Environmental Factors Associated With Natural Methane Occurrence in the Appalachian Basin

by Lisa J. Molofsky[1], John A. Connor[2], Thomas E. McHugh[3], Stephen D. Richardson[4], Casper Woroszlyo[5], and Pedro J. Alvarez[6]

Abstract

The recent boom in shale gas development in the Marcellus Shale has increased interest in the methods to distinguish between naturally occurring methane in groundwater and stray methane associated with drilling and production operations. This study evaluates the relationship between natural methane occurrence and three principal environmental factors (groundwater redox state, water type, and topography) using two pre-drill datasets of 132 samples from western Pennsylvania, Ohio, and West Virginia and 1417 samples from northeastern Pennsylvania. Higher natural methane concentrations in residential wells are strongly associated with reducing conditions characterized by low nitrate and low sulfate ($[NO_3^-] < 0.5$ mg/L; $[SO_4^{2-}] < 2.5$ mg/L). However, no significant relationship exists between methane and iron [Fe(II)], which is traditionally considered an indicator of conditions that have progressed through iron reduction. As shown in previous studies, water type is significantly correlated with natural methane concentrations, where sodium (Na) -rich waters exhibit significantly higher ($p < 0.001$) natural methane concentrations than calcium (Ca)-rich waters. For water wells exhibiting Na-rich waters and/or low nitrate and low sulfate conditions, valley locations are associated with higher methane concentrations than upland topography. Consequently, we identify three factors ("Low NO_3^- & SO_4^{2-}" redox condition, Na-rich water type, and valley location), which, in combination, offer strong predictive power regarding the natural occurrence of high methane concentrations. Samples exhibiting these three factors have a median methane concentration of 10,000 μg/L. These heuristic relationships may facilitate the design of pre-drill monitoring programs and the subsequent evaluation of post-drill monitoring results to help distinguish between naturally occurring methane and methane originating from anthropogenic sources or migration pathways.

[1]Corresponding author: GSI Environmental Inc., 2211 Norfolk St., Suite 1000, Houston, TX 77098; 713 522 6300; fax: 713 522 8010; lmolofsky@gsi-net.com

[2]GSI Environmental Inc., Houston, TX; jaconnor@gsi-net.com

[3]GSI Environmental Inc., Houston, TX; temchugh@gsi-net.com

[4]GSI Environmental Inc., Austin, TX; sdrichardson@gsi-net.com

[5]Department of Civil and Environmental Engineering and Department of Statistics, Rice University, Houston, TX; cerosz@gmail.com

[6]Department of Civil and Environmental Engineering and Department of Statistics, Rice University, Houston, TX; alvarez@rice.edu

Introduction

The historical occurrence of methane in residential water wells in parts of the Appalachian Basin (Pennsylvania, New York, West Virginia) has long been recognized as a natural phenomenon (Blackman 1873; Lohman 1939; Soren 1963; Mathes and White 2006; Breen et al. 2007; Susquehanna Historical Society 2010; Williams 2010; Kappel and Nystrom 2012; Perry et al. 2012; Heisig and Scott 2013; Molofsky et al. 2013; Sloto 2013; Baldassare et al. 2014; Siegel et al. 2015a, 2015b). However, the recent boom in gas development in the Marcellus Shale has led to increased concern over the potential for stray methane impacts from shale gas extraction activities on drinking water resources. This concern, in turn, has led to a greater need for scientific methods to distinguish between natural and anthropogenic occurrence of methane in water wells.

The isotopic and molecular composition of different sources of methane gas has been successfully used to determine the origin of gases in water wells (e.g., Fuex 1977; Schoell 1990; Clayton 1991; Coleman 1994; Baldassare and Laughrey 1997). In addition, the practice of sampling proximate water wells prior to the drilling of shale gas wells (i.e., "pre-drill" sampling) has become

commonplace to establish baseline groundwater quality for comparison to "post-drill" groundwater conditions. Nevertheless, these methods do not always clearly discriminate between naturally occurring methane and stray gas. For example, methane from gas-charged units that are either intersected or targeted for production by shale gas wells can also, in some areas, naturally migrate into shallow groundwater through existing fault and fracture networks (Baldassare et al. 2014). In addition, changes in concentrations of methane and other water quality parameters over time (e.g., an increase in methane from 3 to 10 mg/L) can be associated with natural variability rather than shale gas extraction activities (Coleman et al. 2012). In these instances, additional information regarding the expected range of natural methane concentrations in groundwater can provide another important line of evidence for assessing whether elevated methane concentrations in water wells are related to nearby gas extraction activities or natural phenomena.

In the Appalachian Basin, naturally elevated concentrations of methane in residential wells have been associated with (1) topography, where methane occurs more frequently in valleys than uplands (Mathes and White 2006; Heisig and Scott 2013; Molofsky et al. 2013; Siegel et al. 2015a, 2015b); (2) water type, where methane occurs more frequently in sodium (Na)- and chloride (Cl)-rich waters than calcium (Ca)-rich waters (Perry et al. 2012; Warner et al. 2012; Molofsky et al. 2013; Darrah et al. 2014; McPhillips et al. 2014; Siegel et al. 2015b); and (3) aquifer type, where methane occurs more frequently in confined bedrock aquifers (Kresse et al. 2012; Heisig and Scott 2013; Siegel et al. 2015b). The presence of elevated dissolved methane concentrations in valley settings has been attributed to a number of factors, including (1) enhanced fracture networks in valleys that may serve as natural pathways for methane migration, (2) thinner freshwater flow systems in valleys (vs. uplands) that are in closer proximity to underlying methane-charged brine units, (3) confining units of glacial-drift deposits in valleys that may both slow groundwater flow and impede the loss of methane to the atmosphere, and (4) upward groundwater flow gradients in valleys that assist (and at a minimum do not impede) the upward seepage of methane (Mathes and White 2006; Heisig and Scott 2013; Molofsky et al. 2013; Siegel et al. 2015b). Similarly, it has been suggested that the association of Na- and Cl-rich water types with elevated dissolved methane concentrations implies access to deeper groundwater units that have experienced longer groundwater residence times (and therefore longer rock-water interaction times) and/or are in contact with transitional saline groundwater or brine (Warner et al. 2012; Heisig and Scott 2013; Molofsky et al. 2013; Darrah et al. 2014; McPhillips et al. 2014; Siegel et al. 2015b).

In addition to these factors, dissolved methane concentrations in an aquifer should exhibit a strong relationship with the redox state. Specifically, methane can be produced by methanogenic bacteria under highly reducing conditions (a process known as methanogenesis). However, methane can also be consumed by methanotrophs in the presence of oxygen or other electron acceptors (Table S1, Supporting Information). In this case, the co-consumption of methane and electron acceptors leads to strongly reducing conditions (characterized by a relative lack of electron acceptors) (e.g., Kelly and Matisoff 1985; van Stempvoort et al. 2005). Aerobic degradation reactions typically proceed on a rapid timescale relative to subsurface transport processes (Weidemeier et al. 1999). As a result, either the methane or oxygen is typically completely consumed, leaving the remaining portion of one or the other. Conversely, if methane enters an environment devoid of oxygen or other electron acceptors (i.e., an environment that is already highly reducing), the methane may remain undegraded for an extended period of time.

The relationship between methane concentrations and redox dynamics has not been systematically investigated in residential water wells. This relationship could be confounded by (1) in-well mixing of water from different flow pathways or fractures, and (2) gas-phase transport of methane (e.g., Gorody 2012; Siegel et al., 2015b). These factors may be particularly relevant in the Appalachian Basin, where most water wells are open-borehole completions fed by fracture flow. To increase storage capacity, these water wells often have long boreholes that intersect multiple fractures. In some cases, the primary fracture providing water to the well may not be the same fracture that is yielding the majority of the naturally occurring methane. Furthermore, the rate and extent of methane oxidation occurring within the wellbore itself is generally unknown. Consequently, the relationship between methane and redox conditions in residential water wells remains an open question.

