



# Pre-drilling water-quality data of groundwater prior to shale gas drilling in the Appalachian Basin: Analysis of the Chesapeake Energy Corporation dataset



D.I. Siegel<sup>a,\*</sup>, B. Smith<sup>b</sup>, E. Perry<sup>c</sup>, R. Bothun<sup>d</sup>, M. Hollingsworth<sup>e</sup>

<sup>a</sup> Earth Sciences, Heroy Geology Laboratory, Syracuse University, Syracuse, NY 13244, United States

<sup>b</sup> Enviro Clean Products and Services, 11717 North Morgan Road, P.O. Box 721090, Yukon, OK 73172-1090, United States

<sup>c</sup> AECOM Technology Corporation, 250 Apollo Drive, Chelmsford, MA 01824, United States

<sup>d</sup> AECOM Technology Corporation, 1601 Prospect Parkway, Fort Collins, CO 80525, United States

<sup>e</sup> Chesapeake Energy Corporation, Oklahoma City, OK 73154-0496, United States

## ARTICLE INFO

### Article history:

Received 4 October 2014

Revised 26 June 2015

Accepted 27 June 2015

Available online 2 July 2015

## ABSTRACT

Unconventional natural gas production in the Marcellus and Utica formations of the Northeastern United States raises concerns about potential impacts to shallow groundwater. We examined and interpreted 13,040 analyses from pre-drilling groundwater samples from domestic water wells in northeastern (NE) Pennsylvania and 8004 samples from water wells in the “Western Area” which includes southwest Pennsylvania, eastern Ohio, and north-central West Virginia. These samples were acquired on behalf of Chesapeake Energy Corporation as part of its local pre-drilling water supply monitoring program. We evaluated concentrations of major ions and metals relative to federal drinking-water-quality standards upon which regulatory decisions are often based. Chesapeake's dataset, the most comprehensive for these areas, shows that exceedance of at least one water-quality standard occurs in 63% of water well samples in NE Pennsylvania and 87% in the Western Area. In NE Pennsylvania, 10% of the samples exceeded one or more of the United States Environmental Protection Agency's (USEPA) primary maximum contaminant levels (MCLs) for drinking-water supplies, 46.1% of the samples exceeded one or more of USEPA secondary maximum contaminant levels (SMCLs), and another 7% exceeded one or more of USEPA health advisory or regional screening levels for tap water.

In the Western Area 8% of samples exceeded one or more MCLs, 65% exceeded one or more SMCLs, and 15% exceeded one or more health advisory or regional screening levels for tap water. Chesapeake's dataset, orders of magnitude larger than any in previously published literature, shows that water-quality exceedances relate to factors such as: as where the sample occurs within the groundwater flow system, the natural groundwater chemical type (hydrochemical facies), the geologic unit producing the water, and/or the topographic position (valley versus upland). Our comparison of these results to historical groundwater data from NE Pennsylvania, which pre-dates most unconventional shale gas development, shows that the recent pre-drilling geochemical data is similar to historical data. We see no broad changes in variability of chemical quality in this large dataset to suggest any unusual salinization caused by possible release of produced waters from oil and gas operations, even after thousands of gas wells have been drilled among tens of thousands of domestic wells within the two areas studied. Our evaluation also agrees with early researchers such as Piper (1933) and Lohman (1939, 1937) who found that the saline waters in both areas underlie fresher groundwater. The saline water is naturally-occurring connate brine or salt water which has not been flushed by circulating meteoric water; rather than vertical migration of salt water from deep strata such as the Marcellus shale as suggested by Warner et al. (2012). Elevated metals concentrations, particularly iron and manganese, partly relate to sample turbidity; dissolved metals would provide a more accurate measurement of metals in shallow groundwater than does the total metals analysis typically required by regulations.

© 2015 Elsevier Ltd. All rights reserved.

\* Corresponding author.

E-mail address: [disiegel@syr.edu](mailto:disiegel@syr.edu) (D.I. Siegel).

## 1. Introduction

The recent (post-2005) drilling and production of “unconventional” natural gas from the Marcellus and Utica Shale formations (Devonian and Ordovician age, Appalachian Basin, eastern North American) sparked broad public concern that shallow groundwater may be affected by recent drilling or production activities (e.g. Vidic et al., 2013; Kramer, 2011). Natural gas and oil produced from tight shale, coal beds, and tight sands using horizontal drilling and/or hydraulic fracturing techniques are termed “Unconventional”. State regulatory agencies responded to this concern by generating regulatory frameworks designed to assess pre-drilling (baseline) water supply quality to identify possible changes in solute concentrations related to hydrocarbon extraction. Conventional oil and gas development, coal mining, agricultural development, septic effluent discharges, road salting, commercial development, and industrial development all have occurred in Appalachia for over 100 years; therefore this region cannot be viewed as “pristine” in the strictest sense. Without a proper understanding of pre-existing groundwater geochemical variability, investigators may incorrectly conclude that unconventional oil and gas development has altered shallow groundwater quality when it has not (i.e. a false positive).

