, 52, 7149−7159

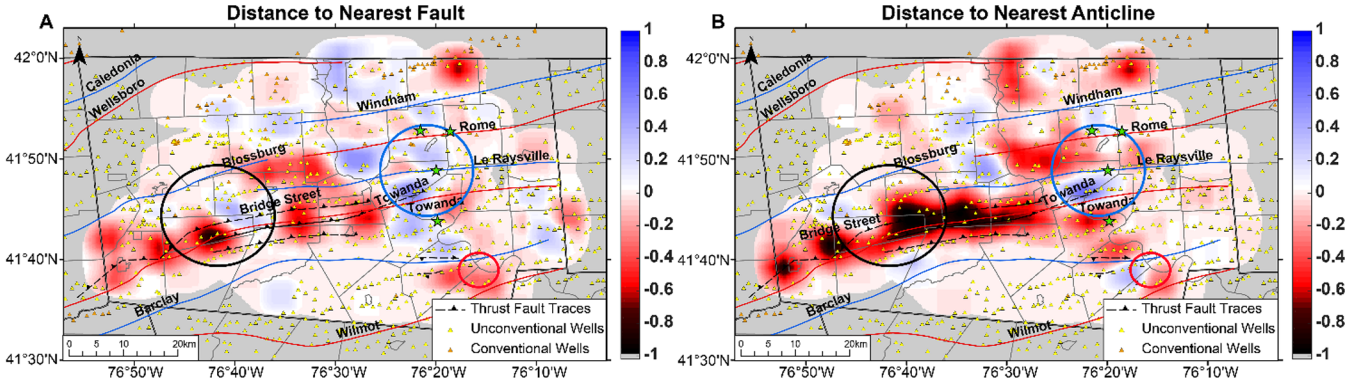


Figure 3. Maps calculated using sliding window technique (see text) showing where dissolved methane concentrations in water samples correlate with distance to (A) faults and (B) anticlines. Colors, indicating where methane increases (red) or decreases (blue) closer to gas wells, are only plotted when correlations are statistically significant. A blue circle denotes the hotspot area in Figure S5a. A black circle represents a sampling area previously identified as being the location of elevated methane correlated with valley bottoms (see the SI).[2,21] A red circle represents the sampling area where a known gas migration incident occurred.[7] From north to south, the faults are the Bridge Street, Towanda, and three unnamed faults.[7,32−34]

The median values of As for Bradford-1980s and Bradford-2010s are 0.01 and 0.005 mg/L, respectively (Tables S1 and S2). However, both of these values are equal to the detection limits for As values in each data set; therefore, we do not compare them. The median for the Rural PA-2010s data is 0.0025 mg/L. For the early time period, the rate of violation of the arsenic MCL (0.01 mg/L) was 9.5%. That value is higher than the failure rate (4.4%) in Bradford-2010s or the value (4%) for 2010/2011 reported in the rural PA data set by Boyer et al.[25] These data suggest that As contaminations might be improving in the Bradford groundwaters.

Unlike Mn and Fe, the median of the Pb concentrations is not statistically different between Bradford-1980s and Bradford-2010s (p = 0.4071). Like Mn, Fe, and As, however, the rate of failure of Bradford groundwater with respect to the EPA Pb action level (0.015 mg/L) has decreased from 7.1% (6/84) in Bradford-1980s to 3.1% (344/10934) in Bradford-2010s. Over the entire state as reported in the rural PA study,[25] this value is 7%.

Ba concentrations in groundwater in Bradford are not statistically different between Bradford-1980s and Bradford-2010s with respect to the median (p = 0.6896) (Tables S1 and S2). Likewise, the very low rate of failure of the EPA standard is almost unchanged: 2.9% (2/69) versus 3.8% (412/10934). In contrast, the Rural PA-2010s study reported an incidence of failing of 1% (2/218) and a median value of 0.13 mg/L.

Organic Analytes. Five volatile organic compounds (VOCs: 1,2-DCE, benzene, ethylbenzene, toluene, and xylene), one glycol (ethylene glycol), three alkanes (methane, ethane, and propane), and oil and grease are all reported in the Bradford-2010s dataset (Table S1), and most are below the RL. EPA has established MCLs for all of these listed organic analytes, except for the three alkanes. Only benzene and methane were ever found to exceed the regulatory standards (see the SI). In these pre-drill data, stable carbon isotopic data for methane, often helpful in determining whether methane is biogenic or thermogenic,are not reported.

Other Analytes. Median values for the other analytes (*E. coli*, Methylene Blue Active Substances, pH, Specific Conductance, Temperature, Total Hardness, Total Coliform, Total Suspended Solids, Turbidity) are listed for Bradford-2010s, and the mean pH is also reported here (Table S1). Except for pH, other analytes are not included in EPA drinking water standards or are not reported in the Bradford-1980s study. Median and mean values of pH for Bradford-1980s and Bradford-2010s are 7.3 versus 7.7 and 7.36 versus 7.64, respectively. Both median and mean pH increased from the 1980s to 2010s (p < 0.01).

DISCUSSION

Natural and Anthropogenic Factors Affect Methane Concentration. Li et al.[9,10] developed a new data mining tool, the sliding window technique, to avoid biases resulting from the size of the data set and/or from some specific local areas. The larger data set that is published with this paper for Bradford County allows us to test the sliding window technique[9,10] for artifacts related to data volume as well as to learn more about groundwater over space and time.

This technique computes the average spatially normalized correlation between variables for a 5 km × 5 km window that is moved across the region of interest. For all water analyses in the window, Kendall rank[30] is calculated for the correlation between methane concentration and proximity to a feature of interest (e.g., previously drilled gas well). If the correlation is statistically significant, +1 (positive correlation) or −1 (negative correlation) is assigned to the window. Spatially normalized significance values, defined by dividing the summation of these values of +1 and −1 by the total number of windows covering each location, are then plotted as a "heat map" (Figure 3). The intensities of red and blue indicate the extent of significant negative or positive correlation, respectively. Correlations were studied separately in this work with respect to faults, anticlines, streams, unconventional gas wells, conventional gas wells, and gas wells (unconventional + conventional). We implicitly assume that the sliding window technique cannot elucidate geographic areas where shale-gas production activities might have exacerbated pre-existing naturally occurring methane. In other words, if natural features appear to explain high methane concentrations, we assume the natural explanation (faults, anticlines, valley/stream) are the dominant control.