This study investigates the relationship between naturally occurring methane concentrations and the prevailing groundwater redox condition in residential water wells located in two areas in the Appalachian Basin: (1) Susquehanna County in northeastern Pennsylvania (1417 water wells), and (2) a larger geographic area consisting of 20 smaller counties in western Pennsylvania, Ohio, and West Virginia (132 water wells). We consider groundwater redox condition in combination with two other principal environmental factors, namely topography and water type, to assess their predictive power regarding natural methane occurrence. These datasets are comprised of water quality samples collected by oil and gas operators prior to starting nearby shale gas extraction activities (i.e., "pre-drill" samples) in the respective areas. Such datasets provide a significant number of samples obtained using standard sampling and analytical methods currently utilized by operators (and likewise regulators) to evaluate baseline conditions. A better understanding of the relevant environmental factors associated with naturally highly dissolved methane concentrations in groundwater can support improved methods for distinguishing between naturally occurring methane and stray gas impacts.

Research Methods

Residential Water Wells Sampled

The dataset for this study includes water quality results from 1549 individual residential water wells, of which 1417 wells are located in northeastern Pennsylvania and 132 are located in western Pennsylvania, Ohio, and West Virginia (Figure 1). In an effort to ensure that the data were representative of natural water quality prior to drilling (i.e., pre-drill) the proposed shale gas wells, only samples from water wells located 2500 ft or more from the nearest shale gas well at the time of sampling were included in the study (consistent with the pre-drill sampling radius established by the Pennsylvania Department of Environmental Protection, PA DEP 2012). The majority of water wells in these areas are open-hole completions penetrating fractured sedimentary bedrock aquifers primarily comprised of sandstone with interbedded siltstone, limestone, shale, and/or coal. These aquifers include the Upper Devonian Catskill and Lock Haven Formations (northeastern Pennsylvania) and Permian and Pennsylvanian Aquifers of western Pennsylvania, West Virginia, and Ohio (Piper et al. 1933; Lohman 1937, 1939; Taylor 1984; Kozar and Mathes 2001; Swistock 2007; Chambers et al. 2012; Ohio EPA 2014; PaGWIS 2014). The movement of groundwater in these formations is dominated by fracture flow (Hollowell and Koester 1975; Geiser and Engelder 1983; Taylor 1984). A smaller proportion of residential water wells draw from surficial deposits of alluvium and/or glacial drift, which can be tens of meters thick in valleys in northeastern Pennsylvania (Aber 1974; Sevon et al. 1975; Inners and Fleeger 2002; Braun 2006; Chambers et al. 2012).

Sample Collection

Pre-drill groundwater samples were collected between August 2010 and April 2014 by independent environmental consultants representing two oil and gas operators. Samples were collected after the water well had been purged at a steady rate (commonly between 3 and 5 gallons per minute) for approximately 15 min or until stabilization of field parameters (i.e., conductivity, pH, and temperature) was achieved as measured by a handheld multiparameter meter (e.g., YSI flow-through cell). Water samples were collected at the closest access point to the wellhead, which included the spigot at the wellhead or before the pressure/treatment tank, the spigot at the base of the pressure tank (most common), or from the sink tap within the residence. Water quality samples were collected after reducing the flow to approximately 0.5 gallons per minute or less. Dissolved gas samples were collected by directly filling 40 mL VOA (volatile organic analyte) vials from the tap or polyethylene tubing connected to the access point. All water samples were analyzed by state-certified commercial laboratories (i.e., Eurofins Lancaster, Quantum, ALS Environmental, TestAmerica, and MicroBac Laboratories) for concentrations of dissolved methane, key redox indicator parameters (i.e., NO_3-, Mn(II), Fe(II), and SO_4^{2-}), and

the following major cations and anions: Na^+, Ca^{2+}, K^+, and Mg^{2+}, Cl- and SO_4^{2-}. In addition, alkalinity (total or bicarbonate) was analyzed in 1483 of the 1549 samples. In addition, alkalinity (total or bicarbonate) was analyzed in 1483 of the 1549 samples. At all laboratories, dissolved gas concentrations were analyzed using a GC-FID (gas chromatography-flame ionization detector) and a modified version of the U.S. EPA Method RSK-175, under which laboratory-specific modifications may include the method of sample extraction and/or calibration. (Comparison of the methane concentrations reported by different state-certified laboratories is the focus of an ongoing Marcellus Shale Coalition research project, Vitale et al. 2015.) Isotopic analyses of methane and ethane were not performed. Water samples were typically analyzed for total rather than dissolved metal concentrations (i.e., not filtered prior to acidification and analysis). Only samples with reported turbidity less than 10 NTU were included in this study. For samples satisfying this criterion, no correlation between turbidity and measured iron concentration was observed (Figures S1A and S1B).

Data Compilation

The study dataset of 1549 samples was developed from a larger pre-drill database ($>10,000$ samples) according to the following screening criteria: (1) for samples with all major cations (Ca^{2+}, Mg^{2+}, Na^+, K^+) and anions (Cl^-, SO_4^{2-}, HCO_3^-) reported, those exceeding a charge balance of $\pm 10\%$ were excluded; and (2) when multiple pre-drill samples were available from the same water well, the sample with the highest methane concentration was used for the study. For samples where methane was not detected (70% of samples), methane concentrations were assumed to be equal to one-half of the detection limit, which ranged from 0.05 to 26 µg/L. For this study, substitution of one-half of the detection limit was determined to be reasonable and convenient for purposes of graphical and statistical analyses without introducing undue bias, given that elevated methane concentrations of concern (i.e., $[CH_4] >1000$ µg/L) were up to five orders of magnitude greater than the analytical detection limits.

Classification of Topographic Location

The topographic location of each water well was defined as either "valley" or "upland" using the Topography Position Index (TPI) (Jenness 2006). The TPI classification calculates a "slope position" for each water well based on the difference between the cell elevation and the average elevation of cells in the surrounding area. A positive slope position represents a location that is higher than its surroundings (i.e., upland), while a negative slope position indicates a location that is lower than its surroundings (i.e., valley).

Classification of Water Type

Water type was initially determined using the relative concentrations in milliequivalents per liter of major cations (Ca^{2+}, Mg^{2+}, Na^+, K^+) and anions (HCO_3^-,
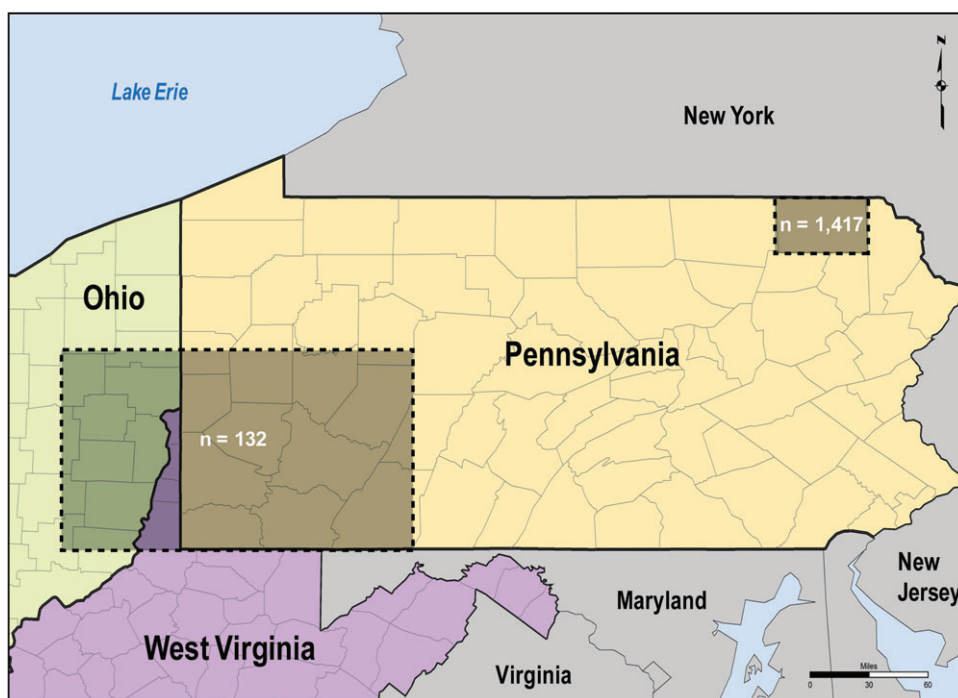
Figure 1. Map of study areas containing 1549 water wells sampled prior to drilling in Pennsylvania, Ohio, and West Virginia.

