Environmental Science & Technology

pubs.acs.org/est

Article

Geochemical Evidence of Potential Groundwater Contamination with Human Health Risks Where Hydraulic Fracturing Overlaps with Extensive Legacy Hydrocarbon Extraction

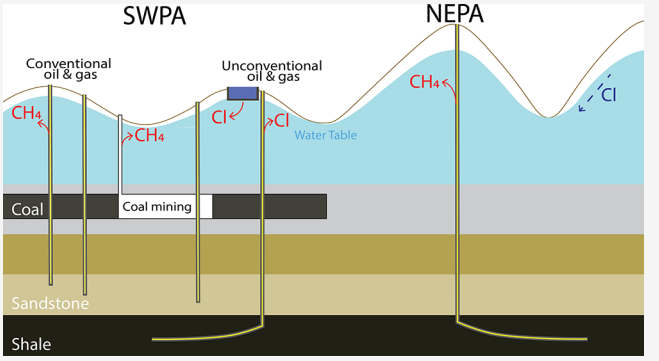
Samuel W. Shaheen,* Tao Wen, Alison Herman, and Susan L. Brantley

Cite This: https://doi.org/10.1021/acs.est.2c00001

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Unconventional oil and gas development (UOGD) sometimes impacts water resources, including incidents of methane (CH_4) migration from compromised wells and spills that degrade water with salts, organics, and metals. We hypothesized that contamination may be more common where UOGD overlaps with legacy coal, oil, and gas extraction. We tested this hypothesis on ~7000 groundwater analyses from the largest U.S. shale gas play (Marcellus), using data mining techniques to explore UOGD contamination frequency. Corroborating the hypothesis, we discovered small, statistically significant regional correlations between groundwater chloride concentrations ([Cl]) and UOGD proximity and density where legacy extraction was extremely dense (southwestern Pennsylvania (SWPA)) but no such correlations where it was minimal (northeastern Pennsylvania). On the other hand, legacy extraction of shallow gas in SWPA may have lessened today's gas leakage, as no regional correlation was detected for [CH_4] in SWPA. We identify hotspots where [Cl] and [CH_4] increase by 3.6 and 3.0 mg/L, respectively, per UOG well drilled in SWPA. If the [Cl] correlations document contamination via brines leaked from wellbores, impoundments, or spills, we calculate that thallium concentrations could exceed EPA limits in the most densely developed hotspots, thus posing a potential human health risk.



KEYWORDS: *groundwater, oil and gas, methane, brine, shale*

INTRODUCTION

The rapid expansion of unconventional oil and gas development (UOGD) in the United States has benefited U.S. energy independence, but also spurred public and scientific debate surrounding its associated environmental impacts, for example, the occasional deterioration of groundwater quality in UOG production areas.[1−3] Compounding such concerns are human health impacts putatively linked to exposure to UOGD, such as increased risk of poor birth outcomes including low birth weights and preterm deliveries.[4−6] While concrete links between negative health effects and contaminated water supplies due to UOGD have not been conclusively demonstrated, the ingestion of UOGD contaminants via water supplies is one possible exposure pathway, particularly considering 37% of UOG wells in the U.S. stimulated in 2014 were located within 2 km of at least one domestic water supply well.[6,7]

The migration of methane (CH_4), the primary component of natural gas, into shallow groundwater is the most cited water quality issue associated with UOGD in the Marcellus Shale of Pennsylvania (PA), the largest U.S. shale gas play, and has been noted in other shale gas basins nationwide.[8,9] CH_4 leakage typically results from improper well construction (i.e., lacking casings to prevent gas migration from overpressured formations) or well integrity issues (i.e., defective casings or gas migration along the well annulus).[9−11] Additionally, CH_4 can reach groundwater wells several kilometers away from a leaking UOG well via migration along faults or fractures.[3,9] CH_4 is nontoxic and common in shallow groundwater throughout hydrocarbon-bearing basins prior to UOGD due to the ubiquity of microbial methanogenesis[12−14] and upward migration of thermogenic CH_4 (produced via the thermal maturation of organic matter at depth) over geologic timescales.[15−17] However, recently migrated CH_4 from leaking UOG wells poses an explosion hazard above 10 mg/L and can induce redox effects that mobilize toxic species (e.g., arsenic) or lead to the formation of toxic compounds (e.g., sulfide).[2,3]

Received: January 1, 2022
Revised: May 20, 2022
Accepted: June 3, 2022

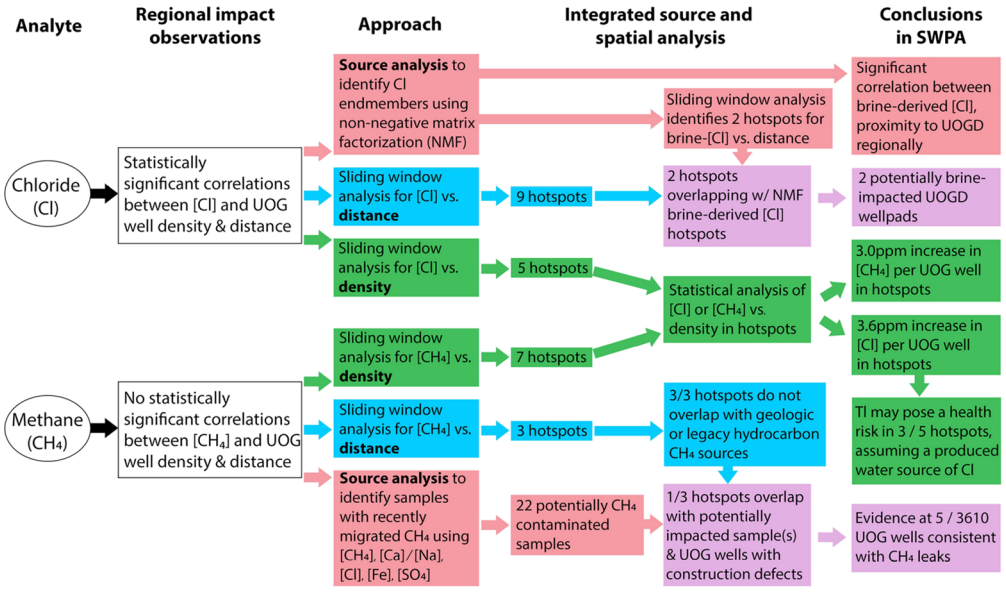


Figure 1. Flowchart displaying the workflow used to determine the potential impacts of UOGD on methane and chloride concentrations in SWPA. Red coloration indicates geochemical source analysis used to constrain the sources of $[CH_4]$ or [Cl] in groundwater samples. Blue coloration denotes spatial analysis considering correlations between $[CH_4]$ or [Cl] and UOG well distance, while green denotes spatial analysis based on UOG well density.

However, studies assessing the relationship between groundwater methane concentrations ($[CH_4]$) and UOGD often reach differing conclusions due to differences in the data set size. Some studies in northeastern Pennsylvania (NEPA) based on small data sets (<200 groundwater sampling sites) reported statistically significant increases in $[CH_4]$ nearby UOG wells,[1,18] while others documented no increase in $[CH_4]$ associated with UOGD.[19] On the other hand, larger data sets (e.g., >5000 samples) typically do not show strong regionwide correlations because contamination tends to be rare and is often obscured by overlapping natural and anthropogenic sources.[13,20] This highlights that although smaller data sets are needed to confirm hotspots of contamination, large data sets and new approaches for spatial data mining are needed to assess contamination frequency. Given this context, data mining techniques are useful because they can highlight hotspots where $[CH_4]$ increases nearby UOG wells within large data sets.[21−23] These sites, if they indeed represent loci of UOGD-driven contamination, may be settings with ongoing leakages that pose risks to those using the aquifer for drinking water. Data mining on large geochemical data sets is thus perhaps the most accurate method of assessing the frequency of UOGD-related water contamination.