For such a large water chemistry data set that is not described as a Gaussian distribution, other geospatial analysis tools such as Geographically Weighted Regression (GWR) or Hot Spot Analysis might also be used to consider spatial autocorrelation. Our sliding window technique has similar advantages and can also handle the situation where water chemistry data are below the detection limits (i.e., censored data) and there are multiple

DOI: 10.1021/acs.est.8b01123
Environ. Sci. Technol. 2018, 52, 7149−7159

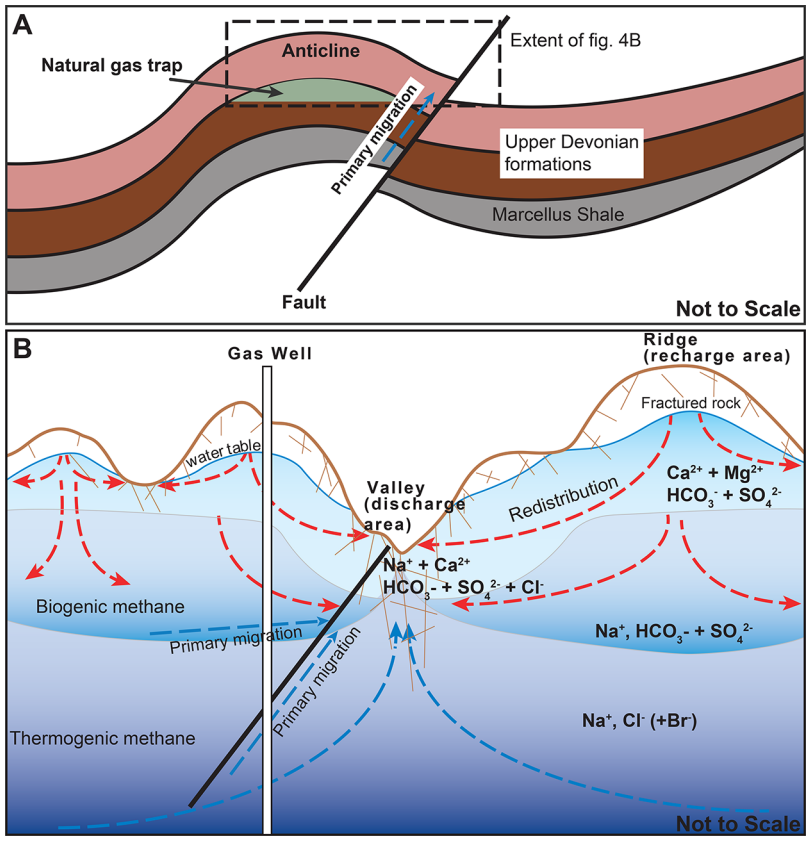


Figure 4. Conceptual diagram of subsurface migration of methane and groundwater along anticlines and faults (modified after Darrah et al.[21] and Gross and Cravotta[38]). (A) Generalized cross-section of anticline and fault in Bradford County and (B) detailed schematic associated with a valley located along a fault (solid line). Methane migrates from deep (thermogenic methane) or shallow (biogenic methane) formations, sometimes with deep brine, toward shallow aquifers as indicated with dashed blue lines. Redistribution of methane and groundwater flow from ridge to valley is shown with dashed red lines. This diagram is not to scale. Type A and B waters might be associated with waters from near the ridge top (Ca^{2+}, Mg^{2+}, HCO_3^-, SO_4^{2-}) or intermediate level (Na^+, HCO_3^-, SO_4^{2-}). Type C water is attributed to wastewater and/or road salt that are not shown in this diagram. Type D water is associated with deep brine and is characterized by high concentrations of Na^+, Cl^- ($+Br^-$).

detection limits. In addition, the output heat map for the sliding window technique is more visually intuitive.

Faults and Anticlines. Li et al.[9,10] observed that methane concentrations in 1,690 water samples have a small tendency to increase in groundwater samples near the Towanda and Bridge Street faults. Figure 3a shows that these earlier observations are maintained and amplified when analyzed for the larger data set of 11,156 analyses. Even with the larger data set, however, no single fault is delineated entirely by red. In other words, methane never increases consistently near any fault across the entire study region. Instead, "hotspots" of red (verging to black) indicate areas where methane concentrations increase closer to the fault. The Bridge Street and Towanda faults host the most hotspots across their extent.

Hotspots could indicate local regions of stronger methane flux or simply be related to sparseness of data. However, methane is well-known to migrate along fractures and faults for kilometers in northern Pennsylvania as illustrated in Llewellyn et al.[7] (red circle in Figure 3 highlights the study area in Llewellyn et al.). Several authors have pointed out that in the Appalachian plateau where strata are roughly parallel to the land surface, fractures can open parallel to strata (bedding plane openings) or across strata (joints), and both can allow methane migration, sometimes in staircase-like upward passageways.[7,31] Both biogenic gas from shallow depths can migrate as well as thermogenic methane from greater depths such as the Marcellus Formation or from other gas-rich and shallower Upper Devonian formations. These geologic observations are likely to explain why red coloration in Figure 3a is often aligned with faults.

For the first time, however, the larger data set documents that methane also has a small tendency to increase near some anticlines – especially the Towanda anticline as shown in Figure 3b. Anticlines are large regional convex-up folds of geologic strata. Strikingly, the correlation between elevated methane and anticlines is more significant than for faults. Like the faults in Figure 3a, hotspots of methane concentration plot along about 60% the Towanda anticline. No such hotspots are observed for the Rome and Wilmot anticlines.