In this paper, we present interpretation of the inorganic chemical composition of over 21,000 samples of groundwater collected by third party contractors from individual domestic or stock water-supply wells before Chesapeake Energy Corporation (Chesapeake) drilled nearby Marcellus and Utica shale oil/gas wells in Pennsylvania, West Virginia, and Ohio. In our analysis, we compare this pre-drilling data to the combination of federal water-quality standards and guidelines, and evaluate whether geology, topographic position, groundwater circulation patterns, groundwater chemical types (hydrochemical facies), and sample turbidity relate to water quality.

Chesapeake's dataset is unparalleled in size and is orders of magnitude larger than prior studies (e.g. Molofsky et al., 2013, 2011; Warner et al., 2012; Boyer et al., 2012). In northeast (NE) Pennsylvania we interpreted 13,040 groundwater sample analyses from 12,844 domestic/stock water wells and in southwest (SW) Pennsylvania, eastern (E) Ohio, and north-central (NC) West Virginia (collectively termed the “Western Area”), we interpreted 8004 sample analyses from 7983 water wells. Combined, these datasets provide the most comprehensive information on general chemistry and trace elemental composition of shallow groundwater in the Appalachian Basin to date.

## 2. Background

Unconventional and conventional oil and natural gas production in the Appalachian Basin commonly extract naturally occurring saline waters from oil and gas source beds and these produced saline waters usually exceed drinking-water standards for trace metals, major solutes, and dissolved solids (e.g. Chapman et al., 2012; Siegel et al., 1991). However, trace metals, major solutes, and dissolved solids also naturally exceed regulatory concentrations in many shallow groundwater aquifers in the Appalachian Basin (Ayotte et al., 2011; DeSimone, 2008). Some of these exceedances occur because shale and other rocks in the Appalachian Basin naturally contain potentially soluble minerals incorporating trace metals (e.g. iron, manganese, strontium, and barium) that are common in deep formation fluids (e.g. Siegel et al., 1987). Road de-icing and septic-system effluent can also contribute salinity to potable groundwater. Saline connate water occurs naturally at the base of all shallow freshwater flow systems in both areas (e.g. Williams et al., 1998; Feth, 1970, 1981; Poth,

1963; Lohman, 1939, 1937; Piper, 1933). Exceedances of drinking-water standards in water wells of the Appalachian region have occurred long before unconventional oil and gas development began (e.g. Gross and Low, 2012; Chambers et al., 2012; Low and Galeone, 2006; Williams et al., 1998; Stoner et al., 1987; Taylor, 1984; Matisoff et al., 1982).

Several recent studies in the Appalachian Basin addressed possible changes in potable groundwater chemical conditions prior to and after unconventional shale gas development. Warner et al. (2012) concluded, from variations in salinity, that deep Marcellus shale water may have naturally migrated upwards in the geologic past to shallow aquifers in NE Pennsylvania; a finding challenged by Engelder (2012). In contrast, Boyer et al. (2012) and Molofsky et al. (2013, 2011) found no evidence that methane or dissolved ions are systematically related to oil and gas operations.

Historical studies (e.g. Razem and Sedam, 1985; Poth, 1973a,b; Poth, 1973c, 1963; Carswell and Bennett, 1963; Lohman, 1939, 1937; Piper, 1933) provide detailed discussions on the natural occurrence of groundwater chemistry and connate brines found at shallow depths throughout Appalachia (including NE Pennsylvania and the Western Area). For example, Lohman (1939) noted that in north-central Pennsylvania: “Most of the waters that contained more than 100 parts of sodium were from the Chemung formation [now referred to as the Lock Haven Formation] or from formations or deposits overlying the Chemung, and probably represent diluted connate water, that is, sea water that became trapped in the marine sediments at the time of deposition and subsequently diluted by meteoric water.” Fresh groundwater of meteoric origin flushes the salt water from the upper-most rocks, but only partially flushed it from deeper rocks, or from strata that have little groundwater circulation (such as in valleys) which Poth (1963) noted is where deep groundwater moves upward.

Eighty years ago, Piper (1933) reported circulating groundwater in bedrock of SW Pennsylvania evolves from young calcium-magnesium bicarbonate ( $\text{Ca-Mg-HCO}_3$ ) water type in the uppermost groundwater, to older sodium-bicarbonate ( $\text{Na-HCO}_3$ ) type at intermediate depths, and still older water deeper (and often in valleys) mixing with residual connate sodium-chloride ( $\text{Na-Cl}$ ) brines. Piper (1933) also found that the presence of coal and carbonate minerals in rocks of Appalachia plays an important role in the occurrence of iron, sulfate, and hardness in groundwater.