Additional water quality concerns surround contamination from UOGD wastewaters that often contain high concentrations of brine salts and toxic species.[2,24] These wastewaters include flowback waters that surface for several weeks after hydraulic fracturing and produced waters that surface throughout the production life of the well.[2] Recent studies have documented small but statistically significant increases in chloride, barium, and strontium concentrations in surface waters associated with UOGD across shale gas basins, suggesting that brine contamination during UOGD can impact regional surface water quality.[25,26] Like CH_4, brine-related species (e.g., sodium, chloride, barium, and strontium) are occasionally present in water supplies prior to UOGD because of the natural upward migration of basin brines across geologic formations[15,17,27] or the slow flushing of connate saline groundwater by meteoric recharge.[28,29] Nonetheless, studies have demonstrated that UOG wastewater leaks or spills in some locations have contaminated nearby surface waters with salts and toxic species.[30−32] Groundwater supplies have also been impacted by leaking UOGD wastewater impoundment pits.[8,33] However, the frequency of groundwater contamination by UOGD brines remains poorly constrained.

Many studies on UOGD and water quality have primarily emphasized NEPA, a predominantly rural area where legacy hydrocarbon production such as coal mining or conventional oil and gas development (COGD) is limited.[1,3,13,23,34] In contrast, UOG wells are often drilled near COGD wellpads or coal mines in many gas-producing regions. Overlap of these land uses increases the difficulty in discerning contaminant sources in groundwater because both COGD and coal beds have also been associated with elevated groundwater $[CH_4]$.[35−37]

We hypothesized that interactions between UOGD and other forms of hydrocarbon extraction may provide additional pathways for contaminants from UOGD to migrate into water resources. For example, 90% of gas migration-related incidents in Alberta, Canada, were attributed to oil and gas wells that penetrated coal beds.[38] Furthermore, poorly maintained or abandoned COG wells can connect aquifers otherwise separated by aquitards and create conduits for CH_4 or brine to reach the surface.[39−42] In this research, we investigated whether the frequency and magnitude of UOGD-related groundwater contamination are exacerbated where UOGD overlaps with areas of dense coal mining and aging COG wells.

To investigate this hypothesis, we utilized a large groundwater chemistry data set of 6991 samples from Washington, Greene, and Beaver counties in southwestern Pennsylvania, U.S. (SWPA). SWPA has >100 years of COG production and coal mining, in addition to some of the highest density of UOGD in the U.S. (see the Supporting Information for more information on the study area). Additionally, ≥99% of

https://doi.org/10.1021/acs.est.2c00001
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

domestic groundwater wells in Greene and Washington counties are located within 1 km of an actively producing oil and gas well as of 2014,[7] highlighting water as a potential exposure pathway for the adverse health effects associated with UOGD in SWPA.[4,43] To search for both regional impacts and specific locations of potential contamination, we utilized a workflow incorporating spatial and source analysis to distinguish the water quality impacts of UOGD (Figure 1). We emphasize CH_4 and chloride (Cl), two analytes considered to be indicative of UOGD-related contamination in national data sets.[9,25] CH_4 contamination typically results from subsurface leakage,[3,9] while Cl contamination typically results from surface leaks or spills.[30,31] Thus, considering both CH_4 and Cl provides insight into subsurface vs surface contamination pathways associated with UOGD.

We first regionally analyzed whether analyte concentrations increase with (a) the density of UOG wells within a 1 km radius or (b) the proximity to UOG wells. Next, we assessed the spatial extent of contamination by applying the sliding window geospatial tool (SWGT), a geostatistics-based data mining tool, to identify subregions where the concentration of a species (e.g., CH_4) correlates with the (a) density of or (b) proximity to UOG wells. We refined our estimate of CH_4 sources using a geochemical protocol based on the concentrations and ratios of salinity- and redox-related species that identifies samples with recently migrated CH_4. Given that multiple Cl sources are known to contaminate Appalachian Basin water resources,[44,45] we used a machine learning approach, non-negative matrix factorization (NMF), to determine Cl sources and the mixing proportions of waters from these sources in our sample set. NMF can delineate contaminant sources in both groundwater and surface water without prior knowledge of source chemical compositions.[46−48] The results of these spatial and source analyses were considered in tandem (Figure 1) to refine estimates of the regional impact of UOGD on $[CH_4]$ and [Cl] and the number of UOGD sites that may be leaking CH_4 or Cl to groundwater.

MATERIALS AND METHODS

Data Sets. The groundwater quality data set was provided to Pennsylvania State University by the Pennsylvania Department of Environmental Protection (PADEP). In work financed by oil and gas companies, samples were collected from private wells or springs by certified environmental consultants prior to UOGD and analyzed by accredited commercial laboratories. While these are "predrill" samples for their associated UOG well, the high density of UOGD in SWPA means that these samples are "postdrill" for neighboring UOG wells, and our calculations consider only UOG wells drilled prior to the collection of the respective water sample. Water sample chemistry data were compiled, cleaned, and quality-controlled (as outlined in Wen et al.[23]) and are added to a master data set (available at https://doi.org/10.4211/his-data-shalenetwork and https://doi.org/10.26208/2nqe-wd53). The data set used in this study consists of 6991 groundwater samples, with 4325 collected from domestic wells and 2666 from springs (Figure S1D). Additional data sources are described in the Supporting Information.

Regional Analysis. We assessed regional correlations between analyte concentrations and UOG well distance or density based on the calculated Kendall rank correlation and, for calculations considering well density, the Akritas−Theil−Sen (ATS) regression slope. For each calculation, p-values were examined. These statistical tests were selected because of their utility in evaluating data sets with multiple detection limits[21] and evaluated using the cenken function in the R package NADA (density) or the ktau_p package within the SWGT (distance).

Sliding Window Geospatial Tool. The SWGT was developed to assess localized impacts by calculation of the spatially averaged correlation between $[CH_4]$ in groundwater and the proximity to features such as UOG wells.[21] Source codes of the SWGT are available online (https://github.com/jaywt/SWGT). Using the SWGT, we investigated groundwater $[CH_4]$ or [Cl] and their relationship with respect to (a) UOG well density within 1 km of the sample or (b) UOG well proximity. The SWGT moves a 5 km x 5 km "window" across the study area in 200 m steps while iteratively calculating the Kendall rank correlation between analyte concentrations and either density or proximity of UOG wells in the window. If the correlation is statistically significant ($p < 0.05$), a value of +1 (positive relationship) or −1 (negative relationship) is assigned to the window and then summed for every location. Relative significance values are calculated for every location by dividing the sum of the windows covering the location by the number of windows and plotted as a heat map. For (a), we sought areas showing positive correlations (i.e., concentrations increase with increased UOG well density). We emphasize "hotspots," defined as localities with a relative frequency of significant correlations ≥0.7. For (b), we sought hotspots with a strong negative correlation (i.e., concentrations increase as the distance to UOG wells decreases), defined as a relative frequency of significant correlations ≤ −0.7. We demonstrated the utility of the SWGT using a test data set from NEPA that includes samples from a well-known locality of putative contamination, Dimock (Supporting Information).