Almost every hotspot on the fault map also appears on the anticline map (Figure 3a and 3b), whereas some hotspots on the anticline map do not display on the fault map. Apparently, either the Towanda anticline is controlling methane emissions or the combination of Bridge Street + Towanda fault + Towanda anticline together control emissions. The Bridge Street and Towanda faults are thrust faults where formations from the northwest were thrust over formations in the southeast during the Alleghanian orogeny.[32] The folding and parallel shortening during the onset of Alleghanian orogeny explains the alignment because it leads to the anticline/syncline sequences along with their associated low angle thrust faults. Synclines are large concave-up folds.

DOI: 10.1021/acs.est.8b01123
Environ. Sci. Technol. 2018, 52, 7149–7159

The stronger correlation of methane with anticlines rather than faults might be because locations of anticlines in Bradford County are better documented than faults in the study area. Faults are noted by small-scale surface observations,[7,32−34] whereas anticlines are large ridgelines that are easily visible and have been imaged in seismic studies. On the other hand, it is well-known that methane moves updip and can be trapped in structures such as anticlinal folds, and the intersection of faults and anticlines, e.g., Bridge Street and Towanda faults vs Towanda anticline, would likely allow methane emission into groundwater. In fact, deeply buried anticlinal folds in oil and gas regions are often targeted by drillers.

Patterns of Natural Methane Migration. Dissolved methane from natural sources in Bradford likely results from both deep and shallow sources.[20,21] Such methane migrates into shallow aquifers in two roughly sequential steps (Figure 4): 1) primary migration where biogenic or thermogenic methane moves along anticlines and faults or other small fractures into shallow aquifers and 2) methane redistribution where topographically driven groundwater flow diverts dissolved methane from ridge to valley. This may explain why we see a strong correlation between higher methane concentrations and the valley bottoms across Bradford County (see the SI). Such patterns have been similarly documented in PA by previous workers.[20,21,35] Specifically, dissolved methane in upflowing waters may be diverted by downward flowing water from the hill tops toward the valleys, redistributing methane in the shallow subsurface. In addition, methane-containing waters may upflow along faults that also outcrop in valley bottoms.

Unconventional and Conventional Gas Wells. Some hotspots on the heat map for unconventional wells do not overlap with hotspots on the fault or anticline maps. One hotspot on Figure S5a was previously identified by Li et al.[9,10] They associated the hotspot with three shale-gas wells that could potentially be leaking (015-20116, 015-21353, 015-20612, see Table S3) because they do not have surface or intermediate casings at the inferred intersection with a nearby fault (see the SI). Leakage of methane has been associated with shale-gas wells that lack such cement or casing in one part of Bradford.[7]

A new finding from the larger data set is another red hotspot (blue circle in Figure S5a; see also Figure 3) in eastern Bradford. Since this hotspot is not observed in Figure 3a and Figure 3b, the migration of methane along faults within an anticline cannot easily explain it. Four shale-gas wells in Bradford (API numbers: 015-21181, 015-20960, 015-20871, and 015-21352) are associated with this hotspot. Two of these (015-21181 and 015-21352) do not have surface or intermediate casings at the inferred intersection with a nearby fault (Table S3). Thus, a conservative estimate leads to the identification of 5 of the 1,385 shale-gas wells in Bradford as potentially leaking methane into shallow aquifers.

Li et al. also identified hotspots on a heat map for conventional wells that did not overlap with hotspots on the fault map. This led to identification of four abandoned or orphan wells that do not report casing or cement at depths where they intersect the nearby fault.[9,10] The heat map calculated here (Figure S5b) reveals that previously identified hotspot as well as new hotspots in northern Bradford. In those hotspots, three conventional gas wells (API number: 015-20017, 015-20023, 015-20029) received notices of violation from PA DEP in 2008 and 2009 (before the date of water sampling for methane) for abandonment of an unplugged well.[36] Methane may emit from old wells that were not completed to modern standards.[37]

Other Organic Analytes. The only organic analyte that was observed above an EPA standard, benzene, was only reported above that standard once. That sample could be associated with oil and gas activity; however, we could not definitively associate the incident with gas development (see the SI).

With respect to other organic analytes, 116 samples were reported to have VOCs such as toluene and glycol concentrations over the RLs. Locations of these samples (Figure S4) occur throughout Bradford County and toluene is the most widespread (n = 101). Concentrations of these organics are all well below EPA limits. They might derive from anthropogenic sources in some cases,[7] but given the limited number of measurements above RLs, we do not discuss these further here (see the SI).

Inorganic Analytes. *Bedrock Geology.* The sliding window technique clearly shows the importance of geology (faults, anticlines) in understanding the distribution of methane in groundwater in Bradford. Here, we use the K-W test to determine if another geological attribute, bedrock composition, is important in determining inorganic chemistry of the groundwaters. We find that Cl, Na, SO_4, TDS, Mn, Fe, As, Ba, CH_4, and pH display different median values in different formations (99% confidence level). This is consistent with other studies where formations are generally the most important natural controlling factors of groundwater quality[38,39] (see the SI). On the other hand, medians of Pb are not statistically different between formations (see the SI and section on *Pb Concentrations*).

Temporal Trends in Inorganic Analytes. As discussed above, data reported from the 1980s to 2010s are consistent with the conclusion that unconventional gas production has not worsened groundwater quality for Mn, Fe, As, Pb, Ba, Na, Cl, TDS, SO_4, and pH. In fact, for those analytes, groundwater concentrations in Bradford County are constant or have improved (i.e., in the case of pH, increased). Here we discuss possible reasons for this observation.