However, there has not yet been a groundwater study that determines if chemical quality in groundwater prior to unconventional gas exploitation has been systematically maintained. To this end, we interpret the largest and densest dataset on groundwater quality available for Appalachia: 21,044 analyses from 20,827 domestic groundwater wells in Pennsylvania, Ohio, and West Virginia, orders of magnitude larger than prior studies. These samples were collected as part of Chesapeake's pre-drilling sampling program. Because many wells were nearby and in place prior to any sampling for a new unconventional gas well, this pre-drilling database also serves as a means to explore if prior drilling has modified natural groundwater quality.

Our study and many of the more recently referenced studies lack well-log data for the water wells sampled. This inhibits certain evaluations, but the overall size of the dataset provides significant information about general water-quality patterns. Reliance on previous work (e.g. Razem and Sedam, 1985; Poth, 1973a,b,c, 1963; Carswell and Bennett, 1963; Lohman, 1939, 1937; Piper, 1933) that described groundwater flow system conceptualization helped put these data into proper context. The United States Geological Survey (USGS) is currently conducting a groundwater availability study of the Appalachian Plateaus (USGS, 2014) where Pennsylvanian- and Mississippian-age groundwater flow systems are being studied.



### 3. Study area and general hydrogeology

Chesapeake's contractors collected and analyzed pre-drilling samples in NE Pennsylvania (Fig. 1) and in a Western Area, covering SW Pennsylvania, NC West Virginia, and E Ohio (Fig. 2). We evaluated these two geographic areas separately since the geology, physiography, and hydrogeology of the two regions differ. Unconventional gas in the Marcellus Formation is being developed in NE Pennsylvania, whereas unconventional gas and hydrocarbon liquids are being developed in the Western Area from both the Utica and Marcellus shales.

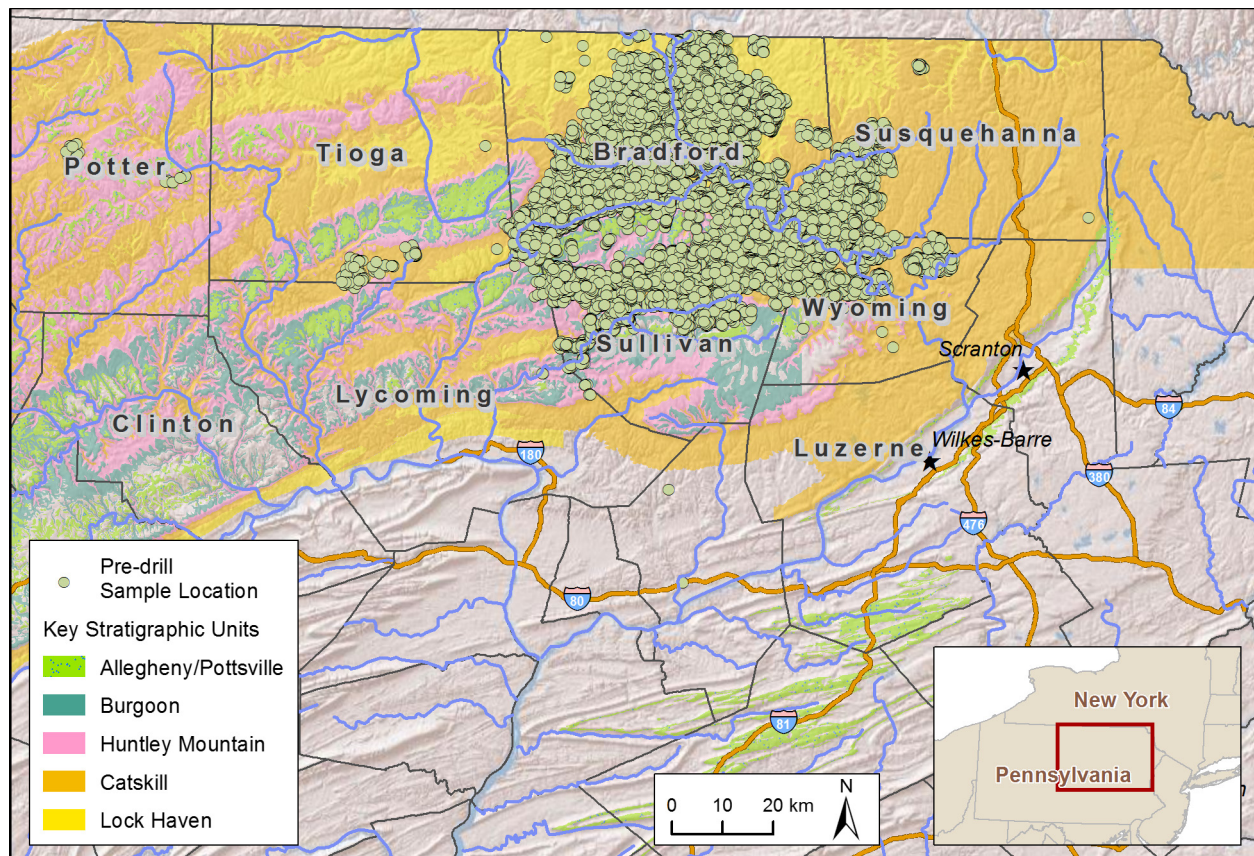
In NE Pennsylvania (Fig. 1) the sampling density averages 1 sample per 1.5 square kilometers ( $\text{km}^2$ ) and approximately 98% of samples were collected from Bradford, Sullivan, Susquehanna, and Wyoming Counties; in the Western Area (Fig. 2) sampling density averages approximately 1 sample per 4.7  $\text{km}^2$ .