NMF Source Attribution. NMF finds patterns in large water chemistry data sets and identifies endmember water types. It then delineates the mixing proportions, α, and compositions of the endmembers in every sample. Unlike traditional mixing models, NMF does not require a priori knowledge of compositions of endmember sources, nor does it require these compositions to be invariant. Adapting a published approach,[48] we used NMF to explore sources of Cl based on the molar ratios of major cations and anions (Ba, Ca, Mg, Na, and SO_4) with respect to Cl. For further methodology of the NMF model, see the Supporting Information. To test the approach, we used NMF to show that we could successfully distinguish Cl sources from a published synthetic data set (Supporting Information).[44]

RESULTS

Regional-Scale Impacts of UOGD on $[CH_4]$ and [Cl]. We first investigated potential regional impacts of UOGD on groundwater $[CH_4]$ or [Cl] using the Kendall rank correlation between species concentrations and UOG well distance or density. A statistically insignificant (p-value > 0.05) negative correlation between $[CH_4]$ and the distance to the closest UOG well and a statistically insignificant positive correlation between $[CH_4]$ and UOG well density were identified (Table 1). However, we identified a statistically significant (i.e., p-value < 0.05) negative correlation between [Cl] and the distance to the nearest UOG well and positive correlation between [Cl] and UOG well density. In other words, we observed increases in [Cl] associated with proximity to the

https://doi.org/10.1021/acs.est.2c00001
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Table 1. Correlations between [CH_4] or [Cl] and the Distance to Density of UOG Wells

calculation	distance to UOG wells vs		density of UOG wells vs	
species	[CH_4]	[Cl][a]	[CH_4]	[Cl][a]
Kendall's τ	−0.009	−0.036	0.011	0.033
p-value	0.219	<0.001	0.084	<0.001
ATS slope			0.038	0.33

[a]Statistically significant correlation.

nearest UOG well and increased density of UOGD across SWPA. We also identified significant correlations between [CH_4] or [Cl] and proximity to COG wells in SWPA (Supporting Information).

Localized Increases in [CH_4] and [Cl] with Unconventional Well Density. Sliding window analysis identified seven hotspots where [CH_4] increases with the number of UOG wells within 1 km (i.e., well density) and five hotspots where [Cl] increases with well density (depicted in green in the Figure 1 workflow, hotspots shown in maroon in Figure 2). When considering only the 506 samples in the seven hotspots for [CH_4] vs UOG well density, a significant correlation was observed (Kendall's τ = 0.029, ATS slope = 3.0, p-value = 0.023). For those 506 samples, the ATS slope predicts a 3.0 mg/L increase in [CH_4] for every additional UOG well drilled within 1 km of a groundwater sample site. Likewise, samples within Cl hotspots (n = 421) show a significant correlation between UOG well density and [Cl] (Kendall's τ = 0.251, ATS slope = 3.6, p-value < 0.001). An ATS slope of 3.6 thus predicts a 3.6 mg/L increase in [Cl] per every UOG well drilled within 1 km of these hotspots.

In Figure 2, hotspots with a strong negative correlation between UOGD density and [CH_4] or [Cl] were identified. These hotspots could result if UOGD contamination occurs in low-UOGD density areas or reflect the long-distance hydrogeologic transport of contaminants away from UOGD wellpads located along ridges. Alternatively, these hotspots may result from natural hydrogeologic processes. For example, the tendency to drill UOGD wellpads along ridgetops could result in a negative correlation if natural brine is upwelling in valley bottoms, as this would result in increasing [CH_4] or [Cl] in typically low-UOGD density localities.

[CH_4] Significantly Increases Nearby Hydrocarbon Infrastructure in Hotspots. Sliding window analysis (blue in Figure 1) revealed three hotspots (~5 to 16 km^2) where [CH_4] significantly increases with proximity to UOG wells (maroon coloration in Figure 3A). Increased [CH_4] in these hotspots could be caused by UOGD, although other explanations are possible. For example, [CH_4] increases with proximity to COG wells in two hotspots (Figure S2) and increases nearby anticlines, large folds of geologic strata that can trap naturally emitted CH_4 migrating upward from depth, in four hotspots (Figure S4).[23] In addition, a few larger hotspots are located near coal mining (Figure S4). We therefore used a source attribution technique (shown in red in the workflow in Figure 1) that can highlight samples with newly migrated CH_4 likely from recent UOGD.[49]

To do this, we focused on the concentrations of redox- and salinity-related species in the 382 /6991 groundwater samples that contained [CH_4] ≥ 1 mg/L (see the Supporting Information for more details). Out of those 382 samples, 350 show evidence for a CH_4 source associated with natural brine migration. Of the 32 samples where CH_4 does not appear associated with natural brine migration, only 22 samples show transiently high concentrations of iron ([Fe]) and sulfate ([SO_4]), as observed after recent migration of CH_4 from leaking UOGD wells.[3,49] Of these 22 samples, 3 were from hotspots in Figure 3A (Figure S5). The two hotspots containing these three samples are hence considered the most likely locations of CH_4 leakage from UOG wells detected in the 6991-sample data set.

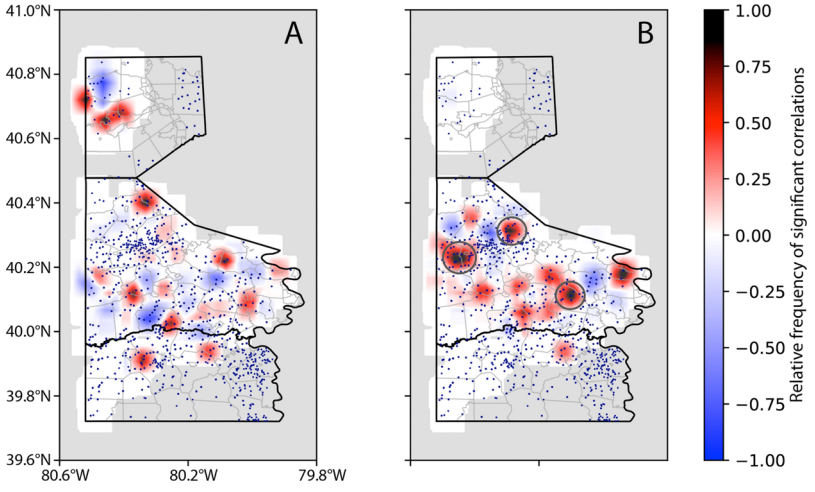


Figure 2. Sliding window heatmaps for the correlation between (A) [CH_4] and (B) [Cl] and the density of unconventional wells within a 1 km radius of water samples in Beaver (top), Washington (middle), and Greene (bottom) counties in Pennsylvania, U.S. The locations of UOG wells are denoted with blue dots on the map, and the locations of water samples considered in this analysis are shown in Figure S1. Areas with sufficient data density for sliding windows are colored in white, while areas lacking sufficient sample coverage (and thus not included in the calculation) are colored in gray. Blue shading indicates negative correlations ([analyte] decreases as UOG well density increases) while red shading indicates positive correlations ([analyte] increases as UOG well density increases). Color intensity corresponds to the relative frequency of significant positive (red) or negative (blue) correlations. If Cl derives from a UOGD wastewater source with average chemistry as reported for SWPA, the four hotspots circled in dark gray in 2B are localities where the contamination may be sufficient to elevate thallium concentrations above the U.S. Environmental Protection Agency maximum contaminant level.

https://doi.org/10.1021/acs.est.2c00001
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

