Environmental Science Processes & Impacts

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d6em00142d

Targeted sampling reveals localized brine signals without evidence of regional water quality impacts from unconventional oil and gas development

Jianfeng Su,*[a] Lara Williams,[b] Allison Fenske,[c] Samuel W. Shaheen,[bd] Susan L. Brantley,[be] Brandon Forsythe,[b] Nathaniel R. Warner,[c] Jennifer Baka[ef] and Tao Wen*[a]

Received 23rd February 2026
Accepted 31st May 2026

DOI: 10.1039/d6em00142d

rsc.li/espi

Unconventional oil and gas development (UOGD) in the Marcellus Shale of southwestern Pennsylvania has raised persistent concerns regarding groundwater quality. Previous work identified potential regional associations between UOGD activity and elevated salinity in heavily drilled counties based on samples collected between 2008 and 2018. Here, we collected 97 water samples from private wells, springs, and streams in 2024 to evaluate whether such associations persist and to assess whether impacts, if present, are regional or highly localized. Sampling targeted previously identified chloride hotspots, decommissioned wastewater impoundments, and recent spills, with comparisons to control locations. Major ions, trace metals, dissolved gases, isotopes, and organic compounds were analyzed and compared with drinking water standards, historical datasets, and proximity to UOGD features while accounting for geologic, hydrogeologic, and topographic controls. Most samples met drinking-water standards, and statistical comparisons revealed no robust regional differences between treatment and control samples after Benjamini–Hochberg correction. Fixed-effects models showed no evidence of widespread groundwater degradation attributable to UOGD. Among spill sites sampled, only one groundwater sample exhibited elevated salinity and brine-associated constituents. Geochemical, isotopic, and topographic evidence suggests this anomaly reflects localized influence from produced water or natural Appalachian Basin brine migrating along deeper flow paths rather than surface transport. Our results indicate that contamination patterns are best explained by highly localized rather than region-wide problems, consistent with rare and localized detectable impacts within the targeted sampling framework used here and highlighting challenges in defining appropriate controls given hydrogeologic heterogeneity. This work emphasizes combining regional screening with targeted, hydrogeology-informed monitoring to assess UOGD-related groundwater contamination.

Environmental significance

Southwestern Pennsylvania hosts a high density of unconventional oil and gas (UOG) wells in the United States, yet the extent of associated impacts under current regulatory conditions remains poorly constrained. This study presents a hydrogeochemical dataset of 97 samples collected in 2024 from Washington and Greene Counties, targeting previously identified chloride hotspots, decommissioned impoundments, and recent spill sites. We show that widespread brine-related contamination attributable to unconventional oil and gas development (UOGD) is not detectable in this targeted dataset, and that persistent impacts are rare, highly localized, and strongly governed by hydrogeologic context. These findings demonstrate that proximity to UOGD features alone is an insufficient predictor of contamination risk. We propose a tiered monitoring framework that combines regional screening with hydrogeology-informed field investigations, offering a transferable model for protecting drinking-water resources in active shale-development regions.

1 Introduction

Over the past two decades, unconventional oil and gas development (UOGD) has transformed global and U.S. energy production. Advances in hydraulic fracturing and horizontal drilling have unlocked hydrocarbon reserves from shale and tight-sand formations, notably in the Marcellus Shale in Pennsylvania (USA), enhancing national energy security while also

---

[a] *Department of Earth and Environmental Sciences, Syracuse University, Syracuse, NY, USA. E-mail: jsu124@syr.edu; twen08@syr.edu*
[b] *Department of Geosciences, Pennsylvania State University, University Park, PA, USA*
[c] *Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, PA, USA*
[d] *Department of Earth & Environmental Sciences, University of Minnesota, Minneapolis, MN, USA*
[e] *Earth and Environmental Systems Institute, Pennsylvania State University, USA*
[f] *Department of Geography, Pennsylvania State University, University Park, PA, USA*

raising environmental and public health concerns. In particular, the potential for groundwater contamination associated with subsurface well integrity and surface handling of produced and flowback fluids has attracted scientific and community concern.[1,2]

Early studies in Pennsylvania documented that shallow aquifers in the Appalachian Basin can contain naturally-occurring methane due to both microbial methanogenesis and natural upwelling of thermogenic methane from deeper formations and, occasionally, thermogenic gas linked to well integrity failures.[2–8] Although nontoxic, methane can pose explosion and asphyxiation hazards in confined spaces once concentration exceeds 10 mg L^{-1} in water.[2,5] Beyond methane, which primarily reflects gas migration and well integrity issues, concerns related to unconventional oil and gas development also include the potential release of liquid wastes associated with hydraulic fracturing operations. In particular, flowback water—produced when injected fluid returns to the surface after hydraulic fracturing—often contains highly concentrated brine salts, reflecting a mixture of the original fracking fluid and naturally occurring formation brines.[9] These wastewaters are usually enriched with naturally occurring radioactive materials and heavy metals, many of which are toxic and can cause adverse health effects even at very low levels.[10] Some species uncommon in shallow groundwater (*e.g.*, barium (Ba), strontium (Sr), bromide (Br)) can act as “fingerprints” for mixing of brines with fresher waters. These fingerprint species are sometimes naturally present in shallow groundwater, attributed to natural vertical migration of basin brines through interformation pathways[7] or infrequent flushing of connate salt water by meteoric recharge.[11] However, leakage from impoundments, where the flowback water is held for unconventional oil and gas (UOG) operation, and surface spills that may occur during temporary storage, transportation and waste handling processes, pose a significant contamination risk.[12]

Using large publicly available groundwater datasets from the Marcellus Shale region, Shaheen *et al.* (2022, 2024)[13,14] demonstrated that UOGD is associated with small but statistically significant regional increases in salt-related groundwater constituents, including chloride (Cl), Ba, and Sr, particularly in southwestern Pennsylvania (SWPA) where UOG activity overlaps with dense legacy coal, oil, and gas extraction. These studies further showed that any regional associations between groundwater chemistry and UOGD are likely driven by subregions with documented spill-related violations and wastewater impoundments, and that such relationships persist after accounting for other natural and anthropogenic sources of salinity. Collectively, these findings were interpreted to suggest that wastewater management issues, rather than routine UOG operations, may contribute to subtle degradation of groundwater in which a small number of localized incidents can produce a regionally detectable effect in SWPA. Shaheen *et al.* (2022)[13] used a geospatial tool to identify “hotspots,” *i.e.*, areas in which concentrations of Cl significantly increased with UOG well density. However, most samples underlying these interpretations were collected early in the history of UOGD in SWPA (*i.e.*, pre-2014), leaving it unclear whether previously observed contamination persists under the current regulatory and operational framework that has evolved since early UOGD.

Here, we evaluate the regional relationships reported by Shaheen *et al.* (2022, 2024)[13,14] using an independent groundwater dataset collected in 2024. Although our sampling targeted fewer locations than the earlier regional studies, it was designed to test whether previously reported salt-related groundwater signals remain detectable under current regulatory and operational conditions. Because the earlier regional associations were hypothesized to arise from localized wastewater-management issues, particularly leakage from impoundments or produced-water spills, we targeted sites near previously identified chloride hotspots, decommissioned wastewater impoundments and documented recent spills. Given the lack of recent spills in the area that we were able to sample, we also targeted brownfield sites (hereafter referred to as “spills” for simplicity), where spills had occurred and been remediated. By applying multivariate statistical analyses and explicitly controlling for geologic, hydrogeologic, topographic, and anthropogenic factors, we assess whether previously reported UOGD-associated groundwater signals are persistent, localized, or no longer detectable.

Our study focuses on Washington and Greene Counties in SWPA, a region with the highest density of UOG wells in the Appalachian Basin.[15] Washington and Greene Counties contain thousands of horizontal gas wells (UOG wells) and a century of legacy conventional drilling and coal mining, making it a key location for assessing potential effects on groundwater. We organized the study around four related hypotheses and corresponding analyses. First, we tested whether regional brine-related contamination signals identified during the early period of UOGD remain detectable in the 2024 dataset by comparing treatment and control samples, and by evaluating results relative to historical groundwater data. Second, we tested whether previously identified chloride hotspots persist as zones of elevated salinity by characterizing groundwater chemistry within hotspot areas and comparing these samples with controls. Third, we tested whether wastewater impoundments and spills produce localized and potentially persistent groundwater contamination by sampling near decommissioned impoundments, documented spills, and brownfield sites and by assessing spatial, geologic, hydrogeologic, and topographic controls on observed chemistry. Fourth, we tested whether any detected impacts correspond to potential human-health concerns by comparing measured concentrations with established or proposed drinking-water thresholds. Together, these analyses provide a framework for determining whether UOGD-related groundwater effects in SWPA are regionally persistent, locally restricted, or not detectable under the current conditions.