A generalized stratigraphic column for NE Pennsylvania and the Western Area is provided in Fig. 3. The general hydrogeologic settings of both study areas (Razem and Sedam, 1985; Poth, 1973a,b,c, 1963; Carswell and Bennett, 1963; Lohman, 1939, 1937; Piper, 1933) have been known for many years as well as the general geochemical variability of shallow groundwater. Potable groundwater in NE Pennsylvania generally occurs in the upper 250 meters (m) of the subsurface, and overlies a freshwater-saline interface approximately 250 m deep under uplands and approximately 60 m deep beneath valleys (Williams, 2010). In NE Pennsylvania, relief is greater and regional flow systems are mostly found in valley bottoms where groundwater flow is controlled by regional dip of formations, topography, and hydrogeological boundaries coincident with large river valleys (Carswell and Bennett, 1963). In NE

Pennsylvania, potable groundwater comes from fractured sandstones, siltstones, and shale of the Catskill and Lock Haven formations of Upper Devonian age (Figs. 1 and 3), alluvial deposits, and glacial outwash. Approximately 1500–3000 m of underlying rock separates the Marcellus and Utica shales producing commercial gas from these drinking-water aquifers in our study areas.

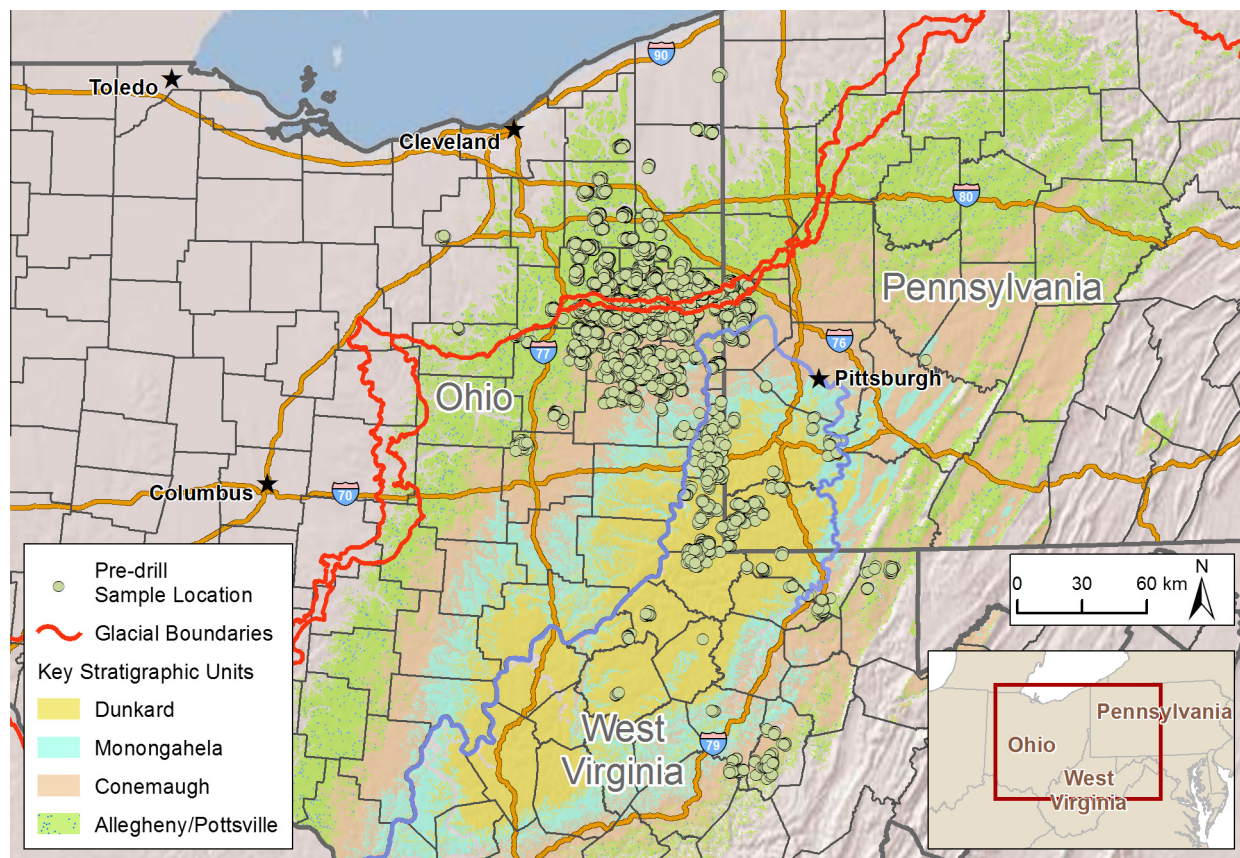
The Catskill and underlying Lock Haven formations consist of deltaic and marine shale, siltstone, sandstone and conglomerate interbedded with carbonaceous sediments and sometimes coal (e.g. Wilson, 2014; Castle, 2000; Taylor, 1984; Piotrowski and Harper, 1979). Precipitation recharges shallow groundwater systems through vertical shallow fractures and unconsolidated materials, and groundwater naturally discharges to streams in intervening valleys or as side-slope springs (e.g. Williams et al., 1998; Seaber et al., 1988).