SO_4^{2-}, and Cl^-) according to the nine characteristic water types identified by Deutsch (1997) (i.e., Ca-HCO$_3$, Ca-HCO$_3$-Cl, Ca-Cl, Ca-Na-HCO$_3$, Ca-Na-HCO$_3$-Cl, Ca-Na-Cl, Na-HCO$_3$, Na-Cl, and Na-HCO$_3$-Cl). For this study, these characteristic water types were consolidated into three general water types, namely Ca-rich, Ca-Na mixed, and Na-rich water, based on the dominant cation(s) (Figure S2A and S2B). Water type was not classified based on chloride concentration (e.g., Warner et al. 2012; Darrah et al. 2014), as the correlation between methane and the sodium/calcium ratio was observed to be stronger than that between methane and chloride in this dataset (Tables S2A and S2B, Figure S3A and S3B). In addition, due to elevated Br detection limits (e.g., 200 µg/L), the Cl/Br ratio did not provide a meaningful classification for this dataset.

Classification of Groundwater Redox Condition

Redox conditions in groundwater are commonly a function of microbial processes that utilize available electron acceptors to oxidize organic molecules. The use of these electron acceptors tends to proceed sequentially in the order of the oxidation potential, where available O_2 is consumed first, followed by NO_3^-, Mn(IV), Fe(III), SO_4^{-2}, and CO_2 (Table S1A and S1B). Reduced by-products of these reactions include N_2, Mn(II), Fe(II), S^{-2}, and CH_4.

Previous redox classification frameworks (e.g., McMahon and Chapelle, 2008) have relied on measurements of the following soluble electron acceptors and reaction by-products to determine the prevailing redox process: O_2, NO_3^-, Mn(II), Fe(II), and SO_4^{-2}. In this study, a modified redox classification scheme was developed that is tailored to residential water well data

collected for pre-drill sampling. Specifically, dissolved O_2 (D.O.) concentrations were not considered to be reliable because most of the water samples were reported to be greater than 0.5 mg/L (the threshold commonly cited for anaerobic conditions), despite the presence of other water quality parameters consistent with highly reduced conditions (i.e., high Fe(II), high methane). Thus, D.O. concentrations were not considered. Given the nature of water supply well sampling, we suspect that O_2 measurements were either artificially elevated due to air entrainment associated with the use of residential water pumps or inaccurately measured by field instrumentation.

Consequently, our initial redox classification approach relied upon measurements of the following four soluble ions: NO_3^-, Mn(II), Fe(II), and SO_4^{-2}. Based on the expected consumption of electron acceptors in the order of energy yield, we initially divided our redox conditions into 10 different categories (Table S3). Within these categories, we expected samples with combined low NO_3^-, high Mn(II), high Fe(II), and low SO_4^{-2} to represent the most reducing conditions, and therefore contain the most elevated methane concentrations. Indeed, samples from northeastern Pennsylvania that met these criteria (far right category on Figure 2A) exhibited significantly elevated median and 90th percentile methane concentrations (i.e., 1200 and 30,200 µg/L, respectively). However, samples from northeastern Pennsylvania with low concentrations of all four ions (i.e., low NO_3^-, Mn(II), Fe(II), and SO_4^{-2}) also exhibited elevated median and 90th percentile methane concentrations (i.e., 5300 and 20,400 µg/L, respectively). In fact, regardless of Mn(II) or Fe(II) concentration, samples with low NO_3^- and low SO_4^{-2} concentrations exhibited similarly
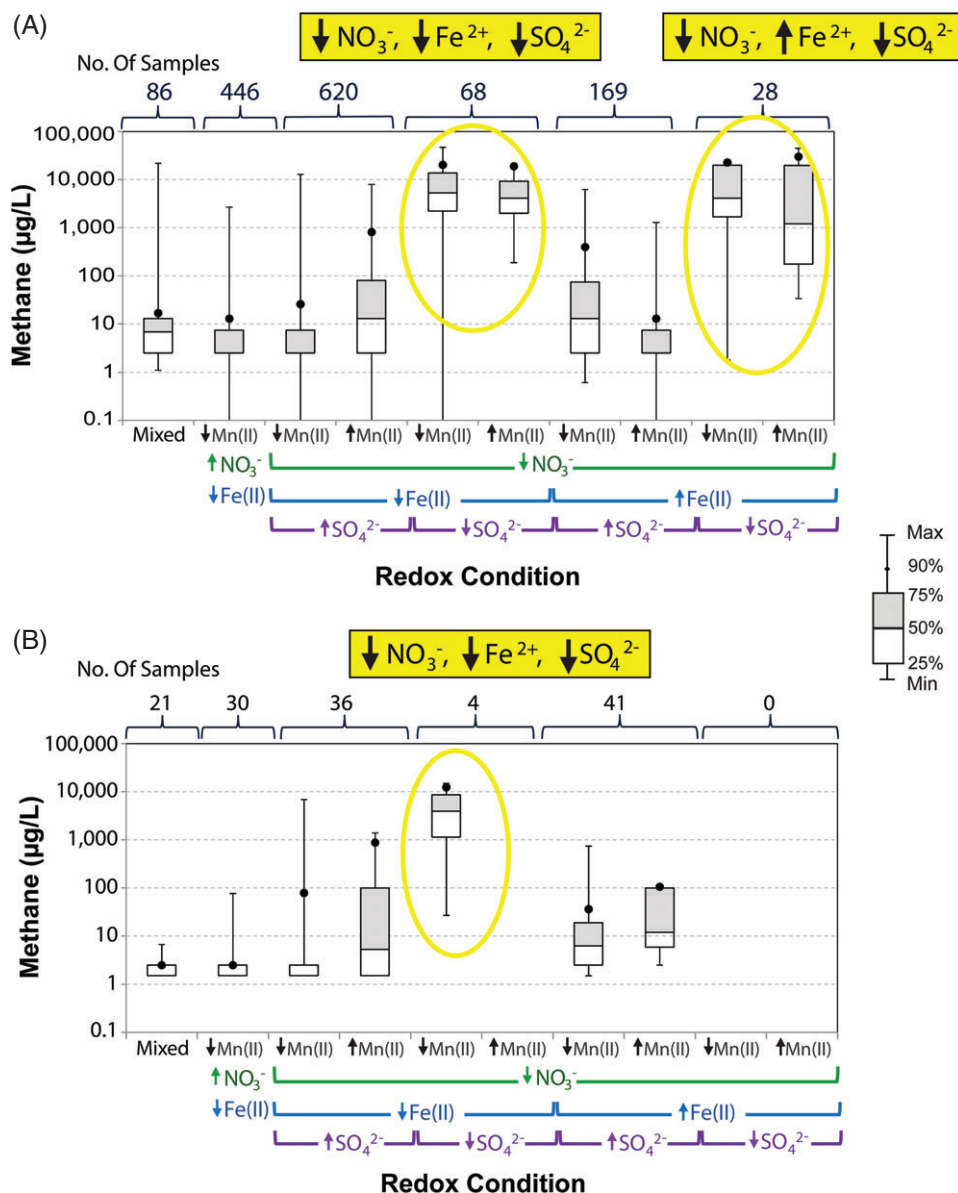
Figure 2. Distribution of methane concentrations among various redox classifications in (A) 1417 residential water well samples in northeastern Pennsylvania, and (B) 132 residential water well samples in western Pennsylvania, Ohio, and West Virginia.

elevated methane concentrations in northeastern Pennsylvania (Figure 2A) and western Pennsylvania, Ohio, and West Virginia (Figure 2B).

This finding departs from expectations of many redox schemes, wherein high Fe(II) is considered an indicator of conditions that have progressed through iron reduction into sulfate reduction or methanogenesis. Given the similarity in methane distributions between groups of samples that exhibit low NO_3^- and low SO_4^{-2}, samples exhibiting these conditions were combined into a single group (where Mn(II) and Fe(II) concentrations were not a discriminating factor). As shown in Table 1 and Figure 3A and 3B, our modified redox classification scheme is comprised of four conditions named on the basis of the observed concentrations of NO_3^- or SO_4^{-2} with respect to the threshold value: (1) High NO_3^-, (2) Low NO_3^-, (3)

Low NO_3^- & SO_4^{-2}, and (4) Mixed. "Mixed" conditions represent environments where the redox process cannot be clearly identified, potentially due to the mixing dynamics within the well, flowpath, or aquifer.