2 Methods

2.1 Study area and water sampling site selection

Water sampling sites are hereafter referred to as “samples,” to distinguish them from UOG sites/features, which are referred to as “sites.” Four types of sites were targeted for water sampling and samples were identified by acronym (Fig. 1): (1) HS samples

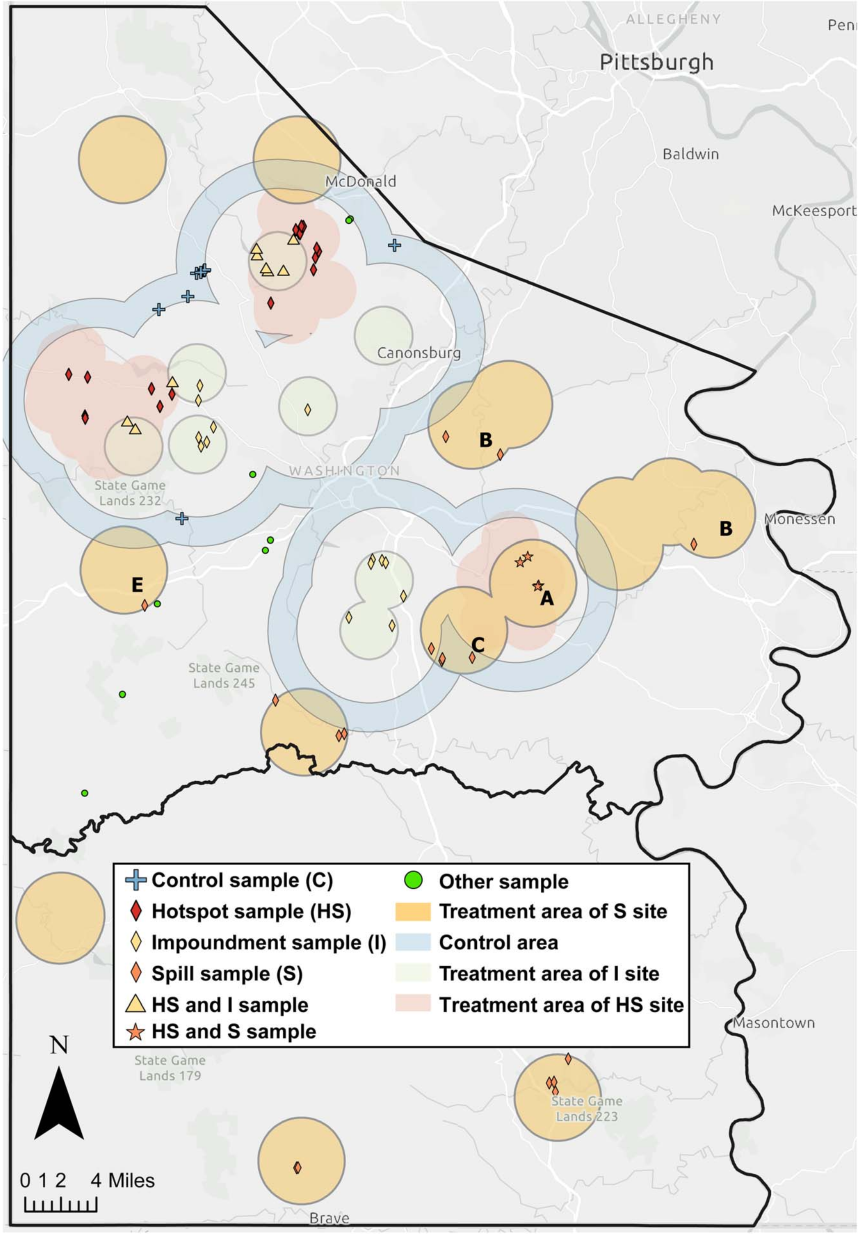

Fig. 1 Map of the study areas in Washington and Greene Counties in SWPA. Locations of collected water samples and sites of impoundments, hotspots, recent spills and brownfields, and control zones are color-coded. Five brownfield sites (all are oil and gas well pads) that are investigated for local spill impacts are marked with letters A–E.

located within 2 km of UOG wells in chloride hotspots as identified previously;[13] (2) I samples located within 2 km of wastewater impoundments decommissioned or restricted by PADEP because of the potential for leakage of UOGD wastewaters;[16] (3) S samples located within 3 km of documented UOG-related spills or brownfields occurring between 2021 and 2024; and (4) C samples chosen as control sites because they are located 5–7 km from any of these UOGD features. An extensive outreach and recruitment campaign was conducted to collect water samples from homes with private wells in each of these categories and the participation was voluntary (see Text S6 for campaign details). Given the rural nature of the study area, the low density of domestic water wells, and the difficulty in finding willing volunteers for water sampling, we also collected 27 supplemental community-engagement samples that did not fall within these four categories (categorized as "Others"; Table 1). These samples were collected from homeowners who expressed interest in the study and were included to support community engagement, return useful water-quality information to participating residents, and provide additional regional groundwater chemistry context. They were therefore not intended to represent a fifth exposure category and were not used to define treatment or control groups. Nevertheless, these samples provide useful contextual information because 7 of the 27 were located within a hydrologic vulnerability zone previously

Table 1 Water sample breakdown by category (I = impoundment; HS = hotspot; S = spill; C = control; I + HS and S + HS = samples located in both I and HS, and S and HS treatment areas, respectively)

Site type	I	HS	I + HS[a]	S	S + HS	C	Other	Total[a]
# wells	20	23	8	16	2	7	23	79
# springs	4	7	2	1	1	1	4	14
# streams	0	1	0	4	1	0	0	4
Total	24	31	10	21	4	8	27	97

[a] Some water samples fell within multiple areas of interest and must not be double counted.

identified as potentially susceptible to contamination from UOGD-related activities by Soriano *et al.* (2022).[17]

In total, 97 samples were collected across 7 field campaigns between May and November 2024, including both groundwater (79 wells and 14 springs) and stream water ($n = 4$). Two groundwater samples strongly influenced by acid mine drainage (AMD) and one rain-dominated sample were excluded from subsequent data analyses, yielding 90 groundwater samples for analysis (Fig. 1 and Table 1). Stream samples were collected to provide local surface water context but since we focus on groundwater in this study, they were also excluded from the analyses.

Water sampling sites were chosen based on locations of hotspots and impoundments as previously reported.[13–16] For example, all UOG well locations within the hotspots as previously identified were used to define an area where the outer perimeter was 2 km from every well. This buffer radius (2 km) was selected for the 31 HS samples based on previous modeling and field-based studies showing putative groundwater impairment caused by UOGD within this radius.[17,18] Like the hotspot target areas, the 24 I samples were collected within zones around impoundments that were also defined with a 2 km buffer radius.

To sample the effect of spills, we targeted spill or leakage events that occurred within the past three years. A 3-year window was selected because impacts from spills were observed in previous work in Pennsylvania as well as in other shale plays within this temporal range.[14,19,20] Initially, we sought to collect groundwater samples within 2 km of these locations, consistent with the buffer used for hotspot and impoundment sampling. However, because domestic wells were sparse near several spill locations and homeowner participation was voluntary, we expanded the spill-sampling buffer from 2 km to 3 km to obtain sufficient samples for exploratory comparison near spill-associated sites. This expansion was also consistent with evidence that at least one documented UOGD-related contamination incident in PA involved migration more than 2 km along fractures.[18] Even within a 3 km radius of the three spills we identified in Greene County that occurred between 2021 and 2024, we collected only six samples.

To increase the potential sampling possibilities near spills, we identified twelve additional sites in Washington County using data from the PADEP Land Recycling Program. Information from that Program allowed us to find brownfield sites associated with major UOG companies. Pennsylvania brownfield sites are former industrial or commercial properties where reuse is complicated by potential contamination from hazardous substances but that are eligible for cleanup and redevelopment. We were able to collect fifteen water samples from 6 of the 12 brownfields. Two of the six brownfields were associated with spills that occurred in 2015 and 2020, respectively. Given the very small number of spills and domestic wells located near spills, samples (Sample_082 and Sample_083) collected near these sites outside the 3-year window were retained in the analysis to evaluate whether longer-term contamination signatures persist beyond the 3-year window, providing a complementary perspective on attenuation timescales. Because these older sites differ in timing from the recent spill sites, they were interpreted as supplemental cases for evaluating possible longer-term persistence rather than as direct equivalents to recent releases. We found no information about exact spillage locations for the 3 spill sites and for the 12 brownfield sites. Therefore, centroids of the wellpads were used to define the areas targeted for S water sampling.

Control samples were selected 5–7 km from the targeted UOGD features and, where possible, from broadly comparable hydrogeologic, topographic, and land-use settings. This distance range was chosen to minimize the likelihood of direct influence from the targeted UOGD features while maintaining sufficient geographic proximity to treatment areas for comparative assessment.

2.2 Sample collection and analytical procedures

At each water sampling site, we collected approximately 20 liters of water following best practice sampling protocols. Field parameters—including pH, temperature, dissolved oxygen (DO), specific conductance (SC), alkalinity, and methane in air—were measured *in situ* using calibrated multiparameter meters. Samples were filtered (0.45 μm) and preserved according to analyte requirements. Cation and Sr isotope samples were acidified with HNO_3, and all samples were kept at 4 °C until analysis. The list of analytes and full methodology for each is described in the SI. As the goal of this study was to compare concentration levels to established or proposed health standards, we did not include analytes that may be elevated near UOGD contamination but do not have established or proposed health standards, such as boron. BTEX compounds were measured using headspace gas chromatography-mass spectrometry (GC-MS). Some samples were not analyzed for BTEX for Trip 1 because of sample bottle breakage and for Trips 6 and 7 due to budgetary constraints.

Quality assurance and quality control included field blanks, trip blanks, duplicate samples, and calibration verification every 10 samples. Analytical precision for major and trace elements was within ±5% (refer to SI for test procedures and details of chemical analyses).