First, we consider possible artifacts that might make it look like groundwater chemistry is improving when it is not. For example, if the two comparison sets of data sampled water in different seasons, seasonable variability of analyte concentrations (e.g., Fe, Mn, As) could be a hidden variable. For example, As shows natural seasonal variations in groundwater.[40,41] In particular, Ayotte et al.[40] observed arsenic concentrations in groundwater in New England in the first half of the year are significantly lower than in the last half. For the waters investigated here, samples that were collected in the first half of the year account for 14.8% and 48.3% of Bradford-1980s and -2010s data sets, respectively. In theory, then, seasonal effects therefore might contribute partially to the observed difference in analyte concentrations between these two data sets. However, within the Bradford-2010s data set, the median values of samples collected in the first and second half-years are the same for As: 0.010 mg/L. The example of As data in Bradford-2010s data set suggests that the impact of seasonal variability on groundwater quality parameters might not be the primary factor. For Fe and Mn in the Bradford-2010s data set, we indeed observe slightly different median values between seasons. To remove the impact of seasonal variability on temporal trend, we compare medians of As, Fe, and Mn of samples collected in the second half year only (i.e., when the

DOI: 10.1021/acs.est.8b01123
Environ. Sci. Technol. 2018, 52, 7149−7159

Bradford-1980s data set has the most samples). We observe these medians are either not changing or declining ($p < 0.01$) from the Bradford-1980s to -2010s data set. This again confirms the observed temporal trend in As, Fe, and Mn concentrations regardless of the seasonal variability.

It is also possible that the analytical or sampling techniques might have changed with time to make concentrations appear lower today than in the past. In particular, most of the cation data in Bradford-2010s were measured with inductively coupled plasma (ICP) atomic-emission spectrometry, while the method used in the Bradford-1980s was atomic absorption (AA) spectrometry.[24] However, where the Bradford-1980s data set reported analytes as "less than" specific threshold reporting values (As, Fe, Mn, Pb, and Ba), those values are comparable to reporting limits in Bradford-2010s. Such consistency would imply that the sensitivity of the methods is similar. If anything, changes in technique over time would be likely to have caused today's concentrations to be higher than those in the past because the earlier samples from Bradford-1980s were filtered (0.45 μm) while the more recent samples were not. Since total concentrations can include particulates, this difference would cause the earlier data to be lower in concentration than the later data.

Alternately, given that groundwater chemistry varies spatially, the earlier data set might simply have sampled wells with higher-concentration groundwaters. If there was such a bias in the Bradford-1980s versus our data set, we might expect statistically different median values in all concentrations. In this regard, the similarity of values of Pb, Ba, Cl, and Na between Bradford-1980s and Bradford-2010s therefore provides some confidence that the 1980s data set was not spatially biased compared to our data set.

Although none of these possible artifacts can be eliminated as possible explanations for the apparent improvement in water quality in Bradford, they also do not yield a convincing explanation for all the improving analytes. In contrast, several explanations for why groundwater chemistry data might be actually improving are consistent with broader trends reported in the literature outside of the shale-gas region, as discussed below.

pH and Sulfate. Increasing pH and decreasing sulfate in groundwater could both be consistent with trends of less acid rain (related to the imposition and subsequent amendments of the Clean Air Act since the 1970s) and decreased acid mine drainage in PA (related to decreasing rates of coal mining).[26,42−45] In addition, the conversion of power plants from burning coal to cleaner natural gas may contribute to the decline of contaminant emissions into the atmosphere. For example, Paulot et al.[44] pointed out that SO_2 emission has declined by over 60% in the eastern U.S. since the late 1990s. Consistent with this, Niu et al.[26] and Raymond and Oh[46] documented that sulfate concentrations in streams in PA steadily decreased from 1965 to 2015. Those researchers attributed the decrease largely to the decline of coal production in PA. In addition, Kaushal et al.[45] observed increasing values of pH in streams and rivers across North America from the 1950s to 2010s. Likewise, Strock et al.[43] concluded that sulfate concentrations in lakes in the northeastern U.S. significantly decreased during the 2000s.

Of course, even if rain became less acidic with time, this change has to propagate into groundwater if it can explain our observations. In fact, time scales of recharge to groundwater are consistent with the temporal changes observed. For example, Yager[47] reported that the vertical hydraulic conductivities of the upper aquifers within the upper Susquehanna River Basin vary between 0.6 to 12 m/day. This is consistent with pore velocities of 0.2 to 4 m/day for a typical porosity for Devonian formations in PA (15%)[48] and a typical hydraulic gradient in the ridge and valley province of 0.05. With these values, recharge waters can reach 72 m (a depth greater than or equal to 75% of the sampled water wells, Figure S1c) within 18−360 days. Although waters will be chemically transformed along such flowpaths, small temporal changes in groundwater may nonetheless be recorded.

Fe, Mn, and As Concentrations. Higher pH could also be the reason why dissolved metal concentrations have decreased in groundwater: metal solubility decreases with increasing pH. In addition, other effects may explain lower metal concentrations. For example, published evidence suggests that Mn was deposited from the atmosphere to soils throughout much of Pennsylvania from approximately 1900 to 2000.[49] This Mn has been flushing out of soils and into rivers especially since the 1970s.[49] Some of this soil Mn is likely to have entered groundwater since Mn is relatively mobile,[49−51] perhaps explaining the decrease in both median Mn concentration and incidences of failing the EPA standard in Bradford from the 1980s to 2010s.

Similarly, atmospheric deposition of Fe likely occurred along with the Mn because most of the industrial activities that have decreased over the U.S. over time that release Mn to the air also release Fe.[52] Once again this could contribute to improvement of groundwater quality with respect to Fe concentrations, especially if some of the deposited Fe is as mobile small particles. Decreasing atmospheric input[49] of Mn and Fe with the rates of recharge discussed above could explain decreasing Mn and Fe in groundwater.

Along with Fe and Mn, arsenic has also been released to the atmosphere through a variety of activities including steel production and coal burning.[53] Arsenic release and deposition may now also be decreasing because of imposition of the Clean Air Act and its amendments since the 1970s. In addition, agricultural use of arsenic is declining.[22] These factors are potential explanations for the improvement in the arsenic failure rate in the relatively nonurban Bradford County. In addition, before the 2000s, the rate of violation of the arsenic MCL of 0.01 mg/L was 9.5%, a value that is consistent with a nationwide estimate of a 10% failure rate[22] for groundwaters sampled across the U.S. from 1987 through 1997. It is possible that today's failure rate for Bradford (4.4%) documents Bradford waters have improved relative to this national value.