In the development of this modified classification scheme, threshold concentrations for NO_3^-, Mn(II), Fe(II), and SO_4^{2-} were consistent with the redox framework developed by McMahon and Chapelle (2008), with one exception. The majority of SO_4^{2-} measurements were reported with an analytical detection limit of 5 mg/L. Consequently, use of a 0.5 mg/L threshold value, as proposed in the McMahon and Chappelle framework, was impractical for this dataset, and the threshold level for characterization of SO_4^{2-} was adjusted to 2.5 mg/L (corresponding one-half of the prevailing analytical detection). For nondetect measurements, concentrations were

Table 1
Modified Groundwater Redox Classification System

Redox Condition	Redox Classification Parameters			
	NO_3^- as N (mg/L)	Mn(II) (mg/L)	Fe(II) (mg/L)	SO_4^{2-} (mg/L)
High NO_3^-	≥0.5	<0.05	<0.1	—
Low NO_3^-	<0.5	—	—	≥2.5
Low NO_3^- and SO_4^{2-}	<0.5	—	—	<2.5
Mixed	≥0.5	≥0.05	—	—

Note: The redox condition is not affected by the species concentration.
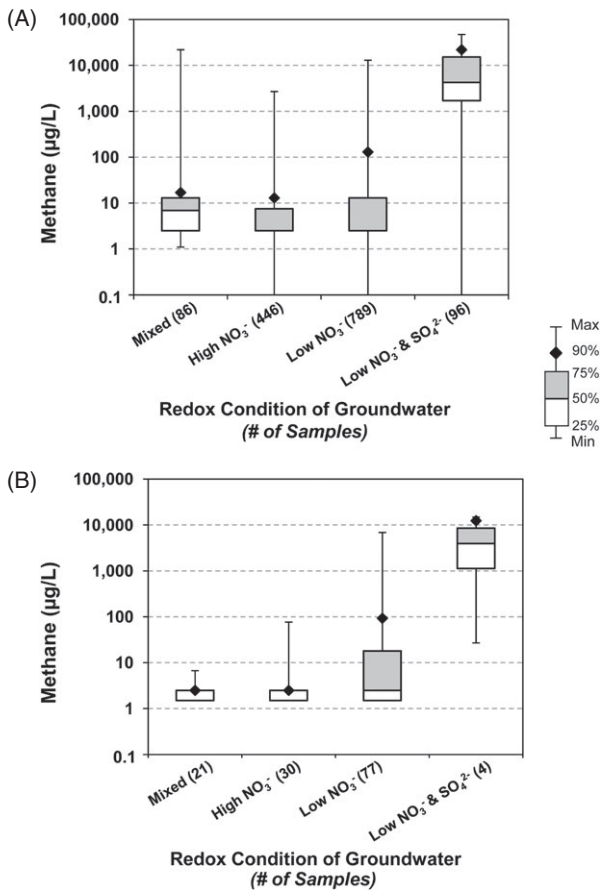

Figure 3. Elevated methane distribution associated with combined Low NO_3^- & SO_4^{2-} concentrations (i.e., more reducing conditions) in (A) 1417 residential water well samples in northeastern Pennsylvania, and (B) 132 residential water well samples in western Pennsylvania, Ohio, and West Virginia.

assumed to be equal to one-half of the detection limit. No samples had a NO_3^-, Mn(II), Fe(II), or SO_4^{2-} detection limit greater than twice the selected threshold values.

Statistical Analyses

A nonparametric Mann-Whitney U test was used to compare methane concentrations in samples exhibiting different water types and redox classifications. A multifactor analysis of variance (ANOVA) was used

to evaluate the relationship between environmental factors and methane concentration. However, because the methane data are not normally distributed, we also fit a logistic regression model to the data by dichotomizing methane values into those less than 1000 µg/L and those greater than or equal to 1000 µg/L. Likelihood ratio tests were then used to evaluate which environmental factor, or combination of factors, best fit the model (i.e., best predicted natural methane concentrations greater than 1000 µg/L).

Results and Discussion

High Methane Concentrations Are Associated with Low NO_3^- and Low SO_4^{2-}, but Not High Fe(II)

The fact that elevated methane was present in groundwater even when Fe(II) concentrations were low (<0.1 mg/L) suggests that, in the study areas, either (1) naturally occurring high methane concentrations are commonly present in certain flowpaths with low levels of iron(III)-oxide minerals, or (2) any Fe(II) generated by iron reduction in groundwater was removed from solution by sorption or precipitation with sulfides generated from sulfate reduction (Pyzik and Sommer 1981; Chapelle and Lovley 1992; Chapelle et al. 1995; He et al. 2009, 2010). Regardless of the driving mechanism, the fact that high Fe(II) concentrations are not uniquely associated with the higher methane concentrations suggests that consideration of Fe(II) as an important indicator of highly reducing conditions should be used with caution in certain hydrogeologic settings.

The clear association of methane with Low NO_3^- & SO_4^{2-} conditions (Figure 3A and 3B) suggests that in-well mixing and gas-phase transport of methane, if present, do not significantly alter the expected relationship between methane and more reducing conditions in residential water wells in the study areas. As an alternative explanation, if the wellbore is being fed by multiple flowpaths, one of which contains elevated methane, the presence of elevated methane may rapidly drive reducing conditions within the wellbore as a whole (e.g., within the timeframe in which the resident uses at least one well volume of water).

Association of Methane with Na–Rich Waters and Valley Settings

In both datasets, methane concentrations were strongly associated (p < 0.001) with Na-rich waters, where water wells with Na-rich water exhibited a significantly higher 90th percentile concentration (18,350 and 6650 μg/L for the datasets from northeastern Pennsylvania and western Pennsylvania, Ohio, and West Virginia, respectively) than water wells with Ca-rich waters (90th percentile values <26 μg/L for both datasets) (Figure S4B). Methane concentrations were also associated with valley water wells, although the difference between methane concentrations in valley and upland water wells was markedly less pronounced and primarily in the upper range of concentrations (i.e., 90th percentile methane concentration of 2240 and 108 μg/L in valley water wells for the datasets from northeastern Pennsylvania and western Pennsylvania, Ohio, and West Virginia, respectively vs. <26 μg/L in upland water wells for both datasets) (Figure S4A).

Co–Occurrence of Low NO_3^- & SO_4^{-2} Conditions, Na–Rich Waters, and Valley Setting

To evaluate whether these three environmental factors associated with naturally elevated methane concentrations simply represent common underlying relationships, we assessed the co-occurrence of different water types, redox conditions, and topographic settings within the 1549 samples.

In both datasets, Na-rich waters and Low NO_3^- & SO_4^{2-} conditions occur more commonly in valley water wells than upland water wells (Figure S5A). Specifically, Na-rich waters are present in 14% and 15% of valley water wells in datasets from northeastern Pennsylvania and western Pennsylvania, Ohio, and West Virginia, respectively, vs. 8% and 9% of upland water wells in both areas. In addition, Low NO_3^- & SO_4^{2-} conditions are present in 20% and 16% of valley wells in northeastern Pennsylvania and western Pennsylvania, Ohio, and West Virginia, respectively, vs. 2% and 0% of upland wells. Conversely, upland water wells contain a higher percentage of Ca-rich waters (81% and 80% of upland water wells in the two datasets, respectively) and High NO_3^- samples (40% and 29% of upland water wells, respectively) than valley water wells (Figure S5B). It follows that Na-rich waters also contain the highest percentage of samples exhibiting Low NO_3^- & SO_4^{2-} conditions (39% and 25% of Na-rich water samples for the two datasets). By comparison, only 1% or less of Ca-rich water samples in both datasets exhibit Low NO_3^- & SO_4^{2-} conditions (Figure S5C).

The fact that all three factors commonly co-occur suggests that they are, to a certain degree, related to the same underlying phenomenon. The co-occurrence of these factors is consistent with both the presence of confined groundwater flowpaths (which are characterized by slower moving water) and/or the mixing of freshwater with deeper saline water or brine. In either scenario, without limited access to oxygen and recharge water, the development of Highly reducing conditions may

be expected in the presence of an organic substrate. Conversely, the prevalence of High NO_3^- conditions in upland water wells is likely related to the presence of groundwater-recharge areas characterized by more rapid groundwater flow and increased transport of oxygen and nitrate (Freeze and Cherry 1979).