Water chemistry data were compared to established standards including the maximum contaminant level (MCL)[10] and secondary contaminant level (SMCL).[21] Where EPA has not established an MCL/SMCL, alternative threshold levels[22,23] were used for comparison (Table S1).

2.3 Statistical and spatial analyses

Nonparametric rank tests (Wilcoxon–Mann–Whitney [WMW] and Brunner–Munzel [BM]) were used to compare analyte concentrations among treatment (HS, I, S) and control (C) samples using one-sided tests. The WMW test performs best when sample variances are similar. The BM test is preferred when variances differ, as it is robust to heteroscedasticity. However, with many tied values—such as when most measurements fall below detection limits (*e.g.*, Aluminum (Al), Beryllium (Be), and Cadmium (Cd))—the BM test may fail, and the WMW test is more reliable. Both tests are well suited for skewed distributions, small samples, and datasets with many values below detection limits.[24] Data below detection limits were replaced with the detection limit for conservative inclusion in rank-based analyses.

Correlations between analyte concentrations and distance to the nearest UOGD feature were evaluated using both linear (ordinary least squares [OLS]) and rank-based (Spearman, Kendall's tau, and Akritas–Theil–Sen [ATS]) regression models for SC and dissolved concentrations of chloride (Cl), sodium (Na), sulfate (SO_4), barium (Ba), strontium (Sr), bromide (Br), iron (Fe), and methane (CH_4). SC, Cl, and Na are indicators of groundwater salinity and serve as proxies for assessing salt impacts, a major component of UOGD wastewaters. Fe and SO_4 were chosen because they are redox-sensitive and have sometimes been used to predict contamination related to UOGD.[8,24] SC, Cl, Ba, Sr, Br, and CH_4 are typically enriched in UOG produced waters, and have been used to identify contamination from deep formation brine, methane leakage, and putative UOG-related incidents.[13,24]

To assess the vulnerability of each sample to contamination that is transported *via* the surface or shallow pathway, watersheds of collected samples were delineated from 10 m × 10 m DEM raster.[25] In ArcGIS Pro, cells with values ≥100 (≥ 0.01 km^2) were used to define stream networks, snap samples within 200 m, and delineate capture zones using the "Watershed" tool. We also examined how groundwater chemistry relates to topography: the topographic position index (TPI)[26] was used to categorize locations into six terrain classes—valley, lower slope, flat slope, middle slope, upper slope, and ridge—which were grouped into "valley" (first three) and "ridge" (last three).

A simplified fixed-effects regression was used to control for potential confounding factors, including proximity to coal mines (<1 km) and conventional oil and gas wells (COGD),[15,27] highways,[28] topographic position (valley *vs.* ridge), and a sampling-season indicator (November *vs.* May–October). These factors were tested because of their reported effects on groundwater chemistry in Pennsylvania.[29,30] In particular, coal mining can increase SO_4 (among other analytes), COGD can affect many of the same contaminants associated with UOGD, and highways in PA are associated with elevated Na and Cl because of road salting. Spatial distances (horizontal planar distances) were calculated using ArcGIS Pro. Statistical significance was assessed at $p = 0.05$. The final equation for the fixed-effects model is:

$$\log C = \beta_1 \mathrm{UOG/I/S} + \mathrm{COG\ 1\ km} + \mathrm{coal\ mining\ 1\ km} + \mathrm{highway\ 1\ km} + \mathrm{TPI} + \mathrm{season} + \varepsilon \quad (1)$$

$\log C$ is the log concentration of brine-related species, UOG/I/S represents the geodesic distance metrics for all water samples, β_1 is the regression coefficient of the distance variables, and ε is the error term.

The 'other' samples were excluded from the primary categorical treatment-control tests because they were not designed to represent either treatment or control groups. They were retained in descriptive statistics and fixed-effects regression models when they met the relevant sample type criteria, because those models used continuous distance metrics rather than categorical treatment assignment.

2.4 Detailed examination of spills

For five of the six spill locations where we collected water samples (henceforth referred to as sites A, B, C, D, and E), we were able to find PADEP incident reports[31] describing spill events and remediation actions *via* file review at the PADEP Southwest Regional Office (Table 2). For four of these five spills, we compared the newly collected water data from this study (hereafter referred to as "2024 SWPA" dataset) to samples from the Shale Network (SN) dataset[32] collected within 1 km during previous years (Fig. 1). This dataset consists of groundwater analyses spanning the region of development of the Marcellus Shale in Pennsylvania, with more than 5000 groundwater samples collected in SWPA primarily before 2017. This is also the dataset used in the previous work that identified hotspots of potential UOGD contamination.[13,14] At each of the five spills, we compared the concentrations of brine-related analytes from the 2024 SWPA dataset with those from the SN dataset for sampling sites within 1 km of our 2024 sampling locations. For one of the spills (*i.e.*, spill A), we also compared 2024 data to water chemistry for samples collected by the PADEP.

3 Results

3.1 General groundwater chemistry and isotopic signatures

Field measurements indicated slightly acidic to slightly basic groundwater (pH 6.4–8.9) across all samples with most SC values ranging from 100 to 2400 μS cm^{-1} (median 380 μS cm^{-1}). Dissolved oxygen was typically <3 mg L^{-1}, consistent with mildly reducing conditions.

Most groundwater samples met EPA drinking-water standards (Tables S1–S6). Among the 90 groundwater samples analyzed, exceedances of MCLs, SMCLs, or other screening thresholds were observed primarily for Cl, Fe, Mn, TDS, and Li. Five of 90 2024 SWPA samples (5.6%) exceeded the EPA SMCL of 250 mg L^{-1} for chloride: two categorized as HS (286 mg L^{-1} and 568 mg L^{-1}; Sample_066 and Sample_077), one as I (307 mg L^{-1}; Sample_026), one as S (697 mg L^{-1}; Sample_082), and one as 'other' (584 mg L^{-1}; Sample_043), *i.e.*, not in any of these categories. The S sample – Sample_082 – also exhibited the highest SC (3485 μS cm^{-1}), Na (622 mg L^{-1}), and elevated levels of several other analytes. Its Ba concentration

Table 2 Compiled information of five selected spills for detailed examination (only Sample_003 has PADEP samples reported in its vicinity)

Spill identifier	Incident detail	Incident date	Remediation detail	2024 SWPA samples	# of SN samples (1 km)	# of PADEP samples (1 km)	# of PADEP samples (1.3 km)
A	Produced fluid release due to washout failure and slow leak of produced water from a dump line (19 213–22 425 barrels)	12/2020–12/2021	Groundwater & soil were impacted; pump and treat were done on site	Sample_003 (HS/S), Sample_013-3 (HS/S), Sample_042 (HS/S)	80 (Sample_003) 20 (Sample_013-3) 18 (Sample_042)	31	292
B	Produced fluid spill during transportation (50–100 gallons)	10/29/2020	Limestone aggregate near the well pad was removed	Sample_082 (S)	17	0	0
C	Produced fluid seeped out of containment	2/12/2021	Impacted soil and fluid were removed	Sample_090 (S), Sample_038 (S)	158 (Sample_090) 0 (Sample_038)	0	0
D	Resue water spill during transportation (11.8 barrels)	1/7/2021	Impacted soil was removed	Sample_081 (S), Sample_084 (S)	0	0	0
E	Produced fluid release while attempting to purge air (15 gallons)	6/6/2015	Impacted soil was removed	Sample_083 (S)	12	0	0

(2.05 mg L^{-1}) was at or slightly above the EPA MCL (2 mg L^{-1}), while F (3.52 mg L^{-1}) and Br (5.81 mg L^{-1}) concentrations exceeded their SMCLs (2 mg L^{-1}).

Fe and Mn surpassed SMCLs at 4 of 90 samples (4.4%) and 15 of 90 samples (16.7%), respectively. These secondary limits are based largely on aesthetic considerations. Li concentrations at 44 samples were higher than the EPA provisional toxicity value of 10 μg L^{-1},[23] and one sample (Sample_082) exceeded the U.S. Geological Survey (USGS) drinking water benchmark (60 μg L^{-1}[33]). All other trace metals (Be, Ni, Cu, As, Tl, Cd, Pb, and U) were below health-based thresholds. A few low-level detections of Li, Cu, and Pb are consistent with possible analytical blank interference.

All BTEX compounds were below EPA MCLs. Dissolved methane was low except for three samples that exceeded the Department of Interior action level (10 mg L^{-1}; Sample_086, Sample_017, Sample_055), including one (57 mg L^{-1}; Sample_086) above the immediate action level.

Methane (CH_4 or C_1) was detectable in 53% of samples, whereas ethane (C_2H_6 or C_2) and propane (C_3H_8 or C_3) were infrequently detected and occurred only where CH_4 was present. $\delta^{13}C–CH_4$ values ranged from −77‰ to −55‰ and $\delta D–CH_4$ from −220‰ to −154‰, and are consistent with published ranges for methane of mixed biogenic and thermogenic origin (Fig. 2A). $^{87}Sr/^{86}Sr$ ratios (0.711–0.713) were similar across all treatment samples and matched UOG brines for Pennsylvania[34] and the control group (Fig. 2B). In addition, these $^{87}Sr/^{86}Sr$ ratios were also significantly lower than those of COG brines from PA (≥0.720 (ref. 34)).