Pb Concentrations. In contrast to all the other inorganic analytes that commonly can derive from bedrock, Pb often enters drinking water from household plumbing. Given that, the decline in violation rate for Pb is likely to result from the 1991 Federal Lead and Copper Rule that required the use of lead-free solder and fixtures in home plumbing[54] rather than changes in groundwater chemistry. This observation matches published conclusions[54,55] that incidence of Pb above 0.015 mg/L declined from 19% in 1989−1991 (over 1600 wells) to 12% in 2006−2007 (701 wells) in data sets from all of PA. Although waters in the Bradford-2010s data set were mostly collected by bypassing or sampling upstream of treatment systems, lead-containing solder and plumbing fixtures can often be found between the well head and the sampling point. In addition, some samples may have been taken inside houses (see

DOI: 10.1021/acs.est.8b01123
Environ. Sci. Technol. 2018, 52, 7149−7159

the SI). Thus, Pb in these groundwaters may be mostly derived from plumbing and fixtures rather than bedrock

Ba Concentrations. The median of Ba concentrations and its incidence of failing the EPA standard are similar between Bradford-1980s and Bradford-2010s, consistent with limited to no discernible impacts from spillage, leakage, or road spreading of Ba in groundwater. Ba is one of the main indicator elements for contamination by oil and gas activity because Ba is contained in oil and gas brines in PA,[11] and Ba is concentrated in drilling muds. However, brine enters water resources not just because of gas development but also by natural processes.[11] Although brines can contaminate groundwater through blow-outs, spills, or leaks,[11] or when brines from well drilling and gas production are spread on roads to suppress dust,[56,57] no evidence for such contamination was documented in the groundwater data sets here. Consistent with this, brines used for road spreading in PA are derived only from conventional oil and gas wells, and their use in road spreading decreased after 2010.[56] In addition, Skalak et al.[57] observed no increase in Ba in sediments near road spreading areas in PA. The majority of road spreading of brines was also conducted in northwestern PA and not in Bradford county.[58]

In contrast to the failure rates for Ba in the two Bradford data sets, Boyer et al.[25] reported a much lower Ba failure rate of 1% for all PA. The lower rate of failure reported in the Rural PA-2010s study compared to both Bradford studies might be because brines migrate naturally to shallow formations in northeastern PA more often than in southwestern PA.[11] Such differences could be because northeastern PA is known to be more naturally fractured as a result of tectonic deformation, glaciation, and subsequent isostatic rebound than the southern, unglaciated part of PA.[7] Thus, the Rural PA-2010 study included water samples from the relatively unfractured south where less brine contaminants have been reported. All of these observations argue against detectable Ba contamination of groundwater by unconventional gas production in Bradford since the 2000s.

In contrast to the groundwater data sets, Niu et al.[26] concluded that the median concentration of Ba in PA streams slightly increased from 1997 to 2006 to 2007−2014, and they could not eliminate contributions from surface spills and leaks from the shale-gas industry as causing this small change.

Cl, Na, and TDS Concentrations. The concentration of Cl, a useful indicator for contamination from road salt[59] and Appalachian brine,[20] is similar between Bradford-1980s and Bradford-2010s. Like Ba, this is consistent with the conclusion that the high Cl in Appalachian brines returned during shale-gas development[2,20] has not impacted groundwaters in Bradford to a statistically significant extent. Instead, the most likely sources of the Ba, Na, and Cl in Bradford groundwaters are natural processes or anthropogenic activities that have been in use prior to the onset of the shale-gas boom.

Unlike the Bradford-based conclusions, some nationwide studies show distinctly different trends. The USGS National Water Quality Assessment (NAWQA) Project[60] suggested that Cl concentrations significantly increased from 1988 to 2001 to 2002−2012 in most principal aquifers within the U.S., including the Early Mesozoic basin aquifers located in southern and eastern PA. Another nationwide surface water study[42] also documented increasing trends of Cl concentrations in U.S. streams and rivers, especially in urban areas. The most likely reason for the difference between these nationwide studies and our Bradford conclusions is differences in land cover. Specifically, only 5.7% of Bradford County is "Developed" land that is related to human activities, e.g., urban area, whereas Pennsylvania and the U.S. overall have 12.7% and 6.9% "Developed". Less road salt (quantity and frequency) is used in more rural areas such as Bradford County.

The median concentration of Na, one of the most concentrated analytes in Appalachian brine, is similar between Bradford-1980s and -2010s studies. Na concentrations (non-censored only) of Bradford-2010s generally correlate with Cl concentration with an r^2 value of 0.69 for the linear regression between these two analytes. Like Cl and Ba, the insignificant change in Na concentration over time in Bradford is consistent with little to no impact on groundwater from Appalachian brine from shale-gas development. Unlike Na and Cl, TDS concentration has improved over time in terms of both the median and the rate of failure of the EPA standard. This is consistent with water quality improvement and also argues against detectable impact of shale gas production on inorganic water quality in Bradford groundwater.

Implications. Our analysis of a data set of ∼11,000 groundwater analyses from Bradford county PA shows no statistically significant deleterious impact on ten analytes related to the aggressive increase in development of unconventional shale-gas since 2008. In fact, groundwater chemistry shows evidence of very minor improvement in comparison to a smaller data set collected prior to 1990. If groundwater is improving, likely explanations include the imposition of the Clean Air Act since the 1970s, decreases in coal mining and steel manufacture in the region, and the trend to transform power plants from coal to gas.[61]

With data mining, however, a few locations of possible methane contamination were identified in Bradford near gas wells exploiting both unconventional and conventional reservoirs. These observations were only possible because of a large set of water chemistry data that was previously unavailable to the public. Currently this large dataset does not include stable carbon isotopic data for methane. Such isotopic data would likely be helpful in determining the origin of methane, i.e., biogenic versus thermogenic, in putatively contaminated areas. Public access to these predrill data nonetheless allow better understanding of the controls, i.e., natural or anthropogenic, on groundwater chemistry. Improved public awareness of groundwater quality issues will surely contribute to designing better strategies for both water resource management and hydrocarbon exploration.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.8b01123.

Supporting Texts S1−S6, Figures S1−S10, and Tables S1−S4 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 734-730-8814. E-mail: tzw138@psu.edu.