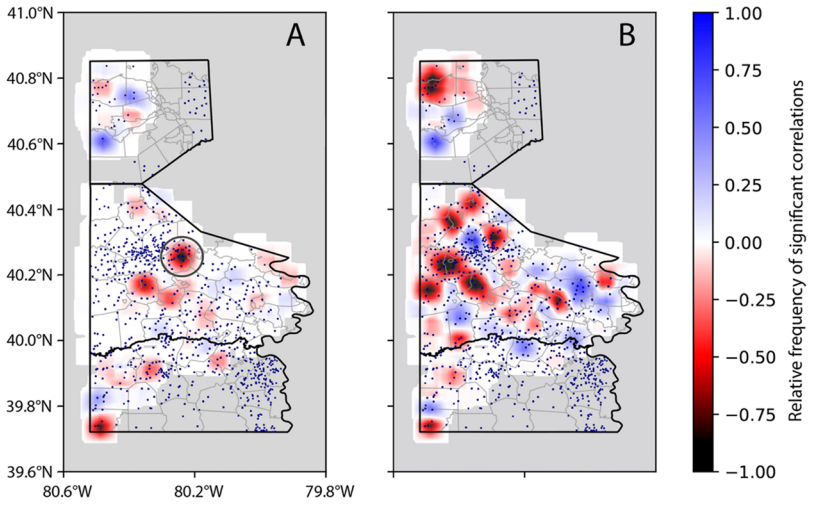


Figure 3. Sliding window heat maps showing the relative frequency of statistically significant correlations between (A) methane concentrations ([CH_4]) or (B) chloride concentrations ([Cl]) and the distance to the nearest unconventional oil and gas well. Red indicates an area with a negative relationship ([CH_4] or [Cl] increases as the distance to a UOG well decreases) and blue areas signify a positive relationship ([CH_4] or [Cl] decreases as the distance to a UOG well decreases). Color intensity corresponds to the relative frequency of significant correlations in the area. UOG well locations are denoted with blue dots on the map. Within the hotspot circled in 3A, multiple lines of evidence are consistent with CH_4 leaking from a UOG well/wells to groundwater.

[Cl] Increases Nearby UOGD in Hotspots and Nearby Highways across SWPA. Sliding window analysis revealed nine subregions where [Cl] significantly increases near UOG wells (Figure 3B). However, [Cl] also increases near COG wells in three hotspots (Figure S2) and nearby highways across large swaths of SWPA (Figure S6). We therefore used NMF to delineate sources of Cl in the water samples. NMF revealed three endmembers for Cl: a Ba-rich endmember, a Na-rich endmember, and a Ca, Mg, and SO_4-rich endmember (Table S1). Based on local geochemical knowledge, we attributed these to brine, road salt, and meteoric recharge water (including other geochemically nondistinct Cl sources such as organic wastes), respectively. The Ba-rich endmember is attributed to a brine source because sedimentary basin brines, flowback, and produced waters often contain characteristically high Ba concentrations.[8,26] Similarly, the Na-rich endmember was attributed to road salting because the rock salt used for PA winter road preparation is ~90 to 98% NaCl and Ba-poor.[50] Finally, the Ca, Mg, and SO_4-rich endmember is chemically consistent with meteoric recharge groundwater in SWPA where Cl is likely rain- and organic waste-derived.[29,44]

Across all samples, the mean mixing proportions for the three endmembers of Cl were 38% from road salting ($\alpha_{\text{road salt}}$), 29% from brine ($\alpha_{\text{brine}}$), and 33% from meteoric recharge (α_{MR}). These proportions indicate that, on average, 29% of the chloride in an individual sample originated from a brine source, rather than the average sample containing 29% brine.

DISCUSSION

UOGD Is Associated with a Slight Increase in Groundwater [Cl] across SWPA. Groundwater [Cl] is statistically correlated with UOG well density and proximity in a region within the Marcellus Shale play (SWPA) where the legacy of oil, gas, and coal development is long and intense. Such a regional correlation in a large groundwater data set has not previously been noted for any shale gas basin although a correlation was reported for [Cl] and UOGD in a national data set of surface water.[25] If the correlation means that Cl is leaking in the study area, the leakage is not evenly distributed because we only detected significant correlations between [Cl] and UOG well density or proximity in five (Figure 2B) and nine hotspots (Figure 3B), respectively. In other words, a small number of localities where UOGD wastewaters migrated or spilled may produce the regional effects we observe, rather than evenly distributed contamination across SWPA.

If brine spills or leaks explain the correlation between [Cl] and UOGD in SWPA, then we should also detect significant correlations for other species present at high concentrations in flowback and produced waters and UOGD. For example, barium (Ba) is present in high concentrations in UOGD brines and is often an effective tracer for oil and gas activity.[26] We tested and discovered a significant regional increase in [Ba] nearby UOG wells across SWPA (Kendall's $\tau = -0.058$, p-value < 0.001). Likewise, [Ba] increased with UOG well density across the study area (Kendall's $\tau = 0.056$, ATS slope = 0.0029, p-value < 0.001). Concentrations of strontium ([Sr]), another species used to investigate UOGD wastewater impacts,[25] similarly increase regionwide with proximity to UOGD (Kendall's $\tau = -0.025$, p-value = 0.003) and with UOG well density (Kendall's $\tau = 0.026$, ATS slope = 0.0050, p-value < 0.001). SWGT tests using [Ba] and [Sr] are consistent with these conclusions (Figure S9A,B).

Not all brine migration is due to human activity, however, as brine is known to upwell naturally. Because brine species are often present in valley bottoms across the Appalachian Basin,[27,29] increases in [Cl] nearby UOG wells could be related to this natural hydrogeologic process if UOG wells are mostly located in topographic lows. However, UOG wells in SWPA are predominantly drilled at topographic highs, whereas mean [Cl] shows no significant difference between topographic position index classes (Figure S8). Thus, topography and hydrogeologic processes alone likely cannot explain the significant increase in [Cl] nearby UOGD wells in SWPA.

On the other hand, some hotspots in Figure 3B overlap with hotspots in the SWGT heatmap for [Cl] vs distance to the nearest highway (Figure S6). To separate UOGD salinization from road salting, the SWGT was re-run for the portion of [Cl] attributed to brine by NMF. When run using only this portion,

https://doi.org/10.1021/acs.est.2c00001
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

two hotspots were still detected (Figure S9C). The regionwide correlation between [Cl] and distance to UOG wells persists for SWPA when analyzed using only the brine-derived [Cl] (Kendall's $\tau = -0.052$, p-value ≤ 0.001), although we detected fewer hotspots than in Figure 3B. As such, road salting does not appear to be the explanation for the correlation we observe.

A defensible explanation for the regional-scale increase in [Cl] associated with UOGD in SWPA is thus UOGD itself, as this correlation appears unrelated to topography and persists even after excluding road salt as a source. Several different pathways, such as boreholes that allow brine leakage upward, wellpad spills, or leaking impoundment pits, may explain the contamination. Most researchers have concluded that the likeliest scenario for brine contamination is large spillage at UOGD wellpads.[25,30,32] The high frequency of UOGD wastewater spills further supports this mechanism, as almost 1300 spills associated with natural gas extraction in PA were reported from 2007 to 2014.[51] In addition, three hotspots in Figure 3B directly overlap with spills in the data set by Patterson et al.,[51] and 5 are within 2 km of a spill (a plausible distance for solute transport along fractures in PA[8]).

While spills may be the likeliest pathway for brines to contaminate water resources during UOGD, we cannot exclude other mechanisms. Infiltration of UOGD wastewaters from poorly lined or unlined impoundment pits has also been associated with Cl contamination in SWPA.[33] Poorly cemented UOG well casings, though responsible for gas migration incidents, are considered unlikely as mechanisms of leakage of flowback or produced waters into groundwater because brine is dense and tends to not migrate quickly upward like buoyant gas.[52] However, if Appalachian Basin brines are currently moving upward naturally, as suggested by some, this mechanism could entrain brine and CH_4 leaking from new wells.[27,45] Modeling studies suggest that abandoned wells may provide a more efficient conduit for deep brines to reach water resources during hydraulic fracturing in instances where "out-of-zone" stimulation enables connectivity between unconventional drilling and conventional wells.[53,54] While some field observations also suggest that brines may surface via leaking abandoned wells,[41,42] the frequency of this phenomenon remains poorly constrained, making it difficult to assess whether this pathway may explain increased [Cl] nearby UOGD in SWPA. Our results emphasize the need for further research into UOGD-induced brine contamination.