Cl/Br mass ratios were also examined to help distinguish the potential sources of elevated chloride (Fig. 2C and D). In hydrocarbon-bearing basins such as the Appalachian, Cl/Br ratios can indicate provenance of salinity: road salt and organic waste (*e.g.*, septic or agricultural runoff) typically contain little bromide, resulting in high Cl/Br ratios, whereas Appalachian Basin formation brines—whether naturally discharging or returning to the surface during UOGD—contain relatively high bromide and therefore much lower Cl/Br ratios.[18] Calculated Cl/Br ratios for this dataset show that most samples with elevated chloride ($>\sim$100 mg L^{-1}) plot within the high-ratio domain characteristic of surface anthropogenic sources rather than deep formation brines. Comparisons further indicate that the majority of samples, regardless of category, exhibit Cl/Br ratios much higher than those typical of UOG wastewaters. These results support the interpretation that most elevated chloride concentrations in the study area primarily reflect surface or shallow anthropogenic inputs rather than persistent influence from UOG-derived brines. The only notable exception is Sample_082, which plots closer to the Appalachian Basin brine zone (Fig. 2C). This high Br is consistent with a non-negligible brine contribution. This anomalous signature is examined in greater detail in subsequent sections.

3.2 Spatial and temporal patterns of groundwater chemistry

3.2.1 Comparison between treatment and control samples. Statistical comparisons between treatment (HS, I, S) and control (C) samples revealed limited systematic differences in groundwater chemistry (Table S7). When all water well and spring samples were included, seven analytes— NO_3, K, U, Cl, Pb, benzene, and Cu—were significantly higher for treatment samples than controls, whereas P and ethylbenzene were lower. For each treatment category, NO_3, K, U were elevated for HS samples; Pb for I samples; and K and Cu for S samples according to both WMW and BM tests. However, because agricultural land use is widespread in the study area, elevated K and NO_3 likely reflect fertilizer inputs rather than UOGD effects.

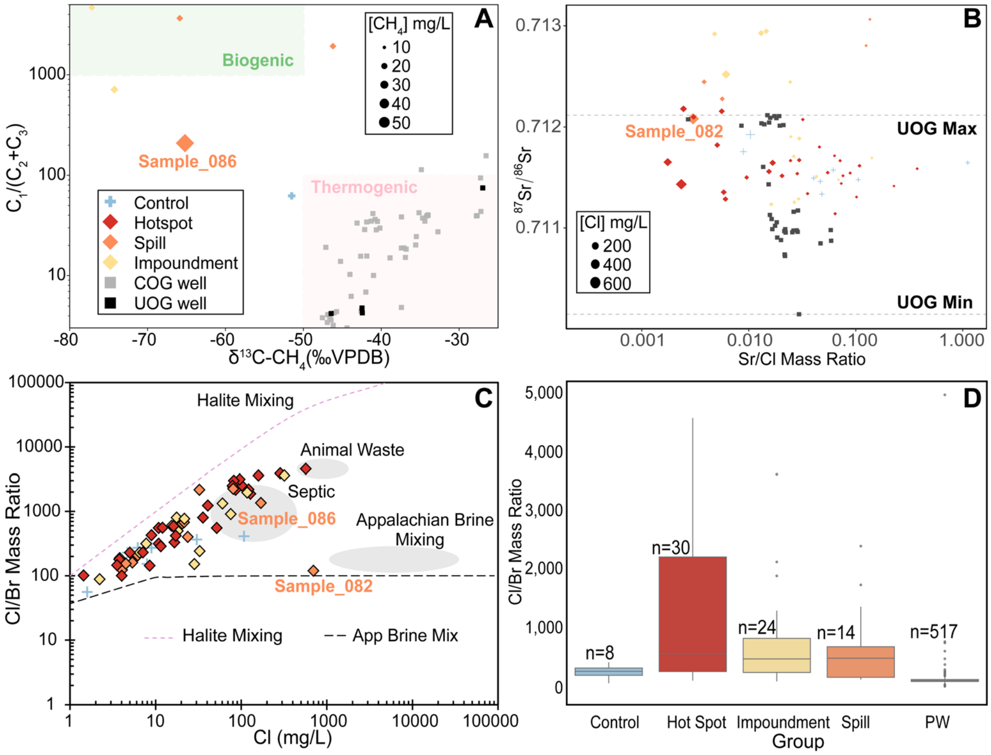

Fig. 2 (A) Concentrations of methane (C_1) *versus* ethane (C_2) + propane (C_3) plotted *versus* $\delta^{13}C-CH_4$ and (B) $^{87}Sr/^{86}Sr$ as a function of Sr/Cl mass ratio (in mg L^{-1}) of groundwater samples. The symbol sizes are proportional to CH_4 concentration in (A) and to Cl concentration in (B), values of production gases and produced waters from COG and UOG wells are compiled from the literature.[34,43] (C) Cl/Br mass ratios for HS, I, S, and C samples. Gray areas for septic and animal waste based on Katz *et al.* (2011).[44] Appalachian Basin brine reported for Pennsylvania based on Blondes *et al.* (2023).[34] (D) Comparison of Cl/Br mass ratio across HS, I, S, and C samples, as well as from the produced waters in Pennsylvania from Blondes *et al.* (2023).[34]

Indeed, spring samples exhibited significantly higher NO_3 than wells, consistent with agriculturally influenced recharge.

To reduce bias related to agricultural inputs, statistical tests were repeated for well samples only (Table S8). In this subset, five analytes—Cl, Mg, K, Pb, and Cu—were higher in treatment wells than in controls, while P and ethylbenzene were lower. Specifically, prior to FDR correction: (1) for HS samples, K, Cl, and Mg remained significantly higher; (2) for I samples, Pb was elevated; and (3) for S samples, K and Cu were higher.

Because multiple tests across 48 analytes increase the risk of false positives, we applied the Benjamini–Hochberg (BH) correction.[35] We set false discovery rate (FDR) = 0.1, meaning that we accepted up to 10% false positives. After correction, no analytes were significantly different between treatment and control groups in either the full dataset or the well-only subset.

3.2.2 Spatial relationships with UOG features and topography. To evaluate spatial relationships with potential contamination sources, correlations between brine-related and redox-sensitive analytes and proximity to UOG wells within

Table 3 Regression analysis results for select analyte concentrations and the distance (km) to hotspot, impoundment, and spill for all groundwater samples (* denotes $p < 0.05$; statistically significant, negative correlations are noted with bold text)

Treatment group	Coefficient	SC	Na	Cl	Ba	SO_4	Sr	Br	Fe	CH_4
HS	OLS slope	2.93	3.055*	0.119	0.003	−0.588	−0.005	0.012	0.002	0.206*
	Spearman	−0.099	0.026	−0.051	−0.041	−0.206	−0.109	0.079	0.139	0.310*
	Kendall's tau	−0.071	0.022	−0.038	−0.029	−0.133	−0.069	0.053	0.104	0.214*
	ATS slope	−3.793	0.121	−0.192	0.000	−0.686	−0.006	0.000	0.000	0.000*
I	OLS slope	11.504	4.617*	1.86	0.007	0.077	−0.005	0.029*	−0.002	0.097
	Spearman	−0.063	0.05	−0.085	−0.129	0.028	**−0.225***	−0.035	−0.156	0.144
	Kendall's tau	−0.034	0.033	−0.05	−0.086	0.03	**−0.151***	−0.026	−0.117	0.095
	ATS slope	−2.310	0.170	−0.236	−0.001	0.160	**−0.014***	−0.000	0.000	0.000
S	OLS slope	−15.816	**−5.954***	−4.165	**−0.014***	1.072	−0.011	−0.026	−0.005	−0.286
	Spearman	0	**−0.307***	−0.119	−0.084	0.208*	−0.048	−0.113	0.034	**−0.259***
	Kendall's tau	−0.002	**−0.209***	−0.084	−0.054	0.145*	−0.037	−0.077	0.024	**−0.172***
	ATS slope	−0.143	**−1.199***	−0.521	−0.001	1.106*	−0.005	−0.000	0.000	**−0.000***

Table 4 Fixed-effects regression coefficients calculated between analytes in all groundwater samples and distance (km) to the closest UOG wells in hotspots, impoundments, and spills (* denotes $p < 0.05$; statistically significant, negative correlations are noted with bold text)

Analyte	UOG well	Impoundment	Spill
SC	0.003 ($p = 0.687$)	0.007 ($p = 0.435$)	−0.01 ($p = 0.487$)
Na	0.044 ($p = 0.091$)	0.039 ($p = 0.176$)	**−0.105* ($p = 0.02$)**
Cl	−0.001 ($p = 0.974$)	−0.023 ($p = 0.465$)	−0.054 ($p = 0.279$)
Ba	0.003 ($p = 0.272$)	0.007* ($p = 0.013$)	−0.008 ($p = 0.071$)
SO_4	**−0.041* ($p = 0.01$)**	−0.028 ($p = 0.115$)	0.088* ($p = 0.001$)
Sr	−0.006 ($p = 0.234$)	−0.009 ($p = 0.116$)	−0.001 ($p = 0.923$)
Br	0.01* ($p = 0.017$)	0.017* ($p < 0.001$)	−0.001 ($p = 0.859$)
Fe	0.003 ($p = 0.309$)	0.002 ($p = 0.577$)	**−0.012* ($p = 0.021$)**
CH_4	0.048* ($p < 0.001$)	0.037* ($p = 0.013$)	**−0.068* ($p = 0.004$)**

hotspots, impoundments, and spills were examined (Table 3). Within hotspot areas, Na and CH_4 concentrations increased with distance from UOG wells, indicating no evidence of enrichment near active UOG wells. Near impoundments, Na, Sr, and Br showed significant correlations with distance, but only Sr increased closer to impoundments. Near spills, Na, Ba, and CH_4 exhibited negative correlations with distance, implying slight enrichment near spill locations. These results were consistent for both the combined (spring and well) and well-only datasets.