In the Western Area, potable groundwater comes from Mississippian- and Pennsylvanian-age fractured rocks of the Allegheny-Pottsville formations and the Conemaugh and Dunkard groups (Fig. 2). These fractured rocks consist mostly of fluvial-deltaic non-marine shale, siltstone, sandstone, and limestone interbedded with commercially minable coal (Stoner et al., 1987; Razem and Sedam, 1985). In the Western Area, where relief is lower than in NE Pennsylvania, there is less driving head necessary to form deep flow systems. Here, groundwater flow is characterized by short flow paths typically extending no more than a few tens of kilometers (encompassing local and intermediate flow systems) with fresh water extending a few tens of meters below land surface before salt water is encountered (USGS, 2014). Permeable coal units above streams can act as drains and can be regional potable water aquifers (e.g. Carswell and Bennett, 1963).



**Fig. 1.** Chesapeake's pre-drilling sample sites and key stratigraphic units in Northeast Pennsylvania. Samples that fall outside of the key stratigraphic units are categorized in the "Other" formation category. Approximately 98% of pre-drilling samples were collected from these 4 counties in NE PA: Bradford, Sullivan, Susquehanna, and Wyoming.





**Fig. 2.** Chesapeake's pre-drilling sample sites and key stratigraphic units in the "Western Area". Samples that fall outside of the key stratigraphic units are categorized in the "Other" formation category.

In both areas, shallow wells typically yield fresh water whereas deeper wells generally find salt water (in valleys, shallow wells also commonly find salt water). According to Piper (1933), Lohman (1937, 1939), and Poth (1963) the salt water is connate water (salt water present during marine deposition) but has been modified by various physical and chemical processes. In the upper-most rocks, connate salt water has been flushed by recharging meteoric water in contrast to deeper water only being partially flushed or not flushed at all by circulating meteoric waters.

In NE Pennsylvania, groundwater typically evolves along flow paths from a  $\text{Ca-HCO}_3$  hydrochemical facies to a  $\text{Na-HCO}_3$  facies and finally sometimes to  $\text{Na-Cl}$  facies groundwater in valley discharge areas (e.g. Molofsky et al., 2013; Williams et al., 1998; Callaghan et al., 1998). Domestic water wells can also intersect isolated coal lamina within the Catskill and Lock Haven formations (e.g. Wilson, 2014; Molofsky et al., 2011). Where the coal occurs, particularly in the Western Area, oxygenated recharge water leads to groundwater with high concentrations of iron and sulfate from pyrite oxidation.

Most private domestic water wells completed in consolidated bedrock in both the Western Area and NE Pennsylvania consist of open-rock wells, with only a short surface casing driven through the shallow, unconsolidated sediments or unstable bedrock. As a result, there is no filter pack to filter out clay- and silt-sized sediment and other particulates since the well bore is open to the formation. In contrast, water wells completed in alluvium and glacial deposits in stream valleys are typically open-ended completions (casing with an open end). Screened completions are rare in domestic wells (Williams et al., 1998).

## 4. Methods

### 4.1. Regulatory framework

The geochemical dataset we describe consists of data compiled from Chesapeake's groundwater quality survey programs designed to sample domestic/stock water wells within a radius of 762–1219 m from proposed unconventional oil and gas well sites prior to drilling operations. The spacing of sampling was done according to accepted state regulatory programs, but Chesapeake often went beyond the requirements of these programs (for example the required pre-drill sampling distance per Pennsylvania Department of Environmental Protection (PADEP) regulations in Pennsylvania is 762 m, but Chesapeake often extended the sampling distance to 1219 m). Chesapeake's contractors collected 13,040 pre-drilling samples from domestic or stock water-supply wells in NE Pennsylvania between June of 2009 and January 2012. In the Western Area 8004 samples were collected by contractors between September of 2009 and June 2012. Pre-drilling sampling continues in both areas and all data is released to relevant state agencies and individuals whose groundwater was sampled.