Estimated 0.14% of UOG Wells in SWPA May be Leaking Methane. To report a defensible estimate of the number of UOGD wells potentially leaking CH_4 into shallow groundwater at the time of water sampling, we followed a four-step workflow (Figure S10). Because of the likelihood of natural CH_4 in SWPA groundwaters, we focused only on hotspots in Figure 3A. Next, areas where significant correlations on the UOG well heatmap overlapped with significant correlations on the COG well, coal mining, or anticline heatmaps (Figures S2 and S4) were attributed to anthropogenic and/or natural background sources (leaking COG wells, coal bed CH_4, and anticlinal folding) and excluded from further analysis. Subsequently, we examined whether any samples located in hotspots show signs of recent CH_4 migration based on our geochemical protocol. Of the 22 samples flagged by the geochemical protocol as potentially containing leaked CH_4, three samples were located within hotspots identified in Figure 3A (circled in Figure S5).

Finally, we investigated whether any aspects of well construction may have resulted in CH_4 leakage in these hotspots. We looked for wells drilled i) without intermediate casings at depths where they intersected with gas-bearing formations or ii) without coal casings in PADEP-designated coal mining areas. Intermediate or coal casings are designed to isolate boreholes from hydrocarbon-bearing formations at intermediate depths or workable coal seams, respectively. Based on well construction and drilling documents in the Pennsylvania Geological Survey's Exploration and Development Well Information Network database,[55] we identified five UOG wells in the hotspot circled in Figure 3A that lack coal or intermediate casings despite spatial overlap with DEP-identified coal mining areas. Although these five wells were allowed to be drilled without coal casings by the regulator despite intersection with coal mining areas, we infer that they may have leaked CH_4 to groundwater. This yields an estimated leakage rate of 0.14% out of 3610 UOG wells in our study area.

Brine Contamination Is More Frequent and Gas Migration Less Frequent in SWPA. We compared our results from SWPA to NEPA, a portion of the Appalachian Basin with little prior hydrocarbon development, to investigate how overlap between UOGD and COGD or coal mining influences the frequency of Cl or CH_4 contamination. While [Cl] increases nearby UOG wells in SWPA, analysis of 11,244 samples from Bradford County in NEPA indicates that [Cl] significantly decreases nearby UOG wells and does not significantly increase with well density (Supporting Information). Differences in hydrogeochemistry may explain this discrepancy between the two regions. Our data sets show significantly higher [Cl] at valley bottoms in NEPA but no significant difference in [Cl] across topographic positions in SWPA (Figure S8). We infer from this that drinking water wells are affected by natural brine migration or incompletely flushed connate water only in valleys in NEPA but in both valleys and ridgetops in SWPA. The most likely reason for this is the absolute depth to brine is shallower in SWPA than NEPA. Saline groundwater is encountered at shallower depths in valleys than uplands in NEPA, whereas the depth to saline groundwater is relatively uniform relative to the land surface across SWPA.[29] Because [Cl] is significantly higher in NEPA valley bottoms but only 9% of UOG wells are drilled in valleys (Figure S8), [Cl] in that area might be expected to increase further from UOG wells. Thus, any UOGD-driven increase in [Cl] in NEPA may be undetectable against the hydrogeologic signal of higher [Cl] in valley bottoms.[27,29,45] While topography may obscure the regional impacts of UOGD on [Cl], we can detect three hotspots where [Cl] increases with the density of or distance to UOGD in NEPA using the SWGT, although we cannot exclude the possibility of a naturally migrated brine source of Cl in these localities (Figures S11 and S12).

Despite evidence for regional-scale brine contamination in the region with intense legacy hydrocarbon extraction (SWPA), fewer UOGD wells may be leaking CH_4. In particular, only 5/3610 UOG wells (0.14%) in SWPA may be leaking CH_4, a factor of ~4 lower than a comparable estimate that 7/1385 UOG wells (0.51%) may leak CH_4 in Bradford County.[23] The difference in frequency between SWPA and NEPA is statistically significant (Fisher exact test, $p < 0.05$) and aligns with other lines of evidence suggesting that UOG wells in SWPA may be less prone to CH_4 migration compared to NEPA. Analyses of integrity of UOG wells in

https://doi.org/10.1021/acs.est.2c00001
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

NEPA show an 8.5-times greater risk of casing/cementing failures (often associated with gas migration) than wells in other parts of the state.[56] Positive determination letters (PDLs), issued when the PADEP determines an oil and gas well has impacted water supplies, also provide regulatory evidence for a higher frequency of CH_4 contamination in NEPA, as 54/67 PDLs issued in Bradford County and 0/8 PDLs issued in SWPA documented groundwater CH_4 contamination from UOGD.[57]

A complex interplay of factors including differences in data quantity/quality, regional hydrogeology and geologic deformation, well construction, and the presence and quantity of shallow gas above the target formation can also contribute to the detection of and propensity for CH_4 leakage from gas wells across hydrocarbon-producing regions. Many of these factors are consistent between SWPA and NEPA, but the hydrogeologic regimes differ and surface faulting is lacking in SWPA relative to NEPA. We argue that hydrogeologic differences between SWPA and NEPA are unlikely to drastically alter contamination frequencies (Supporting Information). Although faults and fractures can provide migration pathways for CH_4 to reach shallow groundwater,[23,34] other UOGD basins with more extensive near-surface faulting than SWPA (e.g., the Eagle Ford) show lower frequencies of CH_4 contamination than PA.[58] We thus argue that the dearth of faults is unlikely to be the primary mechanism governing gas migration frequency.

Gas-bearing formations overlying the Marcellus have been far more exploited in SWPA than in NEPA, likely resulting in greater depletion of shallow gas plays. Given that shallow gas in formations overlying the target formations of unconventional drilling is frequently the CH_4 source in gas migration incidents,[9,10] this prior extraction could have significant implications for gas migration frequency. Abandoned wells, which likely number in the tens of thousands in SWPA, may further vent gas from shallow reservoirs to the atmosphere.[39] As such, we hypothesize that the depletion of shallow gas overlying the Marcellus Shale in SWPA relative to NEPA may explain the reduced frequency of CH_4 migration incidents in SWPA vs NEPA.

We can compare NEPA and SWPA to the Denver-Julesburg basin in Colorado, where shallow gas also overlies target formations and UOGD overlaps with COGD and coal mining. Estimates of CH_4 migration frequency between SWPA and the Denver-Julesburg basin (testing both the lower bound estimate of 0.06% and upper bound estimate of 0.15% of wells leaking CH_4 from Sherwood et al.[35]) do not show a statistically significant difference at the 95% confidence level (Fisher exact test, $p > 0.05$), while the frequency of gas migration in NEPA is significantly higher than in SWPA or the Denver-Julesburg basin (Fisher exact test, $p < 0.05$). Based on this observation, we hypothesize that the prior drawdown of shallow gas via legacy hydrocarbon extraction may be the primary factor diminishing fugitive CH_4 contamination in SWPA. High volumes of shallow gas in NEPA may also partially explain why the gas migration frequency appears more problematic than in other gas-producing regions such as Arkansas, Texas, and North Dakota.[9,58−60]

Potential Health Risks of UOGD in Density Hotspots. CH_4 is nontoxic but can pose an explosion hazard above 10 mg/L.[2,3] Additionally, CH_4 migration can stimulate the growth of methanotrophic microorganisms[61,62] and subsequently produce reducing conditions in aquifers that solubilize iron, manganese, and toxic trace elements such as arsenic.[3] Thus, we calculated the highest expected increase in $[CH_4]$ associated with UOGD in hotspots. To assess this, we multiplied the highest density of UOG wells in a hotspot in Figure 2A (n_{well} = 7) by the increase in $[CH_4]$ per UOG well calculated via the ATS slope (i_{CH4} = 3.0 mg/L). This calculation (n_{well} * i_{CH4}) is consistent with up to a 21.0 mg/L increase in $[CH_4]$ attributable to UOGD. We conclude that, while UOGD does not systematically increase $[CH_4]$ across SWPA, $[CH_4]$ could exceed action limits for explosion hazards (10 mg/L) in hotspots where ≥4 wells are located within 1 km.