We also observed that topography influenced groundwater chemistry. Of the 90 samples, 26 were located in valleys and 64 on ridges. Valley samples had significantly higher Ba and CH_4 but lower SC, SO_4, Ca, Li, and Ni ($p < 0.05$ in both WMW and BM tests) (Table S10). After BH correction, only SO_4 remained lower in valley settings, consistent with enhanced reducing conditions.

After log-transforming analyte concentrations, the fixed-effects linear regression model (Table 4) estimated percentage change per kilometer of distance from each UOGD feature as:

$$\% \text{ increase} = [\exp(\beta) - 1] \times 100\% \quad (2)$$

Because distance was measured from each UOGD feature, negative coefficients indicate higher concentrations closer to the feature and are therefore most relevant for evaluating potential UOGD-related enrichment. After adjustment, distance to impoundments showed no significant negative relationship with any analyte concentration. In contrast, concentrations of SO_4 increased closer to UOG wells within hotspots, while Na, Fe, and CH_4 increased closer to spills. These proximity relationships correspond to estimated concentration increases

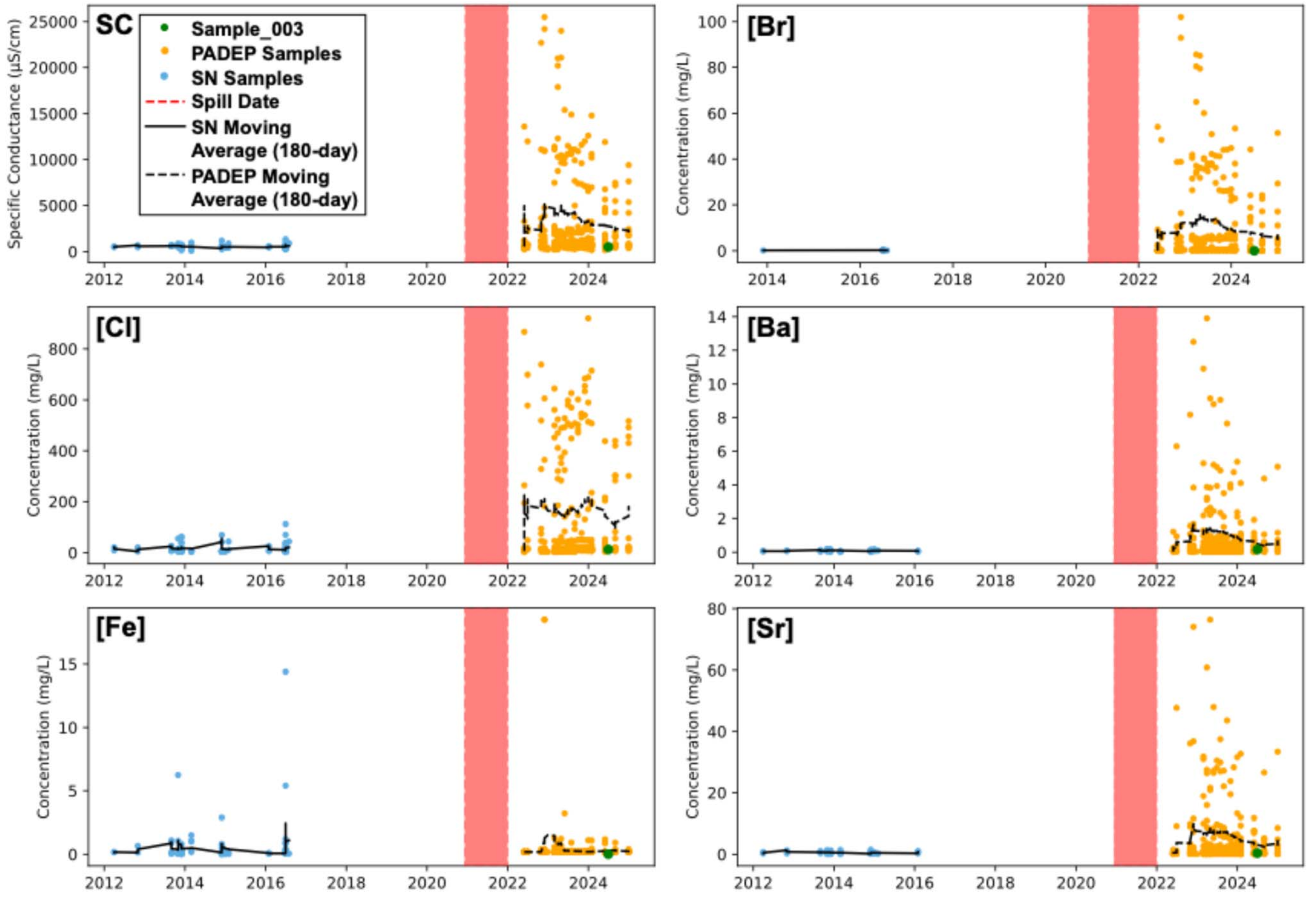

Fig. 3 Scatter plots of brine-related species concentration *vs.* time (year) for 2024 SWPA water sample Sample_003 (collected near spill A), SN dataset samples, and PADEP samples (within 1.3 km).

of 4.2% (SO_4), 11.1% (Na), 1.2% (Fe), and 7.0% (CH_4) per kilometer approaching the feature.

3.3 Investigation of local impacts from spills

Detailed examination of sampling near spills revealed localized and limited impacts (Tables S11–S16 and Fig. 3, S1–S6). Each site is discussed herein.

3.3.1 Spill A. The PADEP report attributes spill A to a washout failure of a buried dump line east of the wellhead and an estimated slow leak of the buried dump line (between December 2020 and December 2021) with the latter leak releasing approximately 19 213–22 425 barrels of produced fluid.

We collected three groundwater samples within the buffer area of spill A as part of the 2024 SWPA dataset but we focus on Sample_003, which was collected at the nearest point. We were able to compare Sample_003 to chemical analyses of 292 water samples collected as part of the PADEP study from 18 monitoring wells between May 2022 and December 2024 (see Text S5 for more details). Although the PADEP sampling overlapped with our sample timing, the PADEP water samples targeted more spill-adjacent locations, *i.e.*, less than ∼700 m downgradient from spill A. In comparison, our Sample_003 was collected 1.2 km downgradient from the spill (our Sample_013-3 and _042 were collected more than 2 km downgradient from spill A and from the PADEP samples).

Brine contamination from spill A is inferred for some of the PADEP water samples even at the same time point (29 months) when we collected Sample_003. In contrast, our Sample_003, collected 1.3 km from the furthest PADEP sampling sites, showed no geochemical indication of spill impact (Table S11): the sample showed markedly lower SC, Br, Cl, Ba, Fe, and Sr than PADEP samples (Fig. 3). Furthermore, Sample_003 exhibited comparable concentrations to the SN sample collected in 2013 at a location only 8 meters away, differing only in its higher methane content (Table S11). Regarding Sample_013-3 and Sample_042, the former showed lower concentrations of most brine-related species, whereas the latter exhibited slightly higher SC, Na, Cl, and Ba relative to its nearest SN sample. The SC, Na, Cl, and Ba concentrations in both samples are comparable to the median values of all nearby SN samples (Table S12, S13 and Fig. S2, S3).

3.3.2 Spill B. The temporal comparison for Sample_082 is shown in Fig. S4; unlike Spill A, no PADEP groundwater-monitoring dataset was available for direct comparison at Spill B.

Spill B was caused while a water truck was loading produced fluid from the production water tank, resulting in an estimated spill of 50–100 gallons onto the well pad surface on 10/29/2020. The release was contained to a small area of the compacted limestone pad, with no reported off-pad migration. Within the buffer radius of spill B, the only 2024 SWPA sample (Sample_082, 2422 m from the spill location; Fig. S4) exceeded drinking-water standards for Cl, Ba and Br and exhibited the highest salinity observed in the 2024 SWPA dataset. Sample_082 exhibited higher SC, Na, Cl, and Ba, and lower SO_4, Fe, and CH_4 (Table S14) compared to (i) the chemistry reported for the same site prior to the spill incident in the SN dataset, and (ii) the median of all prior SN samples within 1 km.

3.3.3 Spills C, D, E. No groundwater sample data were reported in PADEP documents near spills C, D, and E nor near the associated 2024 SWPA water sample locations (Table 2). We were able, however, to compare SN samples to all spills except D as described in the next paragraphs.

Spill C was described in the PADEP report as an incident wherein produced fluid (unknown volume) was found to be seeping from the pad gravel below secondary containment along the northeastern corner of a well pad. Near this site, two samples were collected in our 2024 campaign: Sample_090 (2592 m away) and Sample_038 (1966 m away). Historical SN data were available only near Sample_090. Sample_090 had much lower concentrations of most of the brine-related species compared to the closest SN sample and the median of all nearby SN samples (Table S15 and Fig. S5).

Spill D was described in the PADEP report as occurring on January 7, 2021, when a water hauler veered off the driveway apron while entering the lease road from State Route 519 near the Guyton well pad entrance. The truck overturned on the adjacent fill slope, releasing approximately 11.8 barrels of reuse water (likely brine) onto the surrounding ground. Sample_081 and Sample_084, from sites located near spill D, could not be compared to nearby SN samples because of the lack of samples within a 1 km radius.