Na-rich waters may also be associated with either scenario, where both progressive cation exchange (whereby sodium is exchanged for calcium and magnesium in mineral exchange sites) and mixing of freshwater with saline groundwater or deep brine would enrich the sodium concentration of groundwater. The fact that valley settings contain a greater proportion of wells with both Na-rich water and Low NO_3^- & SO_4^{2-} conditions suggests that certain valley water wells are preferentially accessing groundwater with a longer residence time, potentially related to the presence of deposits of confining glacial drift and alluvium in valleys (Heisig and Scott 2013) and/or the prolonged circulation of groundwater from upland recharge zones (Molofsky et al. 2013; Siegel et al. 2015b). Additionally, thinner fresh water lenses in valleys may result in increased mixing with deeper saline or brine waters, and enhanced valley fracture networks may serve as natural migration pathways for this saline/brine water (Kresse et al. 2012; Warner et al. 2012; Heisig and Scott 2013; Molofsky et al. 2013; Darrah et al. 2014; McPhillips et al. 2014).

Although all three factors are commonly present together, it is important to note that they do not exclusively co-occur. A multifactor ANOVA analysis was conducted to assess whether these three factors (i.e., water type, topographic location, and redox condition) provide individual predictive power regarding natural methane concentration. Evaluating both datasets separately, using either the full datasets or just the samples with detectable methane concentrations, a multifactor ANOVA indicated that there was a statistically significant relationship between each factor and methane concentration (p < 0.01 for each factor). Since the methane data were not normally distributed, we also compared logistic regression models to assess whether the use of any one or two factors (e.g., water type and topographic location) provided the same predictive power for natural methane concentrations equal to or greater than 1000 μg/L as an additive model using all three factors together (i.e., water type, topographic location, and redox condition). Evaluating both datasets separately, models using all three environmental factors were found to provide a better fit to the logistic regression model describing the log-odds of reported natural methane concentrations greater than or equal to 1000 μg/L. Both analyses (multifactor ANOVA and a likelihood ratio test of logistic regression models) suggest that elevated natural methane concentrations are better predicted using all three factors than any one or two factors alone.

When considered in isolation, of the three factors, topographic location is the least predictive of methane concentrations above 1000 μg/L. Specifically, valley and upland water wells exhibited median methane concentrations lower than 10 μg/L in both datasets. The
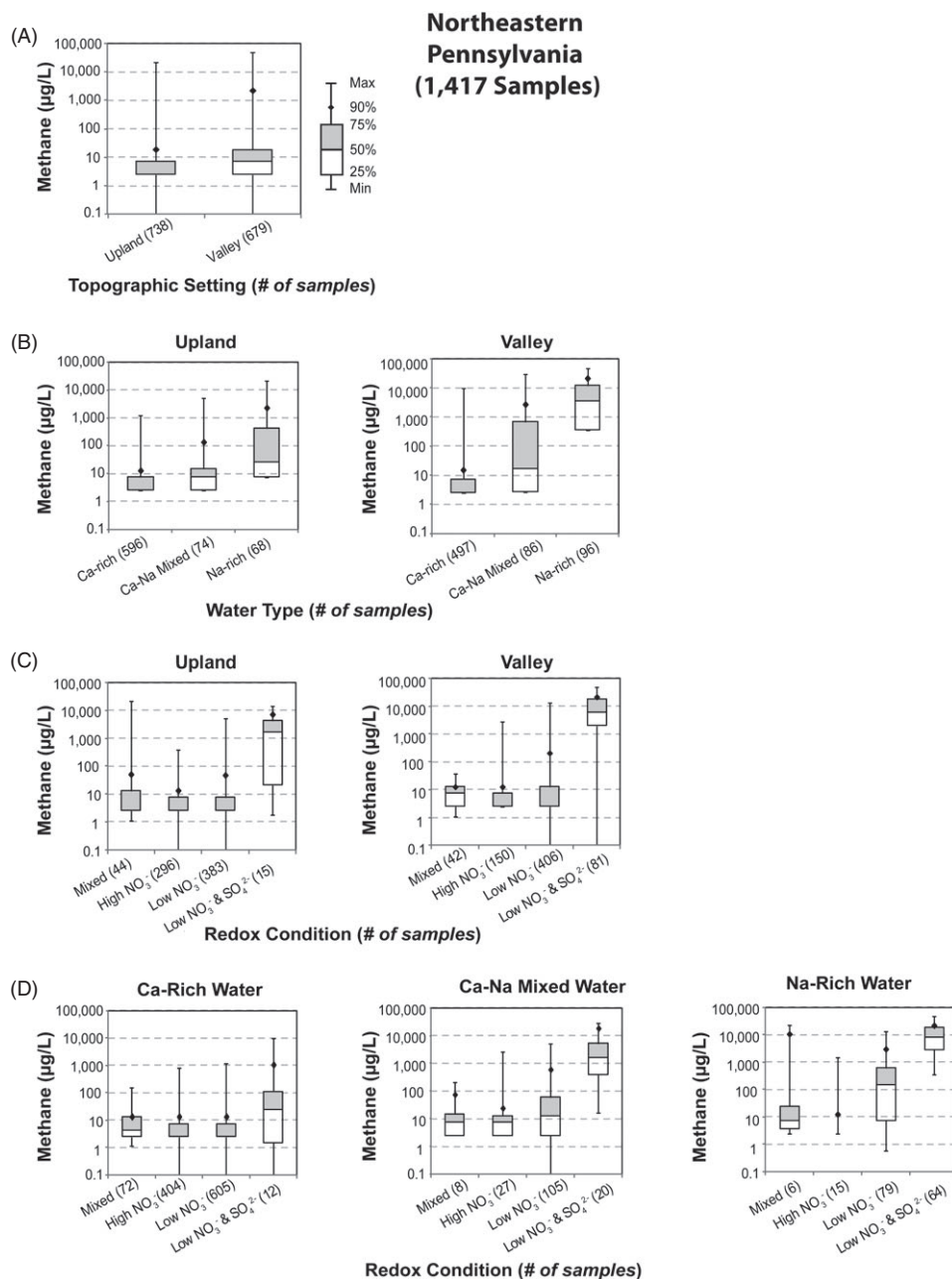
Figure 4. (A) Distribution of methane in valley vs. upland water wells. **(B)** Impact of topographic location on methane concentrations associated with each water type. **(C)** Impact of topographic location on methane concentrations associated with each redox condition. **(D)** Impact of water type on methane concentrations associated with each redox condition.

primary difference between valley and upland water wells is apparent in the upper-end methane concentrations, where the 90th percentile methane concentrations in valley water wells were 2240 and 108 µg/L for datasets from northeastern Pennsylvania vs. western Pennsylvania, Ohio, and West Virginia, respectively (Figure 4A). This suggests that knowing the topographic location of a water well only provides, at most, an approximate 10% chance of predicting methane concentrations of 2000 µg/L or more in northeastern Pennsylvania. However, in combination with information on water type or redox condition, the topographic location provides critical additional information regarding the expected occurrence of elevated

methane concentrations. This is illustrated in Figure 4B, which shows that although Na-rich waters exhibit elevated methane concentrations both in valley and upland water wells, Na-rich waters in valley water wells display the highest range of methane concentrations in both datasets (90th percentile methane = 21,500 and 6900 µg/L in the datasets from northeastern Pennsylvania and western Pennsylvania, Ohio, and West Virginia, respectively).