Although these samples were collected to provide *pre-drilling* water quality reference points at the local scale, conventional oil and gas drilling had occurred in both regions prior to 2007. In the four primary counties of NE Pennsylvania (Bradford, Sullivan, Susquehanna, and Wyoming), where 98% of these pre-drilling sampling data were collected, approximately 78 conventional wells had been drilled from 1900 through 2012, compared to about



System	NE Pennsylvania	West Virginia SW Pennsylvania	Ohio SW Pennsylvania
Permian	Dunkard Group	Dunkard Group	Dunkard Group
Pennsylvanian	Monongahela Group	Monongahela Group	Monongahela Group
	Conemaugh Group	Conemaugh Group	Conemaugh Group
	Allegheny Group	Allegheny Formation	Allegheny Formation
	Pottsville Group	Pottsville Group	Pottsville Group
Mississippian	Mauch Chunk Group		
		Mauch Chunk Group	
		Greenbrier Limestone	
	Pocono Group	Pocono Group	Logan Formation
			Black Hand Sandstone
Devonian	Catskill Formation	Hampshire Formation	Bedford Shale
	Lock Haven Formation	Chemung Formation	Ohio and Olentangy Shale
	Brallier Formation	Brallier Formation	
	Genesee Shale	Harrel Shale	
			Berea Sandstone

**Fig. 3.** Generalized stratigraphic columns for Paleozoic rocks in Pennsylvania, West Virginia, and Ohio that occur at/near the surface or are part of bedrock aquifers in the study areas (modified from Lloyd and William (1995), Trapp and Horn (1997), Ohio Geological Survey (1998), and USGS (2012)). The Dunkard Group includes the Greene, Washington and Waynesburg formations in Pennsylvania, West Virginia, and Ohio. The Pocono Formation is also known as the Burgoon Formation or Burgoon Sandstone in Pennsylvania. The lower portion of the Pocono Formation is also known as the Huntley Mountain Formation in Pennsylvania. The stratigraphy for SW Pa is most similar to West Virginia or Ohio.

1934 unconventional wells from 2007 through 2012 according to PADEP records (PADEP, 2015). Prior to 2007, approximately 5 unconventional wells had been drilled in NE Pennsylvania, and slightly over 30 in the Western Area.

Therefore, while the dataset represents pre-drill samples for specific future unconventional oil/gas wells, the same samples also reflect *post-drilling* samples for oil/gas wells than may have been in place at the time of sampling. Siegel et al. (2015) recently reported from the same dataset in NE Pennsylvania that there is no systematic increase in dissolved methane proximate to existing gas and oil wells. In the Western Area there are a much larger number of pre-existing conventional wells dating back to the mid-1800s.

#### 4.2. Water well sampling protocols

Chesapeake's contractors collected groundwater samples from actively used domestic or stock water wells (from water outlets before treatment systems or pressure tanks, if possible) three to six months before natural gas drilling began in accordance with state regulatory programs. No samples were collected from public water systems. Sampling adhered to individual state regulatory protocols, and samples were analyzed in accordance with established US Environmental Protection Agency (USEPA) test methods and internal laboratory standard operating procedures (SOPs) using best practices accredited by the National Environmental Laboratory Accreditation Program (NELAP) and/or state regulatory authority. Chesapeake used three national contract laboratories for the analytical testing and four national contract environmental engineering firms to collect and process the groundwater samples. Chesapeake personnel did not collect, handle, or transport any pre-drilling samples.

In some instances more than one sample was collected from the same well. In almost all of these instances, only one additional sample was collected. In NE Pennsylvania, approximately 196 additional samples were collected, and in the Western Area approximately 21 additional samples were collected. These additional samples constitute a very minor portion of the dataset, and almost all were still collected as pre-drilling samples. Where these additional samples exist, the water-quality evaluation only used the highest concentration for a particular parameter.

As per protocols, unfiltered water samples were collected after allowing water to run for approximately 15 min to clear water lines and pressure tanks. Field measurements for water temperature, turbidity, pH, and specific conductance were made using standard field instruments (e.g. handheld YSI 556MPS multi-parameter meter and Lamotte 2020 field turbidity meter).

Once stability in field measurements were obtained, the sample aliquot was field-preserved with nitric acid to pH < 2 standard units (S.U.) for metals analyses and one or more unpreserved aliquots were collected for general chemistry parameters. Unpreserved aliquots for dissolved light gases were also collected and those results will be discussed in another paper. All samples were shipped to analytical laboratories overnight under chain-of-custody control and maintained at 6 °C or less.

The standard pre-drilling analytical parameter lists for both NE Pennsylvania and the Western Area are shown in Table S1 in Supplemental Information. This parameter list included 38 different laboratory analytical parameters as well as 4 field parameters. Figs. S1 (NE Pennsylvania) and S2 (Western Area) in the Supplemental Information present box and whisker plots showing percentiles for each parameter. Tables S2 (NE Pennsylvania) and S3 (Western Area) in the Supplemental Information summarizes the key statistics (such as mean, median, minimum value, maximum value, percentiles, and standard deviation) for each parameter. Analytical detection limits were equal to or below applicable water-quality standards noted in Table S1 (Supplemental Information). All analyses were done according to USEPA protocols and quality assurance, with appropriate instrumentation; the reader is referred to USEPA for details on these standard laboratory approaches (USEPA, 2014).