Spill E occurred on June 6, 2015, and consisted of fifteen gallons of brine water released from containment when the company attempted to purge air from the water transfer line running across the well pad. Sample_083, sampled within ∼2800 m of this spill location, showed higher SC, Na, Cl and SO_4 concentrations compared to the median of nearby SN samples (Table S16 and Fig. S6).

4 Discussion

Groundwater chemistry across Washington and Greene Counties reflects a complex interplay of geology, hydrogeology, agricultural activities, and legacy hydrocarbon infrastructure. The absence of statistically robust differences between treatment and control sites after BH correction indicates that widespread or persistent contamination directly attributable to ongoing UOGD is unlikely under current conditions. Nevertheless, localized anomalies—particularly in sodium, barium, bromide, and methane near spills—could point to residual or intermittent effects related to surface fluid handling.

4.1 Legacy and localized impacts of UOGD on groundwater

Hotspots, impoundments, and spills together represent potential pathways for UOGD-related groundwater contamination in SWPA. Reassessing these features with statistical and geospatial approaches reveals little evidence of ongoing regional impacts attributable to UOGD, though isolated anomalies highlight the complexity of local hydrogeochemical systems. Our findings are consistent with a broader set of field-based groundwater studies in Pennsylvania that have generally found limited evidence for widespread groundwater degradation attributable to UOGD in the sampled settings, while recognizing that localized impacts can occur under specific hydrogeologic or operational conditions. In particular, McMahon *et al.* (2019)[36] reported that hydrocarbons in upland groundwater in the Marcellus Shale region of northeastern Pennsylvania and southern New York were not indicative of widespread regional groundwater degradation; similarly, Xiong *et al.* (2022)[37] found few organic chemical impacts in groundwaters of northeastern Pennsylvania near areas of intense hydraulic fracturing activity. These studies provide important context for interpreting our results: regional groundwater quality often appears broadly similar between areas with and without intensive UOGD, whereas detectable impacts, when present, tend to be spatially limited and strongly influenced by local hydrogeologic conditions, legacy infrastructure, and site-specific release histories.

Previously identified chloride hotspots in southwestern Pennsylvania were interpreted as potential indicators of wastewater-related impacts associated with UOGD.[13] Our 2024 sampling provides no evidence that such hotspot-related signals persist at a regional scale under current conditions. Groundwater collected within hotspot areas is geochemically similar to control locations, and no systematic enrichment of brine-associated constituents or trace metals of health concern is observed. Li exceeded its provisional health threshold (10 μg L^{-1}) in 44 samples across all categories, including controls, suggesting that elevated Li reflects regional groundwater quality characteristics rather than hotspot-specific contamination.[38] Exceedances of Mn and Fe were aesthetic and reflect regional lithologic norms, as these metals are ubiquitous in Pennsylvania bedrock and groundwater particulates.[11] Although modest differences appear in uncorrected statistical tests, these do not persist after BH correction, indicating that any earlier hotspot-related impacts are not regionally detectable in the present dataset under current conditions.

Shaheen *et al.* (2024)[14] attributed small increases in brine salts near UOGD operations to occasional, localized surface fluid releases rather than subsurface hydraulic fracturing processes. Consistent with this interpretation, we observed significant correlations between Sr (in 3 of 4 statistical tests) and proximity to impoundments, and between Na (all 4 tests), Ba (1 of 4), and CH_4 (3 of 4) and proximity to spills. After applying fixed-effects regression that accounted for topographic position, coal mining, and conventional oil and gas wells, however, most of these relationships were no longer statistically significant. In the full fixed-effects model, Na, Fe, and CH_4 showed significant negative relationships with distance to spills, indicating higher concentrations closer to spill locations (Table 4). These results suggest that limited enrichment of Na and CH_4 may reflect residual impacts from surface or shallow subsurface wastewater releases rather than well integrity failures or deep formation leakage. Sensitivity analyses excluding Sample_082 showed that negative relationships for Na, Ba, Fe, and CH_4 remained significant, while SO_4 showed a significant positive relationship with distance to spills (Table S17). Thus, the proximity relationships were not driven solely by Sample_082, although Sample_082 likely contributed disproportionately to the strongest brine-like signal.

4.1.1 Effects of spills.

With respect to sample chemistry, only one of the five spills examined in detail (spill B) is associated with a clear brine-like geochemical anomaly (Sample_082). Groundwater samples associated with the other spills (A, C, D and E) from our 2024 sampling show no evidence of persistent salinity, brine-associated constituents, or systematic deviation from nearby historical or background conditions. Of these spills, spill A is nonetheless of interest. PADEP reports documented a brine contamination plume extending from spill A to as far as ~700 m even during the time period when we sampled. But Samples_003, _013-3, and _042, all located at least 1.2 km from spill A, show no elevated salinity. This suggests that any brine contamination associated with spill A was spatially limited, either by containment or dilution, or that, if the contamination had originally extended as far as 1.2 km from the spill, it had attenuated by 2024, about 29 months after spillage. These observations are consistent with the PADEP reports (Text S5) as well as our inferred buffer radius (2 km) and time lag (3 years) for spills.

The most anomalous of our samples is Sample_082, collected near spill B, which exhibits distinctly elevated salinity and brine-associated constituents relative to both nearby controls and co-located historical data. Its ionic composition, Cl/Br ratio, and Sr isotopic signature indicate a non-negligible contribution from Appalachian Basin brine or produced water (Fig. 2B and C). This geochemical fingerprint is unique within the 2024 SWPA dataset and is not observed at other spill sites, underscoring the highly localized nature of this anomaly.

Comparison with historical SN data sheds some light on the interpretation. The 2024 SWPA sample collected at this site exhibits markedly higher specific conductance, Na, Cl, Ba, and Br and lower SO_4 relative to both the co-located SN sample and the median of nearby SN observations, indicating a site-specific increase in brine-associated constituents over time. The decrease in SO_4 is also consistent with the low solubility of barite ($BaSO_4$), which produces inverse behavior for Ba and SO_4 when Ba-containing brines contaminate aquifers.[39] While the chemistry is consistent with either a natural ABB or produced water source, the change over time is more consistent with an explanation involving a localized influence from a produced-water (brine) release.

However, the available data do not allow definitive attribution of the brine source impacting this sample without more detailed site-specific investigation. For example, natural Appalachian Basin brine cannot be fully excluded given that natural brines are known to have impacted waters in SWPA[40] and temporal changes in water chemistry are relatively common. Furthermore, spill contamination may be unlikely given the small volume released (50–100 gallons), the reportedly prompt response, and the ~2.4 km distance between Sample_082 and spill B. In addition, there are other potential explanations. For example, within 2 km of Sample_082, at least 10 active COGD wells and other associated infrastructure (*e.g.*, compressor stations) are present.

Nonetheless, to explore possible sources for contamination, the surface drainage area for Sample_082 (and all other samples in the 2024 dataset) were delineated using the "Watershed" tool in ArcGIS Pro. The drainage area for Sample_082 was small (0.01 km^2), as were the drainage areas for all the samples in the 2024 dataset (maximum area of 1.4 km^2). For example, when each watershed is treated as a circle of equivalent area, the largest diameter is approximately 670 m. None of the drainage areas, including the area for Sample_082, contain an impoundment or spill location. This implies that, if produced waters (brines) explain the chemistry of Sample_082 (or any of the samples in the 2024 dataset), migration likely occurs along deeper groundwater flow paths rather than surficial pathways. Such deeper migration could be facilitated by valley-focused flow convergence and favorable hydrogeologic connectivity such as fractures. This interpretation is consistent with the geologic setting at spill B, which overlies the Waynesburg Formation—a clastic bedrock unit with extremely low primary porosity because the pore spaces are filled with calcareous or siliceous cement.[41] Given these characteristics, it is possible that the local aquifer system at spill B exhibits the same fracture-dominant flow regime observed at spill A (Text S5), and local fractures serve as preferential pathways that facilitate both vertical migration and lateral transport of the brine-associated constituents.

4.2 Hydrogeologic controls and implications for monitoring

Groundwater vulnerability in southwestern Pennsylvania is strongly shaped by topography and hydrogeologic conditions, which together govern flow-path connectivity, residence time, and the potential for interaction with deeper fluids. Valley settings, in particular, represent zones of convergent flow where intermediate and deep groundwater systems may intersect shallow aquifers, increasing susceptibility to both natural Appalachian Basin brine influence and any contaminants introduced at or near the surface. In contrast, ridge settings are more commonly associated with shorter, shallower flow paths and greater dilution by recent recharge. These controls provide essential context for interpreting why potential impacts from UOGD-related activities are not spatially uniform and why proximity alone is an unreliable predictor of groundwater response. For example, in this study, valley samples in the 2024 SWPA dataset exhibited higher Ba and CH_4 and lower SO_4 and Ca relative to ridge samples, consistent with reducing conditions and natural upwelling of deeper fluids.[11,13,42] After BH correction, however, only sulfate remained significantly lower in valley settings—a pattern consistent with Shaheen *et al.* (2022),[13] who attributed the muted topographic influence on groundwater chemistry in SWPA to the region's mild relief, suggesting that much of the observed spatial variability reflects natural hydrogeologic heterogeneity rather than UOGD-related impacts.