Similarly, Low NO_3^- & SO_4^{2-} redox conditions exhibit elevated methane concentrations both in valley and upland water wells, but display the highest methane concentrations in valley water wells (Figure 4C). Of the 107 samples from both datasets with methane
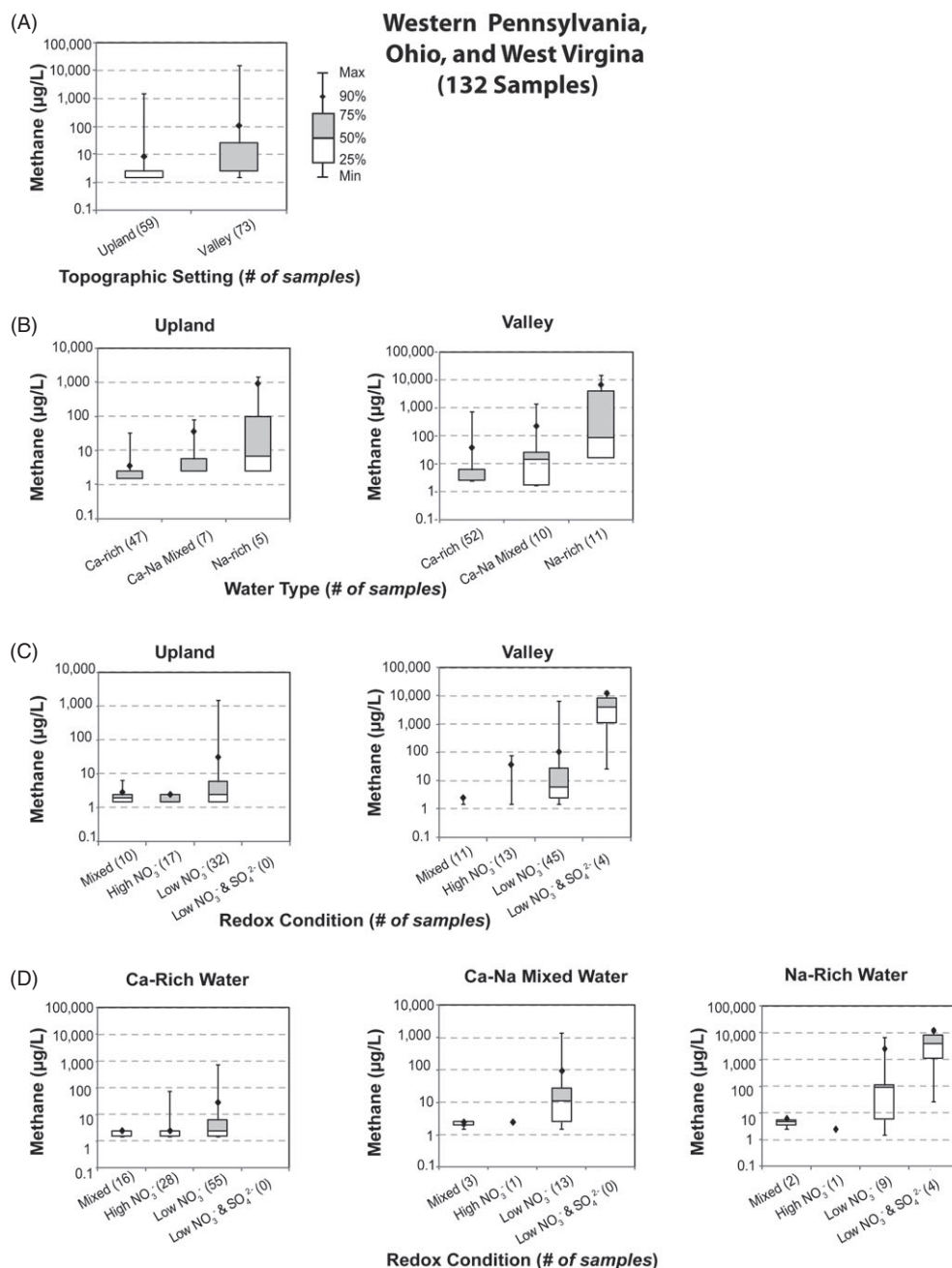
Figure 4. continued

concentrations greater than 1000 µg/L, 98% (105 of 107 samples) were associated with at least one of the three factors (i.e., Na-rich water, Low NO_3^- & SO_4^{2-} redox conditions, and valley setting), 89% (95 of 107 samples) were associated with at least two of the three factors, and 53% (57 of 107 samples) were associated with all three factors. Samples associated with all three factors exhibited a median methane concentration of 10,000 µg/L. For comparison, in the population of 1442 water well samples with methane concentrations less than 1000 µg/L, 50% (726 of 1442 samples) were associated with at least one factor, 3% (47 of 1442 samples) were associated with at least two factors, and 0.1% (2 of 1442 samples) were associated with all three factors. The median and 90th

percentile methane concentrations of samples exhibiting each natural risk factor and combinations of factors are shown in Figure 5.

Similarities Between Datasets from Northeastern Pennsylvania and Western Pennsylvania, Ohio, and West Virginia

Despite being separated by at least 200 miles and different primary aquifers, the two datasets evaluated in this study exhibit remarkably similar relationships between methane and redox conditions, water type, and topographic setting. The dataset from western Pennsylvania, Ohio, and West Virginia does, as a whole, exhibit lower methane concentrations than the dataset from northeastern

Number of Factors Present That Are Associated
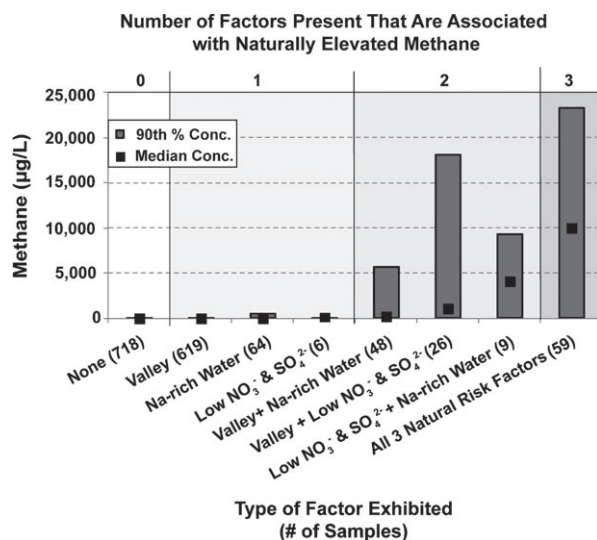with Naturally Elevated Methane

Figure 5. Median and 90th percentile methane concentration in 1549 samples from both datasets combined exhibiting no, only one, only two, or all three environmental factors associated with naturally elevated methane concentrations.

Pennsylvania. Regardless of the differences in the range of methane concentrations, the fact that both datasets demonstrate similar associations between methane and the three environmental factors considered in this study suggest that these factors may inform the natural occurrence of methane throughout the Appalachian Basin.

Implications

Application of Environmental Factors to Discerning Natural Methane Variability from Stray Gas Incidents

Characterizing a pre-drill water well sample according to topographic location, water type, and redox condition can provide valuable information to distinguish between natural variations in water quality vs. changes related to impacts from oil and gas extraction activities. A few of the ways in which this information may be applied during pre-drill and post-drill sampling are as follows:

1. Within a geographic area of interest, evaluation of a baseline dataset according to the three principal environmental factors can define the characteristic range of methane concentrations associated with each factor and combinations thereof. For example, in the dataset evaluated in our study, Ca-rich waters predominantly exhibit low methane concentrations regardless of topographic location (Figure 4B) or redox condition (Figure 4D). In fact, of the 107 samples that exhibit methane concentrations higher than $1000\,\mu g/L$, only three samples (3%) displayed a Ca-rich water type. In comparison, of 1442 water well samples with methane concentrations lower than $1000\,\mu g/L$, 1189 samples (82%) exhibited a Ca-rich water type. This suggests that naturally elevated methane concentrations are unlikely to occur in Ca-rich groundwater.

2. Highly reducing conditions may be the product of both natural methane occurrence (i.e., where both methanogenesis and the oxidation of shallow thermogenic methane can drive reducing conditions) and a stray gas release (where the oxidation of thermogenic or microbial methane that is not naturally present drives reducing conditions). Thus, the redox condition alone cannot discriminate between natural vs. anthropogenic methane sources. However, in combination with information on the water type, the redox condition can lend meaningful insight as to the driver behind changing methane concentrations. For example, if an increase in methane is accompanied by both a change in redox (e.g., from High NO_3^- to a Low NO_3^- & SO_4^2 redox condition) and a change in water type (e.g., from Ca-rich to Ca-Na Mixed waters), it could be related to differential mixing of water within the wellbore, as opposed to a stray gas release.

3. The characterization of the principal environmental factors may enable a better understanding of the duration of a stray gas impact in an individual water well. For example, a water well with Na-rich water and Low NO_3^- & SO_4 conditions, as determined by baseline measurements, is more likely to contain groundwater originating from a longer flowpath already depleted of available electron acceptors. If this water well were to experience a new influx of methane gas, the resulting elevated methane concentrations may persist for an extended period as a result of the restricted ability of the methane to advect out and/or biodegrade. In contrast, a water well with Ca-rich water and High NO_3 conditions, corresponding to shorter flowpaths and/or higher recharge rates, is more likely to experience methane that dissipates comparatively quickly.