ORCID

Tao Wen: 0000-0002-6113-7532

Xianzeng Niu: 0000-0002-1702-5381

Notes

The authors declare no competing financial interest.

DOI: 10.1021/acs.est.8b01123
Environ. Sci. Technol. 2018, 52, 7149−7159

■ ACKNOWLEDGMENTS

Funding derived from the National Science Foundation IIS-16-39150 to S.L.B. and Z.L. We thank Terry Engelder and David Yoxtheimer (Penn State University) for insights, Fei Wu and Alison Herman (Penn State University) for technical support, Bruce Lindsey (USGS) and Seth Pelepko and Stew Beattie (PA DEP) for access and help with data. We thank the editor for handling this paper as well as four anonymous reviewers for their insightful and thorough reviews.

■ REFERENCES

(1) USEIA. *Annual Energy Outlook 2016*; US Energy Information Administration: 2016.

(2) Darrah, T. H.; Vengosh, A.; Jackson, R. B.; Warner, N. R.; Poreda, R. J. Noble Gases Identify the Mechanisms of Fugitive Gas Contamination in Drinking-Water Wells Overlying the Marcellus and Barnett Shales. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (39), 14076−14081.

(3) Hildenbrand, Z. L.; Carlton, D. D.; Fontenot, B. E.; Meik, J. M.; Walton, J. L.; Taylor, J. T.; Thacker, J. B.; Korlie, S.; Shelor, C. P.; Henderson, D.; et al. A Comprehensive Analysis of Groundwater Quality in The Barnett Shale Region. *Environ. Sci. Technol.* **2015**, *49* (13), 8254−8262.

(4) Fontenot, B. E.; Hunt, L. R.; Hildenbrand, Z. L.; Carlton, D. D.; Oka, H.; Walton, J. L.; Hopkins, D.; Osorio, A.; Bjorndal, B.; Hu, Q. H.; et al. An Evaluation of Water Quality in Private Drinking Water Wells near Natural Gas Extraction Sites in the Barnett Shale Formation. *Environ. Sci. Technol.* **2013**, *47* (17), 10032−10040.

(5) Digiulio, D. C.; Jackson, R. B. Impact to Underground Sources of Drinking Water and Domestic Wells from Production Well Stimulation and Completion Practices in the Pavillion, Wyoming, Field. *Environ. Sci. Technol.* **2016**, *50* (8), 4524−4536.

(6) Osborn, S. G.; Vengosh, A.; Warner, N. R.; Jackson, R. B. Methane Contamination of Drinking Water Accompanying Gas-Well Drilling and Hydraulic Fracturing. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (20), 8172−8176.

(7) Llewellyn, G. T.; Dorman, F.; Westland, J. L.; Yoxtheimer, D.; Grieve, P.; Sowers, T.; Humston-Fulmer, E.; Brantley, S. L. Evaluating a Groundwater Supply Contamination Incident Attributed to Marcellus Shale Gas Development. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (20), 6325−6330.

(8) Jackson, R. B.; Vengosh, A.; Darrah, T. H.; Warner, N. R.; Down, A.; Poreda, R. J.; Osborn, S. G.; Zhao, K.; Karr, J. D. Increased Stray Gas Abundance in a Subset of Drinking Water Wells near Marcellus Shale Gas Extraction. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (28), 11250−11255.

(9) Li, Z.; You, C.; Gonzales, M.; Wendt, A. K.; Wu, F.; Brantley, S. L. Corrigendum to "Searching for Anomalous Methane in Shallow Groundwater near Shale Gas Wells" (J. Contam. Hydrol. (2016) 195 (23−30) (S0169772216300985) (10.1016/j.Jconhyd.2016.10.005)). *J. Contam. Hydrol.* **2017**, *207*, 50−51.

(10) Li, Z.; You, C.; Gonzales, M.; Wendt, A. K.; Wu, F.; Brantley, S. L. Searching for Anomalous Methane in Shallow Groundwater near Shale Gas Wells. *J. Contam. Hydrol.* **2016**, *195*, 23−30.

(11) Brantley, S. L. L.; Yoxtheimer, D.; Arjmand, S.; Grieve, P.; Vidic, R.; Pollak, J.; Llewellyn, G. T. T.; Abad, J.; Simon, C. Water Resource Impacts during Unconventional Shale Gas Development: The Pennsylvania Experience. *Int. J. Coal Geol.* **2014**, *126*, 140−156.

(12) Wen, T.; Castro, M. C.; Nicot, J.-P.; Hall, C. M.; Larson, T.; Mickler, P.; Darvari, R. Methane Sources and Migration Mechanisms in Shallow Groundwaters in Parker and Hood Counties, Texas—A Heavy Noble Gas Analysis. *Environ. Sci. Technol.* **2016**, *50* (21), 12012−12021.

(13) Wen, T.; Castro, M. C.; Nicot, J. P.; Hall, C. M.; Pinti, D. L.; Mickler, P.; Darvari, R.; Larson, T. Characterizing the Noble Gas Isotopic Composition of the Barnett Shale and Strawn Group and Constraining the Source of Stray Gas in the Trinity Aquifer, North-Central Texas. *Environ. Sci. Technol.* **2017**, *51* (11), 6533−6541.

(14) Molofsky, L. J.; Connor, J. A.; Wylie, A. S.; Wagner, T.; Farhat, S. K. Evaluation of Methane Sources in Groundwater in Northeastern Pennsylvania. *Groundwater* **2013**, *51* (3), 333−349.

(15) Siegel, D. I.; Azzolina, N. A.; Smith, B. J.; Perry, A. E.; Bothun, R. L. Methane Concentrations in Water Wells Unrelated to Proximity to Existing Oil and Gas Wells in Northeastern Pennsylvania. *Environ. Sci. Technol.* **2015**, *49* (7), 4106−4112.

(16) Vidic, R. D. D.; Brantley, S. L. L.; Vandenbossche, J. M. M.; Yoxtheimer, D.; Abad, J. D. D. Impact of Shale Gas Development on Regional Water Quality. *Science (Washington, DC, U. S.)* **2013**, *340* (6134), 1235009.