With a similar calculation for Cl (n_{well} * i_{Cl}) based on a calculated i_{Cl} = 3.6 mg/L per UOG well drilled and n_{well} = 12 in hotspots in Figure 2B, the highest expected increase in [Cl] resulting from UOGD is 43.2 mg/L in hotspots. This value is much lower than the U.S. Environmental Protection Agency (USEPA) secondary maximum contaminant level (MCL) of 250 mg/L for Cl. However, if flowback or produced water is the source of Cl contamination from UOGD, drinking water may also be impacted by species more threatening to human health than Cl. Many of these species are not typically analyzed in predrill data sets (e.g., thallium) or include species like arsenic that may only be detected by commercial analytical labs at concentrations harmful to human health. To investigate such species, we compiled the average chemistry of produced water from SWPA UOG wells in the U.S. Geological Survey Produced Water database[63] and calculated the expected increase in species concentrations per UOG well drilled in Figure 2B hotspots. We used the calculated ratio of [species]/[Cl] and the per-well increase in [Cl] in hotspots (i_{Cl}). Of the 44 species for which we have data, only one (thallium) yielded a predicted concentration exceeding the USEPA MCL (0.002 mg/L) in 3/5 hotspots (Table S2). These hotspots are circled in dark gray in Figure 2B. In the hotspot with the highest density of UOGD, predicted arsenic, beryllium, and cadmium concentrations exceed 75% of the EPA MCL and could pose health risks in more acutely impacted water supplies.

Our results point to a regional increase in groundwater [Cl] nearby UOG wells in SWPA, where the Marcellus Shale play overlaps with a long legacy of oil, gas, and coal extraction. However, rather than every UOG well producing higher [Cl], this regional impact is likely driven by hotspots where UOGD brines reach groundwater via leaks or spills, or, possibly, where subsurface features or well integrity allows leaking. If the Cl contamination is accompanied by contaminants such as thallium, this could be of concern. Notably, the ingestion of thallium above the USEPA MCL in drinking water has previously been associated with increased risk of low birth weights,[64] and low birth weights have been documented in populations in SWPA exposed to a high density of UOGD.[43]

These results point toward potential human exposure risk via water supplies contaminated by UOGD in limited areas. Our methods, which highlight localities in which water supplies may be impacted, are not a "gold standard" for exposure and must be followed by case studies. Nonetheless, our methods are important because the low frequency of UOGD contamination and complexity of hydrogeologic transport of contaminants into water supplies means that randomly collected water samples are unlikely to document exposure via drinking water. Methods such as ours should thus guide exposure studies.

We also identify infrequent CH_4 contamination in SWPA and emphasize that UOGD near coal mining necessitates

https://doi.org/10.1021/acs.est.2c00001
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

enhanced attention to prevent gas migration from coal seams. At the same time, the frequency of gas migration from UOG wells appears lower than the frequency associated with COG wells, as we identified a significant correlation across SWPA between [CH_4] and proximity to COG wells but no such correlation for UOG wells. Perhaps, the prior OGD in SWPA that removed gas explains today's lessened gas leakage during UOGD. We nonetheless emphasize that contamination from UOGD can translate into small regional impacts on groundwater chemistry that could expose human populations to low-level concentrations of harmful contaminants in highly localized areas.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.2c00001.

Extended study area background (Text S1), additional databases used (Text S2), additional spatial analysis details (Text S3), geochemical protocol to identify recently migrated methane (Text S4), additional details of NMF (Text S5), correlations between conventional wells and groundwater species (Text S6), assessing the relationship between topography, unconventional gas well locations, and [Cl] (Text S7), sliding window analysis of [Cl] vs. UOG well distance and density in Bradford County, PA (Text S8), testing the SWGT on Susquehanna County, PA groundwater methane data (Text S9), testing the NMF model using the Lautz et al.[44] synthetic data set (Text S10), groundwater flow in SWPA vs. NEPA (Text S11), gas geochemistry in a small subset of sample sites (Text S12), study area maps, additional sliding window heatmaps, locations of samples flagged by the geochemical protocol, histograms of NMF-derived mixing proportions, topographic index position class vs chloride concentrations and UOG well locations, schematic diagram of the workflow to flag potentially leaking UOG wells, methane isotopic compositions and gas chemistry data (Figures S1−S15), NMF-derived endmember sources, summary of additional brine species calculations, and correlations between [CH_4] or [Cl] and the distance to or density of COG wells (Tables S1−S4) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Samuel W. Shaheen − *Department of Geosciences, Pennsylvania State University, University Park, Pennsylvania 16802, United States;* orcid.org/0000-0002-0111-7361; Phone: (612)-229-0850; Email: samuelwshaheen@gmail.com

Authors

Tao Wen − *Department of Earth and Environmental Sciences, Syracuse University, Syracuse, New York 13244, United States;* orcid.org/0000-0002-6113-7532

Alison Herman − *Earth and Environmental Systems Institute, Pennsylvania State University, University Park, Pennsylvania 16802, United States*

Susan L. Brantley − *Department of Geosciences and Earth and Environmental Systems Institute, Pennsylvania State University, University Park, Pennsylvania 16802, United States;* orcid.org/0000-0002-9574-2693

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.2c00001

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the PADEP for data access and A. Shaughnessy for help developing NMF models. Additionally, we acknowledge assistance from D. Yoxtheimer, P. Hammond, T. Engelder, M. Lloyd, F. Baldassare, J. Hooker, R. Slingerland, and D. Catalano. This work was supported by National Science Foundation IIS award 16-39150 to S.L.B. and Zhenhui Li. AH was partially supported by the Institutes of Energy and the Environment and the College of Earth and Mineral Sciences at Penn State.

■ REFERENCES

(1) 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*, 8172−8176.

(2) Vidic, R. D.; Brantley, S. L.; Vandenbossche, J. M.; Yoxtheimer, D.; Abad, J. D. Impact of Shale Gas Development on Regional Water Quality. *Science* **2013**, *340*, No. 1235009.

(3) Woda, J.; Wen, T.; Oakley, D.; Yoxtheimer, D.; Engelder, T.; Castro, M. C.; Brantley, S. Detecting and Explaining Why Aquifers Occasionally Become Degraded near Hydraulically Fractured Shale Gas Wells. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 12349.

(4) Currie, J.; Greenstone, M.; Meckel, K. Hydraulic Fracturing and Infant Health: New Evidence from Pennsylvania. *Sci. Adv.* **2017**, *3*, No. e1603021.

(5) Hill, E. L. Shale Gas Development and Infant Health: Evidence from Pennsylvania. *J. Health Econ.* **2018**, *61*, 134−150.