The analytical data was reviewed for Quality Control and Quality Assurance by third party contractors who uploaded the data into an electronic database maintained by Chesapeake.

#### 4.3. Drinking-water standards and benchmarks

We compared the pre-drilling groundwater data to USEPA drinking-water standards (USEPA, 2015a) and guidelines: primary maximum contaminant levels (MCLs), secondary maximum contaminant levels (SMCLs), health advisory levels (HALs), or regional screening levels (RSLs) for tap water (USEPA, 2015b). The USEPA primary drinking water MCLs are enforceable health-based standards applied to public-water systems whereas SMCLs are non-enforceable standards for contaminants that may cause cosmetic or esthetic effects to the water. The USEPA HALs provided information on constituents (such as sodium) that may affect human-health. Finally, USEPA RSLs are risk-based screening levels considered protective for human health. Although we recognize that these standards are locally enforced differently, they have in practice been considered the same in the public debate over unconventional gas development (ATSDR, 2011).

If there was no MCL for a parameter, we used the SMCL. If no MCL or SMCL was available, we used a HAL or RSL for tap water. All regulatory drinking-water standards or guidelines used in this evaluation are provided in Table S1 (Supplemental Information).

#### 4.4. Water well topographic position

Heisig and Scott (2013), Williams et al. (1998), and Stoner et al. (1987) identified possible relationships between water chemistry and topographic position in the Appalachian Basin and we similarly assessed this by classifying water chemistry in wells in “upland” or “valley” topographic settings. We include hilltops and intervening slope areas in the upland category; valleys include valley floors.

Topographic position can be broadly used as a proxy for groundwater residence time, or age based on subsurface flow paths through aquifers. Fig. 4 shows a generalized conceptual hydrochemical flow model for the Appalachian Plateau. Typically, older groundwater becomes more mineralized due to water–rock interactions along flow paths and mixing with connate brines. Streams in valley settings of the Appalachian Plateau serve as local- to intermediate-scale groundwater discharge zones, along with side-slope springs. Groundwater in major river valleys, the terminus of local, intermediate and regional groundwater flow paths, can be more mineralized than groundwater in upland settings. The left-hand side of the Fig. 4 reflects conditions more representative of the Western Area where interlayered coal seams containing pyrite lead to more sulfate in shallow groundwater than found in NE Pennsylvania.

We evaluated topographic position in NE Pennsylvania and the Western Area by two automated approaches. The first, the Topographic Position Index (TPI) method (Jenness et al., 2013; Weiss, 2001) compares (LiDAR) elevation at a particular location to the mean elevation in a neighborhood around the location. The second method uses average valley widths (after Molofsky et al. (2011)) and then applies buffers of corresponding size to stream centerlines from the National Hydrography Dataset (NHD) to validate the locations of valleys. We used the same buffers used by Molofsky et al. (2011); a 305 m buffer for the larger streams and a 152 m buffer for the smaller streams. In our

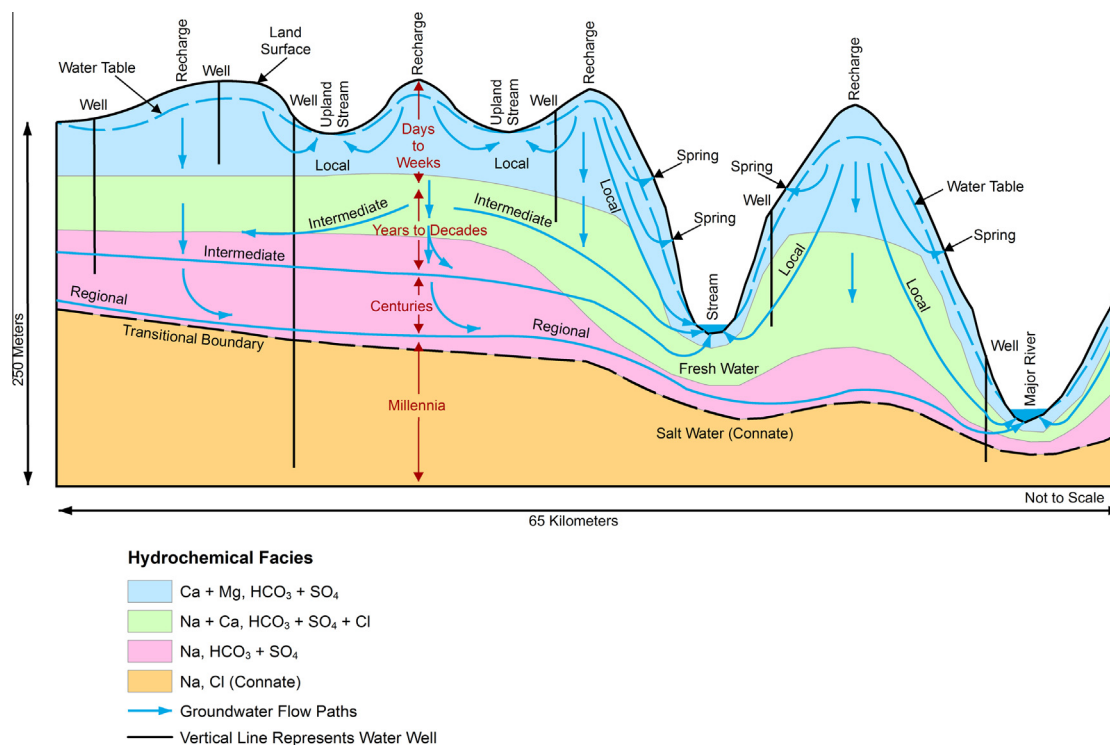
analysis, we only used those wells where both methods agree with respect to occurrence of water type relative to topographic position.