The strong heterogeneity of geology, topography, land use, and legacy infrastructure in this region also complicates both statistical inference and the definition of appropriate control samples. In settings where lithology, redox conditions, agricultural inputs, road salting, coal mining, and conventional oil and gas development co-vary over short distances, background groundwater chemistry can vary as much as or more than any potential UOGD-related signal. As a result, no universal sample size guarantees detection of subtle effects, and control samples are inherently imperfect. This heterogeneity increases the sample size required to resolve small regional trends and reduces the power of broad spatial comparisons, particularly when contamination events are rare and spatially constrained.

In this study, only one of the five spills examined in detail (spill B) is associated with a groundwater sample exhibiting a clear brine-like geochemical anomaly (Sample_082). Our groundwater samples associated with the other four spills (A, C, D, and E) show no evidence of persistent salinity, brine-associated constituents, or systematic deviation from nearby historical or background conditions: this observation is especially interesting in that samples taken by the state regulator (PADEP) during the same time period as our sampling but 500 m closer to spill A did show evidence of brine contamination. This pattern indicates that persistent impacts were not detected at most of the spill/brownfield sites examined here, and that when impacts were detected, they appeared highly localized and strongly conditioned by site-specific hydrogeologic setting rather than spill occurrence alone.

In this sense, the present study extends prior field-based work by using previously identified chloride hotspots and documented spill or impoundment locations as a targeted sampling framework for evaluating whether regional signals identified in large datasets correspond to persistent local groundwater impacts. Our findings underscore the complementary roles of regional datasets and targeted, hydrogeology-informed field investigations in evaluating groundwater impacts associated with UOGD. Large datasets are essential for identifying subtle regional patterns and for flagging areas of potential concern, but they cannot resolve rare, spatially constrained impacts. Targeted, hydrogeology-informed sampling is required to directly test hypotheses generated from regional analyses and to detect localized anomalies when they occur. In heterogeneous settings such as southwestern Pennsylvania, a tiered approach—regional screening followed by focused field investigation—provides a useful framework for distinguishing widespread trends from isolated, site-specific impacts.

4.3 Study design limitations and implications for interpretation

Several aspects of the sampling design reflect practical constraints inherent to field-based groundwater studies in regions dominated by private domestic wells. Although the study was designed to target previously identified chloride hotspots, decommissioned impoundments, documented spills, and control areas, final sampling locations depended on domestic-well availability, homeowner participation, and access permission. Therefore, the sampling design should be interpreted as targeted and feasibility-constrained rather than random, stratified, or spatially exhaustive. The inclusion of

"Other" samples improved spatial coverage and provided additional regional context, but these samples were not intended to represent a formally defined exposure category. Similarly, the expansion of the spill sampling radius from 2 to 3 km and the inclusion of a small number of samples associated with older brownfield or spill sites were necessary to obtain sufficient samples near potential release locations, but these decisions also introduce uncertainty in exposure classification and limit source-specific inference. In addition, because control-site selection was constrained by homeowner participation and well availability, control samples could not always be matched precisely with treatment samples in terms of hydrogeologic, topographic, and land-use setting (Fig. S7). These constraints may reduce treatment-control contrast, introduce participation bias, and limit the ability to detect either regional patterns or highly localized impacts. Consequently, null regional results should be interpreted as evidence that persistent impacts were not detectable under the sampling conditions of this study, rather than as proof that impacts are absent everywhere.

5 Conclusions

Based on our sample set, groundwater in Washington and Greene Counties of southwestern Pennsylvania largely meets U.S. Environmental Protection Agency drinking-water standards and shows no evidence of widespread or persistent degradation attributable to ongoing UOGD under current conditions. Groundwater samples collected within previously identified chloride hotspots, near decommissioned impoundments, and near recently documented spill sites were geochemically similar to control locations after BH correction was applied. However, identification of control sites was challenging because of the highly heterogeneous nature of geology, topography, and land use in the area.

Despite the absence of a regional signal, only one of the five spill sites examined in detail showed a localized anomaly—a single groundwater sample with elevated salinity and brine-like signatures relative to nearby controls and co-located historical data, suggesting increased brine influence over time. Samples from the other four sites show no evidence of persistent salinity or enrichment in brine-associated constituents, indicating that persistent groundwater contamination from surface releases was uncommon and highly localized among the spill/brownfield sites examined here. The isolated anomaly observed at Sample_082 is consistent with localized influence from produced water or natural Appalachian Basin brine migration along deeper or intermediate flow paths; however, definitive attribution to a specific source or event is not possible based on the available data.

Together, these results directly address the central questions motivating this study. We find no evidence that regional brine-related contamination signals identified during the early period of UOGD persist under current regulatory and operational conditions, nor that previously identified chloride hotspots remain detectably enriched relative to controls after controlling FDR. With respect to localized impacts, only one of five examined spill sites exhibited a clear brine-like geochemical anomaly, indicating that persistent groundwater contamination from surface releases is rare and highly localized. Exceedances of enforceable primary MCLs were limited. Other exceedances or screening-threshold exceedances, including Li, occurred across multiple sample categories and were largely unrelated to proximity to UOGD features, suggesting that present-day human health concerns associated with groundwater contamination in this region are limited and site-specific rather than regional in scale.

These results also reinforce that, within our targeted sample set, detectable groundwater impacts associated with UOGD-related features are rare, highly localized, and strongly conditioned by site-specific hydrogeologic setting rather than proximity alone. Large regional datasets are essential for identifying subtle patterns and flagging areas of potential concern, but targeted, hydrogeology-informed field investigations are required to detect and characterize localized impacts when they occur. A tiered monitoring strategy that combines regional screening with focused sampling in hydrogeologically vulnerable settings provides an effective framework for distinguishing isolated, site-specific contamination from widespread trends.

Author contributions

Conceptualization: J. S., S. S., S. B., J. B., T. W.; data curation: J. S., L. W., B. F.; formal analysis: J. S., S. S., N. W., T. W.; funding acquisition: S. B., J. B., N. W., T. W.; investigation: J. S., L. W., S. S., S. B., T. W.; methodology: J. S., S. S., S. B., N. W., J. B., T. W.; project administration: S. B., J. B., T. W.; resources: J. S., L. W., S. S., T. W.; software: J. S., S. S.; supervision: S. B., J. B., N. W., T. W.; validation: J. S., L. W., S. S., T. W.; visualization: J. S., L. W., A. F., S. S., T. W.; writing – original draft: J. S., T. W., S. B.; writing – review & editing: J. S., L. W., A. F., S. S., S. B., B. F., N. W., J. B., T. W.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data and scripts discussed and used in this paper are freely accessible through the data https://doi.org/10.5281/zenodo.20213462. Supplementary information: detailed field and laboratory methods, quality control, PADEP record review and incident characterization, domestic well recruitment and community engagement, supplemental figures, and supporting statistical and geochemical tables. See DOI: https://doi.org/10.1039/d6em00142d.

Acknowledgements

This work was supported by the Health Effects Institute (grant number 143815 to Jennifer Baka, Pennsylvania State University, with subcontract to T. W. at Syracuse University). T. W. would also like to acknowledge support from the National Science

Foundation (grant number OAC-2209864). Research described in this article was conducted under contract to the Health Effects Institute (HEI), an organization jointly funded by the United States Environmental Protection Agency (EPA) (Contract No. 68HERC19D0010) and certain oil and natural gas companies. Although the research was produced with partial funding by EPA and industry, they have not been subject to their review, and therefore the research does not necessarily reflect the views of the Agency or the oil and natural gas industry, and no official endorsement by the Agency or the industry should be inferred.

References

1 O. Harrington, J. Baka and S. Shaheen, An anatomy of inscrutability of the unconventional oil and gas development-water-health nexus in Pennsylvania, *Energy Res. Soc. Sci.*, 2025, **130**, 104413, DOI: 10.1016/j.erss.2025.104413.

2 R. D. Vidic, S. L. Brantley, J. M. Vandenbossche, D. Yoxtheimer and J. D. Abad, Impact of Shale Gas Development on Regional Water Quality, *Science*, 2013, **340**(6134), 1235009, DOI: 10.1126/science.1235009.

3 P. A. Hammond, T. Wen, J. Woda and D. Oakley, Pathways and Environmental Impacts of Methane Migration: Case Studies in the Marcellus Shale, USA, *Geofluids*, 2024, **2024**, e9290873, DOI: 10.1155/2024/9290873.

4 P. A. Hammond, T. Wen, S. L. Brantley and T. Engelder, Gas well integrity and methane migration: evaluation of published evidence during shale-gas development in the USA, *Hydrogeol. J.*, 2020, **28**(4), 1481–1502, DOI: 10.1007/s10040-020-02116-y.

5 G. Lackey, H. Rajaram, J. Bolander, *et al.*, Public data from three US states provide new insights into well integrity, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**(14), e2013894118, DOI: 10.1073/pnas.2013894118.

6 S. G. Osborn, A. Vengosh, N. R. Warner and R. B. Jackson, Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**(20), 8172–8176, DOI: 10.1073/pnas.1100682108.

7 N. R. Warner, R. B. Jackson, T. H. Darrah, *et al.*, 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, DOI: 10.1073/pnas.1121181109.

8 J. Woda, T. Wen, D. Oakley, *et al.*, Detecting and explaining why aquifers occasionally become degraded near hydraulically fractured shale gas wells, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**(49), 12349–12358, DOI: 10.1073/pnas.1809013115.