4. The three principal environmental factors identified as predictive of natural methane concentrations in

Table 2
Recommended Data for Characterizing Water Source Samples According to the Three Key Environmental Factors

Environmental Factor	Data Required
Topographic location	Water source coordinates
Redox condition	NO_3^-, [1]$Fe(II)$, [1]$Mn(II)$, SO_4^{2-}
Water type	Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, SO_4^{2-}, Cl^-

[1]Fe(II) and Mn(II) concentrations are not utilized in modified redox classification scheme presented in this study. Nevertheless, they still may provide valuable information if significant changes in concentrations are observed from pre-drill to post-drill sampling.

this study (topography, water type, and redox condition) can be easily determined during pre-drill and post-drill sampling through simple measurements of (1) water source coordinates (valley vs. upland) and (2) the following 10 common water quality analytes that provide information on groundwater origin, dominant geochemical processes, and electroneutrality balance for data quality assurance: Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, SO_4^{2-}, Cl^-, NO_3^-, $Fe(II)$, and $Mn(II)$ (where filtration is recommended for metals analysis in the presence of elevated turbidity). Based upon this short list of analytes (Table 2), the water type can be determined and the redox condition characterized using the modified system presented in our paper. Baseline monitoring programs for shale gas operations often consist of a much longer list of water quality parameters; however, the value of these additional parameters for discriminating between naturally occurring methane and stray gas impacts has not been clearly demonstrated. Our research suggests that focusing on the three environmental factors discussed in this paper may provide a simple and cost-effective method for assessing the likelihood that elevated methane concentrations in a water well are naturally occurring vs. related to stray gas migration.

Acknowledgments

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Section S1. Sequential Consumption of Electron Acceptors in Aquifer
Section S2. Evaluation of Relationship Between Iron and Turbidity
Section S3. Classification of Water Type
Section S4. Statistical Evaluation of Relationships Between Methane and Inorganic Parameters Utilized to Determine Water Type
Section S5. Redox Classification
Section S6. Comparison of Data from Northeastern Pennsylvania to that from Western Pennsylvania, Ohio, and West Virginia
Section S7. Co-occurrence of Environmental Factors
Figure S1. (A) Iron concentration vs. turbidity; and (B) methane concentration vs. turbidity for the combined dataset of 1549 samples from northeastern Pennsylvania, western Pennsylvania, Ohio, and West Virginia.
Figure S2. Distribution of methane in water types classified by (A) major cation and anion; and (B) major cation for the combined dataset samples from northeastern Pennsylvania, western Pennsylvania, Ohio, and West Virginia.
Figure S3. Relationship between methane and (A) sodium/(sodium + calcium) ratio, in meq/L; (B) chloride, in mg/L, for the combined dataset of 1549 samples from northeastern Pennsylvania, Western Pennsylvania, Ohio, and West Virginia.
Figure S4. Comparison of relationship between methane and environmental factors in datasets from northeastern Pennsylvania vs. western Pennsylvania, Ohio, and West Virginia: (A) methane distribution within valley vs. upland water wells; (B) methane distribution within different water types; (C) methane distribution within different redox conditions.
Figure S5. (A) Occurrence of different water types within valley vs. upland water wells; (B) occurrence of different redox conditions within valley vs. upland water wells; (C) occurrence of different redox conditions within different water types.
Table S1. Redox reactions that consume organic molecules in an aquifer.
Table S2. Results of Spearman's rank correlation tests.
Table S3. Ten initial redox conditions.

References

Aber, J.S. 1974. Upland glacial stratigraphy in the Binghamton-Montrose region of New York and Pennsylvania. M.Sc. thesis, University of Kansas, Lawrence, Kansas.

Baldassare, F.J., M. McCaffrey, and J.A. Harper. 2014. Geochemical context for stray gas investigations in the N. Appalachian Basin: Implications of analyses of natural gases from Neogene-through-Devonian-age strata. *American Association of Petroleum Geologists Bulletin* 98, no. 2: 341–372.

Baldassare, F.J., and C.D. Laughrey. 1997. Identifying the sources of stray methane by using geochemical and isotopic fingerprinting. *Environmental Geosciences* 4, no. 2: 85–94.

Blackman, E.C. 1873. *History of Susquehanna County, Pennsylvania*. Philadelphia, Pennsylvania: Claxton, Remsen & Haffelfinger.

Braun, D.D. 2006. Surficial geology of the springville 7.5-minute quadrangle, Susquehanna and Wyoming Counties, Pennsylvania. The Pennsylvania Geological Survey, 4th Series, vol. 13 and oversized map.

Breen, K.J., K. Revesz, F.J. Bladasare, and S.D. McAuley. 2007. Natural gases in ground water near Tioga Junction, Tioga County, North-Central Pennsylvania—Occurrence and use of isotopes to determine origins, 2005. U.S. Geological Survey Scientific Investigations Report Series 2007-5085, 65.

Chambers, D.B, M.D. Kozar, J.S. White, and K.S. Paybins. 2012. Groundwater quality in West Virginia, 1993-2008. U.S. Geological Survey Scientific Investigations Report 2012-5186. http://pubs.usgs.gov/sir/2012/5186/ (accessed April 2015).

Chapelle, F.H., P.B. McMahon, N.M. Dubrovsky, R.F. Fujii, E.T. Oaksford, and D.A. Vrobleshy. 1995. Deducing the distribution of terminal electron-accepting processes in hydrologically diverse groundwater systems. *Water Resources Research* 31: 359–371.

Chapelle, F.H., and D.R. Lovley. 1992. Competitive exclusion of sulfate reduction by Fe(III)-reducing bacteria: A mechanism for producing discrete zones of high-iron ground water. *Groundwater* 30, no. 1: 29–36.

Clayton, C. 1991. Carbon isotope fractionation during natural gas generation from kerogen. *Marine and Petroleum Geology* 8: 232–240.

Coleman, D.D. 1994. *Advances in the Use of Geochemical Fingerprinting for Gas Identification*. Champaign, Illinois: Isotech Laboratories.

Coleman, N.P., D. McElreath, and B. Smith. 2012. Short-term natural variability in groundwater quality from domestic water wells in Northeastern Pennsylvania. In *Battelle Ninth International Conference on Remediation of Chlorinated and Recalcitrant Compounds*, Monterey, California, May 21–24, 2012.

Darrah, T.H., A. Vengosh, R.B. Jackson, N.R. Warner, and R.J. Poreda. 2014. Noble gases identify the mechanisms of fugitive gas contamination in drinking water wells overlying the Marcellus and Barnett Shales. *Proceedings of the National Academy of Sciences USA* 111, no. 39: 14076–14081.

Deutsch, W.J. 1997. *Groundwater Geochemistry: Fundamentals and Applications to Contamination*. Boca Raton, Florida: CRC Press.

Freeze, R.A., and J.A. Cherry. 1979. *Groundwater*. Englewood Cliffs, New Jersey: Prentice-Hall.

Fuex, A.N. 1977. The use of stable carbon isotopes in hydrocarbon exploration. *Journal of Geochemical Exploration* 7: 155–188.

Geiser, P., and T. Engelder. 1983. The distribution of layer parallel shortening fabrics in the Appalachian Foreland of New York and Pennsylvania: Evidence for two noncoaxial phases of the Alleghenian orogeny. *Geological Society of America Memoir* 158: 161–175.

Gorody, A. 2012. Factors affecting the variability of stray gas concentration and composition in groundwater. *Environmental Geosciences* 19, no. 1: 17–31.

He, Y.T., J.T. Wilson, and R.T. Wilkin. 2010. Impact of iron sulfide transformation on trichloroethylene degradation. *Geochimica et Cosmochimica Acta* 74: 2025–2029.

He, Y.T., J.T. Wilson, and R.T. Wilkin. 2009. Transformation of reactive iron minerals in a permeable reactive barrier (biowall) used to treat TCE in groundwater. *Environmental Science & Technology* 42, no. 17: 6690–6696.