(17) Brantley, S. L.; Vidic, R. D.; Brasier, K.; Yoxtheimer, D.; Pollak, J.; Wilderman, C.; Wen, T.; Brantley, B. S. L.; Vidic, R. D.; Brasier, K.; et al. Engaging over Data on Fracking and Water Quality. *Science* **2018**, *359* (6374), 395−397.

(18) Meik, J. M.; Lawing, A. M. *Considerations and Pitfalls in the Spatial Analysis of Water Quality Data and Its Association With Hydraulic Fracturing*; Elsevier: 2017; Vol. *1*, DOI: 10.1016/bs.apmp.2017.08.013.

(19) Rhodes, A. L.; Horton, N. J. Establishing Baseline Water Quality for Household Wells within the Marcellus Shale Gas Region, Susquehanna County, Pennsylvania, USA. *Appl. Geochem.* **2015**, *60*, 14−28.

(20) Warner, N. R.; Jackson, R. B.; Darrah, T. H.; Osborn, S. G.; Down, A.; Zhao, K.; White, A.; Vengosh, A. Geochemical Evidence for Possible Natural Migration of Marcellus Formation Brine to Shallow Aquifers in Pennsylvania. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (30), 11961−11966.

(21) Darrah, T. H.; Jackson, R. B.; Vengosh, A.; Warner, N. R.; Whyte, C. J.; Walsh, T. B.; Kondash, A. J.; Poreda, R. J. The Evolution of Devonian Hydrocarbon Gases in Shallow Aquifers of the Northern Appalachian Basin: Insights from Integrating Noble Gas and Hydrocarbon Geochemistry. *Geochim. Cosmochim. Acta* **2015**, *170*, 321−355.

(22) Welch, A. H.; Westjohn, D. B.; Helsel, D. R.; Wanty, R. B. Arsenic in Ground Water of the United States: Occurrence and Geochemistry. *Groundwater* **2000**, *38* (4), 589−604.

(23) ShaleNetwork. DOI: 10.4211/his-data-shalenetwork.

(24) Williams, J. H.; Taylor, L. E.; Low, D. J. *Hydrogeology and Groundwater Quality of the Glaciated Valleys of Bradford, Tioga, and Potter Counties, Pennsylvania*; 1998, No. *68*.

(25) Boyer, E. W.; Swistock, B. R.; Clark, J.; Madden, M.; Rizzo, D. E. *The Impact of Marcellus Gas Drilling on Rural Drinking Water Supplies*; Center for Rural Pennsylvania: 2012.

(26) Niu, X.; Wendt, A.; Li, Z.; Agarwal, A.; Xue, L.; Gonzales, M.; Brantley, S. L. Detecting the Effects of Coal Mining, Acid Rain, and Natural Gas Extraction in Appalachian Basin Streams in Pennsylvania (USA) through Analysis of Barium and Sulfate Concentrations. *Environ. Geochem. Health* **2018**, *40* (2), 865−885.

(27) Field, A. P. *Discovering Statistics Using SPSS*; SAGE Publications: 2009.

(28) Balashov, V. N.; Engelder, T.; Gu, X.; Fantle, M. S.; Brantley, S. L.; et al. A Model Describing Flowback Chemistry Changes with Time after Marcellus Shale Hydraulic Fracturing. *AAPG Bull.* **2015**, *99* (1), 143−154.

(29) USEPA. *Drinking Water Advisory: Consumer Acceptability Advice and Health Effects Analysis on Sodium*; Washington, DC, 2003.

(30) Helsel, D. R. *Statistics for Censored Environmental Data Using Minitab and R*; John Wiley & Sons: 2011; Vol. *77*, DOI: 10.1002/9781118162729.

(31) Reese, S. O.; Neboga, V. V.; Pelepko, S.; Kosmer, W. J.; Beattie, S. *Groundwater and Petroleum Resources of Sullivan County, Pennsylvania*; Water Resource Report; 2014.

(32) Berg, T. M.; Edmunds, W. E.; Geyer, A. R.; Glover, A. D.; Hoskins, D. M.; MacLachlan, D. B.; Root, S. I.; Sevon, W. D.; Socolow, A. A.; Miles, C. E. *Geologic Map of Pennsylvania: Pennsylvania Geological Survey*, 4th Series; 1980.