9 L. O. Haluszczak, A. W. Rose and L. R. Kump, Geochemical evaluation of flowback brine from Marcellus gas wells in Pennsylvania, USA, *Appl. Geochem.*, 2013, **28**, 55–61, DOI: 10.1016/j.apgeochem.2012.10.002.

10 U.S. Environmental Protection Agency, *National Primary Drinking Water Regulations*, U.S. Environmental Protection Agency, 2024, accessed December 16, 2024, https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations.

11 D. I. Siegel, B. Smith, E. Perry, R. Bothun and M. Hollingsworth, 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, DOI: 10.1016/j.apgeochem.2015.06.013.

12 K. O. Maloney, S. Baruch-Mordo, L. A. Patterson, *et al.*, Unconventional oil and gas spills: Materials, volumes, and risks to surface waters in four states of the U.S, *Sci. Total Environ.*, 2017, **581–582**, 369–377, DOI: 10.1016/j.scitotenv.2016.12.142.

13 S. W. Shaheen, T. Wen, A. Herman and S. L. Brantley, Geochemical Evidence of Potential Groundwater Contamination with Human Health Risks Where Hydraulic Fracturing Overlaps with Extensive Legacy Hydrocarbon Extraction, *Environ. Sci. Technol.*, 2022, **56**(14), 10010–10019, DOI: 10.1021/acs.est.2c00001.

14 S. W. Shaheen, T. Wen, Z. Zheng, L. Xue, J. Baka and S. L. Brantley, Wastewaters Coproduced with Shale Gas Drive Slight Regional Salinization of Groundwater, *Environ. Sci. Technol.*, 2024, **58**(40), 17862–17873, DOI: 10.1021/acs.est.4c03371.

15 Pennsylvania Department of Environmental Protection, *PADEP SPUD Data Report*, 2025, accessed November 7, 2025, https://cedatareporting.pa.gov/Reportserver/Pages/ReportViewer.aspx?/Public/DEP/OG/SSRS/Spud_External_Data.

16 Pennsylvania Department of Environmental Protection, *Pennsylvania DEP Fines Range Resources $4.15 Million for Violating Environmental Regulations*, September 18, 2014, accessed September 21, 2025, https://www.prnewswire.com/news-releases/pennsylvania-dep-fines-range-resources-415-million-for-violating-environmental-regulations-275638111.html.

17 M. A. Soriano, N. C. Deziel and J. E. Saiers, Regional Scale Assessment of Shallow Groundwater Vulnerability to Contamination from Unconventional Hydrocarbon Extraction, *Environ. Sci. Technol.*, 2022, **56**(17), 12126–12136, DOI: 10.1021/acs.est.2c00470.

18 G. T. Llewellyn, F. Dorman, J. L. Westland, *et al.*, Evaluating a groundwater supply contamination incident attributed to Marcellus Shale gas development, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**(20), 6325–6330, DOI: 10.1073/pnas.1420279112.

19 T. E. Larson, J. P. Nicot, P. Mickler, *et al.*, Monitoring Stray Natural Gas in Groundwater With Dissolved Nitrogen. An Example From Parker County, Texas, *Water Resour. Res.*, 2018, **54**(9), 6024–6041, DOI: 10.1029/2018WR022612.

20 T. Wen, M. C. Castro, J. P. Nicot, *et al.*, 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, DOI: 10.1021/acs.est.6b01494.

21 U.S. Environmental Protection Agency, *Secondary Drinking Water Standards: Guidance for Nuisance Chemicals*, U.S.

Environmental Protection Agency, June 2, 2025, accessed May 6, 2025, https://www.epa.gov/sdwa/secondary-drinking-water-standards-guidance-nuisance-chemicals.

22 K. K. Eltschlager, J. W. Hawkins, W. C. Ehler and F. Baldassare. Technical Measures for the Investigation and Mitigation of Fugitive Methane Hazards in Areas of Coal Mining, Office of Surface Mining Reclamation and Enforcement (DI), Pittsburgh, PA, 2001, accessed September 25, 2025, https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/PB2002106133.xhtml.

23 C. D. Smith and L. H. Nowell, *Health-Based Screening Levels for Evaluating Water-Quality Data*, U.S. Geological Survey, 3rd edn, 2024, DOI: 10.5066/F71C1TWP, https://water.usgs.gov/water-resources/hbsl/.

24 T. Wen, A. Agarwal, L. Xue, *et al.*, Assessing changes in groundwater chemistry in landscapes with more than 100 years of oil and gas development, *Environ. Sci.: Processes Impacts*, 2019, **21**(2), 384–396, DOI: 10.1039/C8EM00385H.

25 U.S. Geological Survey, *NHDPlus High Resolution*, U.S. Geological Survey, 2018, accessed October 13, 2025, https://www.usgs.gov/national-hydrography/nhdplus-high-resolution.

26 A. Weiss, Topographic position and landforms analysis, in *Poster Presentation, ESRI User Conference*, San Diego, CA, 2001, vol. 200.

27 Pennsylvania Department of Environmental Protection, *PA Department of Environmental Protection Open Data Portal*, 2025, accessed October 14, 2025, https://newdata-padep-1.opendata.arcgis.com/.

28 US Census Bureau, *TIGER/Line Shapefiles*, The United States Census Bureau, 2025, accessed October 13, 2025, https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html.

29 F. Epuna, S. W. Shaheen and T. Wen, Road salting and natural brine migration revealed as major sources of groundwater contamination across regions of northern Appalachia with and without unconventional oil and gas development, *Water Res.*, 2022, **225**, 119128, DOI: 10.1016/j.watres.2022.119128.

30 J. D. Johnson, J. R. Graney, R. C. Capo and B. W. Stewart, Identification and quantification of regional brine and road salt sources in watersheds along the New York/Pennsylvania border, USA, *Appl. Geochem.*, 2015, **60**, 37–50, DOI: 10.1016/j.apgeochem.2014.08.002.

31 Pennsylvania Department of Environmental Protection, *Land Recycling Reports*, 2025, accessed October 14, 2025, https://www.pa.gov/agencies/dep/data-and-tools/reports/land-recycling-reports.

32 S. L. Brantley, *Shale Network Database*, 2015, DOI: 10.4211/his-data-shalenetwork.

33 B. D. Lindsey, K. Belitz, C. A. Cravotta, P. L. Toccalino and N. M. Dubrovsky, Lithium in groundwater used for drinking-water supply in the United States, *Sci. Total Environ.*, 2021, **767**, 144691, DOI: 10.1016/j.scitotenv.2020.144691.

34 M. S. Blondes, K. J. Knierim, M. R. Croke, *et al.U.S. Geological Survey National Produced Waters Geochemical Database (ver. 3.0, December 2023)*, Published online, 2023, DOI: 10.5066/P9DSRCZJ.

35 Y. Benjamini and Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, *J. R. Stat. Soc., Ser. B*, 1995, **57**(1), 289–300, DOI: 10.1111/j.2517-6161.1995.tb02031.x.

36 P. B. McMahon, B. D. Lindsey, M. D. Conlon, *et al.*, Hydrocarbons in Upland Groundwater, Marcellus Shale Region, Northeastern Pennsylvania and Southern New York, U.S.A, *Environ. Sci. Technol.*, 2019, **53**(14), 8027–8035, DOI: 10.1021/acs.est.9b01440.

37 B. Xiong, M. A. Soriano, K. M. Gutchess, *et al.*, Groundwaters in Northeastern Pennsylvania near intense hydraulic fracturing activities exhibit few organic chemical impacts, *Environ. Sci.: Processes Impacts*, 2022, **24**(2), 252–264, DOI: 10.1039/D1EM00124H.

38 M. A. Lombard, E. E. Brown, D. M. Saftner, *et al.*, Estimating Lithium Concentrations in Groundwater Used as Drinking Water for the Conterminous United States, *Environ. Sci. Technol.*, 2024, **58**(2), 1255–1264, DOI: 10.1021/acs.est.3c03315.

39 X. Niu, A. Wendt, Z. Li, *et al.*, 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, DOI: 10.1007/s10653-017-0031-6.

40 W. C. Dzombak, *Salt in the Conemaugh Valley : the Importance of the Conemaugh Salt Works in the Early American Salt Industry (1800-1860)*, Saltsburg Historical Soc., 2004.

41 U.S. Geological Survey, Water resources and the hydrologic effects of coal mining in Washington County, Pennsylvania, 1993, accessed January 19, 2026, https://www.usgs.gov/publications/water-resources-and-hydrologic-effects-coal-mining-washington-county-pennsylvania.

42 T. Wen, X. Niu, M. Gonzales, G. Zheng, Z. Li and S. L. Brantley, 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**(12), 7149–7159, DOI: 10.1021/acs.est.8b01123.

43 O. A. Sherwood, S. Schwietzke, V. A. Arling and G. Etiope, Global Inventory of Gas Geochemistry Data from Fossil Fuel, Microbial and Burning Sources, version 2017, *Earth Syst. Sci. Data*, 2017, **9**(2), 639–656, DOI: 10.5194/essd-9-639-2017.

44 B. G. Katz, S. M. Eberts and L. J. Kauffman, Using Cl/Br ratios and other indicators to assess potential impacts on groundwater quality from septic systems: A review and examples from principal aquifers in the United States, *J. Hydrol.*, 2011, **397**(3), 151–166, DOI: 10.1016/j.jhydrol.2010.11.017.