(6) Deziel, N. C.; Brokovich, E.; Grotto, I.; Clark, C. J.; Barnett-Itzhaki, Z.; Broday, D.; Agay-Shay, K. Unconventional Oil and Gas Development and Health Outcomes: A Scoping Review of the Epidemiological Research. *Environ. Res.* **2020**, *182*, No. 109124.

(7) Jasechko, S.; Perrone, D. Hydraulic Fracturing near Domestic Groundwater Wells. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 13138−13143.

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

(9) Hammond, P. A.; Wen, T.; Brantley, S. L.; Engelder, T. Gas Well Integrity and Methane Migration: Evaluation of Published Evidence during Shale-Gas Development in the USA. *Hydrogeol. J.* **2020**, *28*, 1481−1502.

(10) Dusseault, M.; Jackson, R. Seepage Pathway Assessment for Natural Gas to Shallow Groundwater during Well Stimulation, in Production, and after Abandonment. *Environ. Geosci.* **2014**, *21*, 107−126.

(11) Lackey, G.; Rajaram, H.; Bolander, J.; Sherwood, O. A.; Ryan, J. N.; Shih, C. Y.; Bromhal, G. S.; Dilmore, R. M. Public Data from Three US States Provide New Insights into Well Integrity. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, No. e2013894118.

(12) Martini, A. M.; Walter, L. M.; Ku, T. C. W.; Budai, J. M.; McIntosh, J. C.; Schoell, M. Microbial Production and Modification of Gases in Sedimentary Basins: A Geochemical Case Study from a Devonian Shale Gas Play, Michigan Basin. *Am. Assoc. Pet. Geol. Bull.* **2003**, *87*, 1355−1375.

(13) 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*, 4106−4112.
(14) McMahon, P. B.; Barlow, J. R. B.; Engle, M. A.; Belitz, K.; Ging, P. B.; Hunt, A. G.; Jurgens, B. C.; Kharaka, Y. K.; Tollett, R. W.; Kresse, T. M. Methane and Benzene in Drinking-Water Wells Overlying the Eagle Ford, Fayetteville, and Haynesville Shale Hydrocarbon Production Areas. *Environ. Sci. Technol.* **2017**, *51*, 6727−6734.
(15) 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.
(16) 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*, 12012−12021.
(17) Kreuzer, R. L.; Darrah, T. H.; Grove, B. S.; Moore, M. T.; Warner, N. R.; Eymold, W. K.; Whyte, C. J.; Mitra, G.; Jackson, R. B.; Vengosh, A.; Poreda, R. J. Structural and Hydrogeological Controls on Hydrocarbon and Brine Migration into Drinking Water Aquifers in Southern New York. *Groundwater* **2018**, *56*, 225−244.
(18) 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*, 11250−11255.
(19) Barth-Naftilan, E.; Sohng, J.; Saiers, J. E. Methane in Groundwater before, during, and after Hydraulic Fracturing of the Marcellus Shale. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 6970−6975.
(20) 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*, 333−349.
(21) 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.
(22) 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. (195) (December 2016) 23−30]. *J. Contam. Hydrol.* **2017**, *207*, 50−51.
(23) Wen, T.; Niu, X.; Gonzales, M.; Zheng, G.; Li, Z.; Brantley, S. L. Big Groundwater Data Sets Reveal Possible Rare Contamination Amid Otherwise Improved Water Quality for Some Analytes in a Region of Marcellus Shale Development. *Environ. Sci. Technol.* **2018**, *52*, 7149−7159.
(24) Elliott, E. G.; Ettinger, A. S.; Leaderer, B. P.; Bracken, M. B.; Deziel, N. C. A Systematic Evaluation of Chemicals in Hydraulic-Fracturing Fluids and Wastewater for Reproductive and Developmental Toxicity. *J. Expo. Sci. Environ. Epidemiol.* **2017**, *27*, 90−99.
(25) Bonetti, P.; Leuz, C.; Michelon, G. Large-Sample Evidence on the Impact of Unconventional Oil and Gas Development on Surface Waters. *Science* **2021**, *373*, 896−902.
(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*, 865−885.
(27) Warner, N.; Jackson, R.; Darrah, T.; Osborn, S.; 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*, 11961−11966.
(28) Long, D. T.; Wilson, T. P.; Takacs, M. J.; Rezabek, D. H. Stable-Isotope Geochemistry of Saline near-Surface Ground Water: East-Central Michigan Basin. *Geol. Soc. Am. Bull.* **1988**, *100*, 1568−1577.
(29) Siegel, D. I.; Smith, B.; Perry, E.; Bothun, R.; Hollingsworth, M. Pre-Drilling Water-Quality Data of Groundwater Prior to Shale Gas Drilling in the Appalachian Basin: Analysis of the Chesapeake Energy Corporation Dataset. *Appl. Geochem.* **2015**, *63*, 37−57.
(30) Lauer, N. E.; Harkness, J. S.; Vengosh, A. Brine Spills Associated with Unconventional Oil Development in North Dakota. *Environ. Sci. Technol.* **2016**, *50*, 5389−5397.
(31) Cozzarelli, I. M.; Skalak, K. J.; Kent, D. B.; Engle, M. A.; Benthem, A.; Mumford, A. C.; Haase, K.; Farag, A.; Harper, D.; Nagel, S. C.; Iwanowicz, L. R.; Orem, W. H.; Akob, D. M.; Jaeschke, J. B.; Galloway, J.; Kohler, M.; Stoliker, D. L.; Jolly, G. D. Environmental Signatures and Effects of an Oil and Gas Wastewater Spill in the Williston Basin, North Dakota. *Sci. Total Environ.* **2017**, *579*, 1781−1793.
(32) Agarwal, A.; Wen, T.; Chen, A.; Zhang, A. Y.; Niu, X.; Zhan, X.; Xue, L.; Brantley, S. L. Assessing Contamination of Stream Networks near Shale Gas Development Using a New Geospatial Tool. *Environ. Sci. Technol.* **2020**, *54*, 8632−8639.
(33) Wilkin, R. T.; Lee, T. R.; Ruybal, C. J.; Rectenwald, D. J. *Retrospective Case Study in Southwestern Pennsylvania: Study of the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources*; U.S. Environmental Protection Agency, 2015.
(34) 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*, 6325−6330.
(35) Sherwood, O. A.; Rogers, J. D.; Lackey, G.; Burke, T. L.; Osborn, S. G.; Ryan, J. N. Groundwater Methane in Relation to Oil and Gas Development and Shallow Coal Seams in the Denver-Julesburg Basin of Colorado. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 8391−8396.
(36) Kelly, W. R.; Matisoff, G.; Fisher, J. B. The Effects of a Gas Well Blow out on Groundwater Chemistry. *Environ. Geol. Water Sci* **1985**, *7*, 205−213.
(37) LeDoux, S. T. M.; Szynkiewicz, A.; Faiia, A. M.; Mayes, M. A.; McKinney, M. L.; Dean, W. G. Chemical and Isotope Compositions of Shallow Groundwater in Areas Impacted by Hydraulic Fracturing and Surface Mining in the Central Appalachian Basin, Eastern United States. *Appl. Geochem.* **2016**, *71*, 73−85.
(38) Bachu, S. Analysis of Gas Leakage Occurrence along Wells in Alberta, Canada, from a GHG Perspective − Gas Migration Outside Well Casing. *Int. J. Greenh. Gas Control* **2017**, *61*, 146−154.
(39) Kang, M.; Christian, S.; Celia, M. A.; Mauzerall, D. L.; Bill, M.; Miller, A. R.; Chen, Y.; Conrad, M. E.; Darrah, T. H.; Jackson, R. B. Identification and Characterization of High Methane-Emitting Abandoned Oil and Gas Wells. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 13636−13641.
(40) Lacombe, S.; Sudicky, E. A.; Frape, S. K.; Unger, A. J. A. Influence of Leaky Boreholes on Cross-Formational Groundwater Flow and Contaminant Transport. *Water Resour. Res.* **1995**, *31*, 1871−1882.
(41) Woda, J.; Wen, T.; Lemon, J.; Marcon, V.; Keeports, C. M.; Zelt, F.; Steffy, L. Y.; Brantley, S. L. Methane Concentrations in Streams Reveal Gas Leak Discharges in Regions of Oil, Gas, and Coal Development. *Sci. Total Environ.* **2020**, *737*, No. 140105.
(42) Townsend-Small, A.; Hoschouer, J. Direct Measurements from Shut-in and Other Abandoned Wells in the Permian Basin of Texas Indicate Some Wells Are a Major Source of Methane Emissions and Produced Water. *Environ. Res. Lett.* **2021**, *16*, No. 054081.
(43) Stacy, S. L.; Brink, L. L.; Larkin, J. C.; Sadovsky, Y.; Goldstein, B. D.; Pitt, B. R.; Talbott, E. O. Perinatal Outcomes and Unconventional Natural Gas Operations in Southwest Pennsylvania. *PLoS One* **2015**, *10*, No. e0126425.
(44) Lautz, L. K.; Hoke, G. D.; Lu, Z.; Siegel, D. I.; Christian, K.; Kessler, J. D.; Teale, N. G. Using Discriminant Analysis to Determine Sources of Salinity in Shallow Groundwater Prior to Hydraulic Fracturing. *Environ. Sci. Technol.* **2014**, *48*, 9061−9069.