Heisig, P., and T.M. Scott. 2013. Occurrence of methane in groundwater of south-central New York State, 2012—Systematic evaluation of a glaciated region by hydrogeologic setting. U.S. Geological Survey Scientific Investigations Report 2013–5190. http://pubs.usgs.gov/fs/2013/3118/ (accessed April 2015).

Hollowell, J.R., and H.E. Koester. 1975. Ground-water resources of Lackawanna County, Pennsylvania. Pennsylvania Geologic Survey, 4th Series, Water Resource Report 41, 106.

Inners, J.D., and G.M. Fleeger. 2002. From Tunkhannock to Starrucca: Bluestone, glacial lakes, and great bridges in the "Endless Mountains" of northeastern Pennsylvania: Guidebook. In *67th Annual Field Conference of Pennsylvania Geologists*, Tunkhannock, Pennsylvania.

Jenness, J. 2006. Topographic Position Index (tpi_jen.avx) extension for ArcView 3.x, v. 1.3a. Jenness Enterprises. http://www.jennessent.com/arcview/tpi.htm (accessed April 2015).

Kappel, W.M., and E.A. Nystrom. 2012. Dissolved methane in New York groundwater, 1999—2011: U.S. Geological Survey Open-File Report 2012-1162, 6, Reston, Virginia: USGS. http://pubs.usgs.gov/of/2012/1162 (accessed April 2015).

Kelly, W.R., and G. Matisoff. 1985. The effects of a gas well blow out on groundwater chemistry. *Environmental Geology and Water Sciences* 7, no. 4: 205–213.

Kozar, M.D., and M.V. Mathes. 2001. Aquifer-characteristics data for West Virginia. USGS Water-Resources Investigations Report 01-4036. http://pubs.er.usgs.gov/usgspubs/wri/wri014036 (accessed April 2015).

Kresse, T.M., N.R. Warner, P.D. Hays, A. Down, A. Vengosh, and R. Jackson. 2012. Shallow groundwater quality and geochemistry in the Fayetteville shale gas-production area, North-Central Arkansas, 2011. USGS Scientific Investigations Report 2012-5273, U.S. Geological Survey, 31, Reston, Virginia: USGS. http://pubs.usgs.gov/sir/2012/5273/ (accessed April 2015).

Lohman, S.W. 1939. Groundwater in North-Central Pennsylvania: Commonwealth of Pennsylvania. *Topographic and Geologic Survey Bulletin* W4: 312.

Lohman, S.W. 1937. Ground water in northeastern Pennsylvania. Progress Report W4, Pennsylvania Geological Survey, 4th Series, 312.

Mathes, M.V., and J.S. White. 2006. Methane in West Virginia groundwater. United States Geological Survey (USGS) Factsheet 2006–3011. http://pubs.usgs.gov/fs/2006/3011/pdf/Factsheet2006_3011.pdf (accessed April 2015).

McMahon, P.B., and F.H. Chapelle. 2008. Redox processes and water quality of selected principal aquifer systems. *Groundwater* 46, no. 2: 259–271.

McPhillips, L.E., A.E. Creamer, B.G. Rahm, and M.T. Walter. 2014. Assessing dissolved methane patterns in central New York groundwater. *Journal of Hydrology: Regional Studies* 1: 57–73.

Molofsky, L.J., J.A. Connor, A.S. Wylie, T. Wagner, and S.K. Farhat. 2013. Evaluation of methane sources in groundwater in Northeastern Pennsylvania. *Groundwater* 51, no. 3: 333–349.

Ohio EPA. 2014. Major aquifers and ground water quality in Ohio. Fact sheet: Division of drinking and ground waters, February 2014. http://epa.ohio.gov/Portals/28/documents/gwqcp/MajorAquifers-GWQinOhio.pdf (accessed April 2015).

Pennsylvania Department of Environmental Protection (PA DEP). 2012. Act 13 of 2012, House Bill No. 1950. http://www.ctbpls.com/www/PA/11R/PDF/PA11RHB0 1950CC1.pdf (accessed April 2015).

Pennsylvania Ground Water Information System (PaGWIS). 2014. Microsoft access relational database containing data for wells, springs, and ground water quality throughout Pennsylvania. http://www.dcnr.state.pa.us/topogeo/groundwater/PaGWIS/help.aspx (accessed August 2014).

Perry, E.A., R. Bothun, B. Smith, and M. Hollingsworth. 2012. The occurrence of methane in shallow groundwater from extensive pre-drill sampling. Presented at the Groundwater

Protection Council (GWPC) 2012 Stray Gas Incidence and Response Forum, Cleveland, Ohio, July 24–26, 2012.

Piper, A.M., M.D. Foster, and C.S. Howard. 1933. Ground water in southwestern Pennsylvania. Pennsylvania Geological Survey, 4th Series, 406.

Pyzik, A.J., and S.E. Sommer. 1981. Sedimentary iron monosulfides: Kinetics and mechanisms of formation. *Geochimica et Cosmochimica Acta* 45: 687–698.

Schoell, M. 1990. The hydrogen and carbon isotopic composition of methane from natural gases of various origins. *Geochemica et Cosmochemica Acta* 44: 649–661.

Sevon, W.D., G.H. Crowl, and T.M. Berg. 1975. The late Wisconsinan drift border in Northeastern Pennsylvania. 40th Annual Field Conference of Pennsylvania Geologists, Pennsylvania Bureau of Topographic and Geologic Survey, Harrisburg, Pennsylvania, 108.

Siegel, D., N.A. Azzolina, B.J. Smith, E.A. Perry, and R.L. Bothun. 2015a. Methane concentrations in water wells unrelated to proximity to existing oil and gas wells in Northeastern Pennsylvania. *Environmental Science and Technology* 49, no. 7: 4106–4112.

Siegel, D., B.J. Smith, E.A. Perry, R.L. Bothun, and M. Hollingsworth. 2015b. Pre-drilling water-quality data of groundwater prior to shale gas drilling in the Appalachian Basin: Analysis of the Chesapeake Energy Corporation dataset. *Applied Geochemistry* 63: 37–57.

Sloto, R.A. 2013. Baseline groundwater quality in southern Sullivan County, Pennsylvania, 2012. U.S. Geological Survey Scientific Investigations Report 2013–5085, 34. http://pubs.usgs.gov/sir/2013/5085/ (accessed April 2015).

Soren, J. 1963. The ground-water resources of Sullivan County, New York: U.S. Geological Survey Bulletin GW-46, prepared in cooperation with the New York State Water Resources Commission, 58.

van Stempvoort, D.V., H. Maathuis, E. Jaworski, B. Mayer, and K. Rich. 2005. Oxidation of fugitive methane in ground water linked to bacterial sulfate reduction. *Ground Water* 43, no. 2: 187–199.

Susquehanna Historical Society Website. 2010. www.susqcohistsoc.org (accessed July 19, 2010).

Swistock, B. A. 2007. *Quick Guide to Groundwater in Pennsylvania*. State College, Pennsylvania: Pennsylvania State University. http://pubs.cas.psu.edu/FreePubs/pdfs/uh183.pdf (accessed April 2015).

Taylor, L.E. 1984. Groundwater resources of the upper Susquehanna River Basin, Pennsylvania. Pennsylvania Geological Survey, 4th Series, Water Resources Report, vol. 58, 136.

Vitale, R., D. Yost, J. Smelko, L. Anderson, D. Gratson, S. Brower, and L. Work. 2015. Marcellus shale coalition dissolved methane study. Presented at National Environmental Monitoring Conference (NEMC), Hyatt Regency, Chicago, Illinois, July 13–15, 2015.

Warner, N.R., R.B. Jackson, T.H. Darrah, S.G. Osborn, A. Down, K. Zhao, A. White, and A. Vengosh. 2012. Geochemical evidence for possible natural migration of Marcellus Formation brine to shallow aquifers in Pennsylania. *Proceedings of the National Academy of Sciences USA* 109, no. 30: 11961–11966.

Weidemeier, T.H., H.S. Rifai, C.J. Newell, and J.W. Wilson. 1999. *Natural Attenuation of Fuels and Chlorinated Solvents*. New York: John Wiley & Sons.

Williams, J.A. 2010. Evaluation of well logs for determining the presence of freshwater, saltwater, and gas above the Marcellus Shale in Chemung, Tioga, and Broome Counties, New York. U.S. Geological Survey, Scientific Investigations Report 2010–5224.

Authors' Note

The authors do not have any conflicts of interest.