DOI: 10.1021/acs.est.8b01123
Environ. Sci. Technol. 2018, 52, 7149−7159

(33) Pohn, H.; Purdy, T. A Major (?) Thrust Fault at Towanda. *Pennsylvania An Ex. faulting with some Specul. Struct. Allegheny Plateau 46th Annu. F. Conf. Pennsylvania Geol. Guid. Geol. Tioga Bradford Counties, Pennsylvania* 1981, 45−56.
(34) Woodrow, D. L. *Stratigraphy, Structure, and Sedimentary Patterns in the Upper Devonian of Bradford County, Pennsylvania*; 1968.
(35) Molofsky, L. J.; Connor, J. A.; McHugh, T. E.; Richardson, S. D.; Woroszylo, C.; Alvarez, P. J. Environmental Factors Associated With Natural Methane Occurrence in the Appalachian Basin. *Groundwater* **2016**, *54* (5), 656−668.
(36) PADEP. PA DEP Oil&Gas Compliance Report http://www.depreportingservices.state.pa.us/ReportServer/Pages/ReportViewer.aspx?/Oil_Gas/OG_Compliance (accessed Jan 1, 2018).
(37) Kang, M.; Kanno, C. M.; Reid, M. C.; Zhang, X.; Mauzerall, D. L.; Celia, M. A.; Chen, Y.; Onstott, T. C. Direct Measurements of Methane Emissions from Abandoned Oil and Gas Wells in Pennsylvania. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (51), 18173−18177.
(38) Gross, E. L.; Cravotta, C. A. *Groundwater Quality for 75 Domestic Wells in Lycoming County, Pennsylvania*; 2014 Scientific Investigations Report 2016-5143; 2017; DOI: 10.3133/sir20165143.
(39) Edmunds, W. M.; Shand, P.; Hart, P.; Ward, R. S. The Natural (Baseline) Quality of Groundwater: A UK Pilot Study. *Sci. Total Environ.* **2003**, *310* (1), 25−35.
(40) Ayotte, J. D.; Belaval, M.; Olson, S. A.; Burow, K. R.; Flanagan, S. M.; Hinkle, S. R.; Lindsey, B. D. Factors Affecting Temporal Variability of Arsenic in Groundwater Used for Drinking Water Supply in the United States. *Sci. Total Environ.* **2015**, *505*, 1370−1379.
(41) Thundiyil, J. G.; Yuan, Y.; Smith, A. H.; Steinmaus, C. Seasonal Variation of Arsenic Concentration in Wells in Nevada. *Environ. Res.* **2007**, *104* (3), 367−373.
(42) Stets, E. G.; Lee, C. J.; Lytle, D. A.; Schock, M. R. Increasing Chloride in Rivers of the Conterminous US and Linkages to Potential Corrosivity and Lead Action Level Exceedances in Drinking Water. *Sci. Total Environ.* **2018**, *613-614*, 1498−1509.
(43) Strock, K. E.; Nelson, S. J.; Kahl, J. S.; Saros, J. E.; McDowell, W. H. Decadal Trends Reveal Recent Acceleration in the Rate of Recovery from Acidification in the Northeastern U.S. *Environ. Sci. Technol.* **2014**, *48* (9), 4681−4689.
(44) Paulot, F.; Fan, S.; Horowitz, L. W. Contrasting Seasonal Responses of Sulfate Aerosols to Declining SO2emissions in the Eastern U.S.: Implications for the Efficacy of SO2emission Controls. *Geophys. Res. Lett.* **2017**, *44* (1), 455−464.
(45) Kaushal, S. S.; Likens, G. E.; Pace, M. L.; Utz, R. M.; Haq, S.; Gorman, J.; Grese, M. Freshwater Salinization Syndrome on a Continental Scale. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E574−E583.
(46) Raymond, P. A.; Oh, N. H. Long Term Changes of Chemical Weathering Products in Rivers Heavily Impacted from Acid Mine Drainage: Insights on the Impact of Coal Mining on Regional and Global Carbon and Sulfur Budgets. *Earth Planet. Sci. Lett.* **2009**, *284* (1−2), 50−56.
(47) Yager, R. M. *Estimation of Hydraulic Conductivity of a Riverbed and Aquifer System on the Susquehanna River in Broome County, New York*; USGPO; US Geological Survey, Book and Open-File Report Sales [distributor]: 1993.
(48) Manger, G. E. *Porosity and Bulk Density of Sedimentary Rocks*; 1963.
(49) Herndon, E. M.; Brantley, S. L. Movement of Manganese Contamination through the Critical Zone. *Appl. Geochem.* **2011**, *26*, S40−S43.
(50) Gillispie, E. C.; Austin, R. E.; Rivera, N. A.; Bolich, R.; Duckworth, O. W.; Bradley, P.; Amoozegar, A.; Hesterberg, D.; Polizzotto, M. L. Soil Weathering as an Engine for Manganese Contamination of Well Water. *Environ. Sci. Technol.* **2016**, *50* (18), 9963−9971.
(51) Herndon, E. M.; Jin, L.; Andrews, D. M.; Eissenstat, D. M.; Brantley, S. L. Importance of Vegetation for Manganese Cycling in Temperate Forested Watersheds. Global Biogeochem. *Cycles* **2015**, *29* (2), 160−174.
(52) Bilinski, H.; Bruins, R. J. F.; Erdreich, L.; Fugas, M.; Kello, D. *Health Assessment Document for Manganese. Final Report*; Environmental Protection Agency, Cincinnati, OH (USA). Office of Environmental Criteria and Assessment: 1984.
(53) Matschullat, J. Arsenic in the Geosphere - A Review. *Sci. Total Environ.* **2000**, *249*, 297−312.
(54) Swistock, B. R.; Clemens, S. Water Quality and Management of Private Drinking Water Wells in Pennsylvania. *J. Environ. Health* **2013**, *75* (6), 60.
(55) Swistock, B. R.; Sharpe, W. E.; Robillard, P. D. A Survey of Lead, Nitrate and Radon Contamination of Private Individual Water Systems in Pennsylvania. *J. Environ. Health* **1993**, *55* (5), 6−13.
(56) Rahm, B. G.; Bates, J. T.; Bertoia, L. R.; Galford, A. E.; Yoxtheimer, D. A.; Riha, S. J. Wastewater Management and Marcellus Shale Gas Development: Trends, Drivers, and Planning Implications. *J. Environ. Manage.* **2013**, *120*, 105−113.
(57) Skalak, K. J.; Engle, M. A.; Rowan, E. L.; Jolly, G. D.; Conko, K. M.; Benthem, A. J.; Kraemer, T. F. Surface Disposal of Produced Waters in Western and Southwestern Pennsylvania: Potential for Accumulation of Alkali-Earth Elements in Sediments. *Int. J. Coal Geol.* **2014**, *126*, 162−170.
(58) PADEP. PA DEP Waste Report. http://www.depreportingservices.state.pa.us/ReportServer/Pages/ReportViewer.aspx?%2FOil_Gas%2FOil_Gas_Well_Waste (accessed Jan 1, 2018).
(59) Conway, T. M. Impervious Surface as an Indicator of PH and Specific Conductance in the Urbanizing Coastal Zone of New Jersey, USA. *J. Environ. Manage.* **2007**, *85* (2), 308−316.
(60) Lindsey, B. D.; Johnson, T. D.; Belitz, K. Decadal changes in groundwater quality: USGS Web page. https://nawqatrends.wim.usgs.gov/Decadal/ (accessed May 29, 2018).
(61) Massetti, E.; Brown, M. A.; Lapsa, M.; Sharma, I.; Bradbury, J.; Cunliff, C.; Li, Y. *Environmental Quality and the U.S. Power Sector: Air Quality, Water Quality, Land Use and Environmental Justice*; 2017.

DOI: 10.1021/acs.est.8b01123
Environ. Sci. Technol. 2018, 52, 7149−7159