https://doi.org/10.1021/acs.est.2c00001
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

(45) Llewellyn, G. T. Evidence and Mechanisms for Appalachian Basin Brine Migration into Shallow Aquifers in NE Pennsylvania, USA. *Hydrogeol. J.* **2014**, *22*, 1055−1066.
(46) Alexandrov, B. S.; Vesselinov, V. V. Blind Source Separation for Groundwater Pressure Analysis Based on Nonnegative Matrix Factorization. *Water Resour. Res.* **2014**, *50*, 7332−7347.
(47) Vesselinov, V. V.; Alexandrov, B. S.; O'Malley, D. Contaminant Source Identification Using Semi-Supervised Machine Learning. *J. Contam. Hydrol.* **2018**, *212*, 134−142.
(48) Shaughnessy, A. R.; Gu, X.; Wen, T.; Brantley, S. L. Machine Learning Deciphers CO2 Sequestration and Subsurface Flowpaths from Stream Chemistry. *Hydrol. Earth Syst. Sci.* **2021**, *25*, 3397−3409.
(49) Wen, T.; Woda, J.; Marcon, V.; Niu, X.; Li, Z.; Brantley, S. L. Exploring How to Use Groundwater Chemistry to Identify Migration of Methane near Shale Gas Wells in the Appalachian Basin. *Environ. Sci. Technol.* **2019**, *53*, 9317−9327.
(50) Titler, R. V.; Curry, P. *Chemical Analysis of Major Constituents and Trace Contaminants of Rock Salt*; Pennsylvania Department of Environmental Protection, Bureau of Water Standards and Facility Regulation, 2011.
(51) Patterson, L. A.; Konschnik, K. E.; Wiseman, H.; Fargione, J.; Maloney, K. O.; Kiesecker, J.; Nicot, J.-P.; Baruch-Mordo, S.; Entrekin, S.; Trainor, A.; Saiers, J. E. Unconventional Oil and Gas Spills: Risks, Mitigation Priorities, and State Reporting Requirements. *Environ. Sci. Technol.* **2017**, *51*, 2563−2573.
(52) Flewelling, S. A.; Sharma, M. Constraints on Upward Migration of Hydraulic Fracturing Fluid and Brine. *Groundwater* **2014**, *52*, 9−19.
(53) Montague, J. A.; Pinder, G. F. Potential of Hydraulically Induced Fractures to Communicate with Existing Wellbores. *Water Resour. Res.* **2015**, *51*, 8303−8315.
(54) Taherdangkoo, R.; Tatomir, A.; Anighoro, T.; Sauter, M. Modeling Fate and Transport of Hydraulic Fracturing Fluid in the Presence of Abandoned Wells. *J. Contam. Hydrol.* **2019**, *221*, 58−68.
(55) Exploration and Development Well Information Network (EDWIN) https://edwin.dcnr.pa.gov/ (accessed October 8, 2021).
(56) Ingraffea, A. R.; Wells, M. T.; Santoro, R. L.; Shonkoff, S. B. C. Assessment and Risk Analysis of Casing and Cement Impairment in Oil and Gas Wells in Pennsylvania, 2000-2012. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 10955−10960.
(57) Water Supply Determination Letters. Pennsylvania Department of Environmental Protection. https://files.dep.state.pa.us/OilGas/BOGM/BOGMPortalFiles/OilGasReports/Determination_Letters/Regional_Determination_Letters.pdf (accessed October 8, 2021).
(58) Wen, T.; Liu, M.; Woda, J.; Zheng, G.; Brantley, S. L. Detecting Anomalous Methane in Groundwater within Hydrocarbon Production Areas across the United States. *Water Res.* **2021**, *200*, No. 117236.
(59) Warner, N. R.; Kresse, T. M.; Hays, P. D.; Down, A.; Karr, J. D.; Jackson, R. B.; Vengosh, A. Geochemical and Isotopic Variations in Shallow Groundwater in Areas of the Fayetteville Shale Development, North-Central Arkansas. *Appl. Geochem.* **2013**, *35*, 207−220.
(60) McMahon, P. B.; Caldwell, R. R.; Galloway, J. M.; Valder, J. F.; Hunt, A. G. Quality and Age of Shallow Groundwater in the Bakken Formation Production Area, Williston Basin, Montana and North Dakota. *Groundwater* **2015**, *53*, 81−94.
(61) Cahill, A. G.; Steelman, C. M.; Forde, O.; Kuloyo, O.; Emil Ruff, S.; Mayer, B.; Ulrich Mayer, K.; Strous, M.; Cathryn Ryan, M.; Cherry, J. A.; Parker, B. L. Mobility and Persistence of Methane in Groundwater in a Controlled-Release Field Experiment. *Nat. Geosci.* **2017**, *10*, 289−294.
(62) Stanish, L. F.; Sherwood, O. A.; Lackey, G.; Osborn, S.; Robertson, C. E.; Harris, J. K.; Pace, N.; Ryan, J. N. Microbial and Biogeochemical Indicators of Methane in Groundwater Aquifers of the Denver Basin, Colorado. *Environ. Sci. Technol.* **2021**, *55*, 292−303.
(63) Engle, M. A.; Saraswathula, V.; Thorsden, J. J.; Morissey, E. A.; Gans, K. D.; Blondes, M. S.; Kharaka, Y. K.; Rowan, E. L.; Reidy, M. E. U. S. Geological Survey National Produced Waters Geochemical Database v2.3 [Data Set]. U.S. Geological Survey Data Release, 2019, DOI: 10.5066/F7J964W8.
(64) Nuvolone, D.; Petri, D.; Aprea, M. C.; Bertelloni, S.; Voller, F.; Aragona, I. Thallium Contamination of Drinking Water: Health Implications in a Residential Cohort Study in Tuscany (Italy). *Int. J. Environ. Res. Public Health* **2021**, *18*, 4058.

https://doi.org/10.1021/acs.est.2c00001
Environ. Sci. Technol. XXXX, XXX, XXX−XXX