#### 4.5. Water well completion depth and geology

We estimated the probable formation that each well was set in by overlaying the well locations onto GIS shape files of USGS or state geologic survey bedrock maps (Nicholson et al., 2007; Pennsylvania Bureau of Topographic and Geologic Survey, 2001).

Information about well depth was provided by the well owner. Williams et al. (1998) noted that there are three water well types in common use in NE Pennsylvania: an open-hole well, an open-end well, and a screened well. Williams et al. (1998) noted that screened wells are mostly used in non-domestic wells completed in unconsolidated deposits, while the first two are used in most domestic completions.

We did an extensive review to match water-well logs in the Pennsylvania Groundwater Information System (Pennsylvania Department of Conservation and Natural Resources, 2014) database with the domestic/stock water wells from which the pre-drill samples were obtained, with limited success. Taylor (1984) published well completion data from 890 wells in Bradford, Susquehanna, Sullivan, and Wyoming counties of NE Pennsylvania wherein he showed completion information for wells based upon topographic setting (i.e. valleys, side slopes, or hilltops). Almost all wells completed on hilltops and side slopes were open-hole completions in the Catskill and Lock Haven formations, and in the valleys over 71% of the domestic wells after penetrating alluvium were still open-hole completions in the bedrock. Water wells in valley settings therefore cannot be assumed to be completed only in unconsolidated materials. Most high capacity wells for public, irrigation, industrial or commercial needs are completed in unconsolidated materials, but this is not the case with domestic water wells, which make up the entirety of the pre-drilling dataset.



**Fig. 4.** Conceptual diagram showing groundwater flow zones in the Appalachian Plateau for different hydrogeologic settings and relationship with hydrochemical facies (modified after Fleeger (1999) and Wunsch (1995)). In the “Western Area”, flow systems are shallower with waters in both regions discharging to streams possibly having variable chemistries based on the nature of the local convergence of flow systems at different scales—local, intermediate, and regional.



For our analysis we assumed that domestic water wells on side and upland areas overwhelmingly consisted of open-hole wells obtaining water from fractured bedrock. We knew the mapped bedrock formation underlying valleys and assigned valley wells per that bedrock type, knowing that there will be some unknown percentage of wells that would actually be completed only in the overlying alluvial or glacial deposits. However, our review showed that the common practice in the four-county area of NE Pennsylvania is to drill through unconsolidated sediments, set steel surface casing, and complete an open-hole well in the bedrock formations.

We also used well depth estimates to evaluate completion intervals in the valleys. Information provided by Taylor (1984) showed that the median depth for water wells completed within unconsolidated materials in valleys was 18.3 m (75th percentile was 28.3 m) for the four-county area representing the majority of the pre-drilling samples. When we evaluate the water well depth database for NE Pennsylvania, approximately 75% of wells completed in valleys have depths that exceed the median value of 18.3 m for unconsolidated deposit wells.

For geological evaluations involving these data, we have removed those wells that are more likely to be completed in unconsolidated deposits in valley settings which have reported well depths less than 28.3 m (approximately 3000 wells removed). Removing these shallow valley wells removes most of the potential influence of the unconsolidated deposits. After completing this process, we found essentially no difference in median parameter values of concentrations measured between the two datasets, showing that wells in valleys are mostly completed in bedrock formations, or that unconsolidated deposits have nearly identical water quality.

Table S4 (Supplemental Information) provides our estimate of the number of samples from each geological group or formation in NE Pennsylvania and in the Western Area. Table S5 (Supplemental Information) provides estimates of the number of samples from each geological formation in a particular topographic position (upland or valley setting).

We assumed that the bedrock unit supplying the groundwater to the water well was the same as the bedrock unit mapped at the ground surface or sub-cropping beneath the valley sediments. This is a reasonable assumption for the majority of wells because of the near horizontal attitude of the formations and overall thicknesses. For example, the Catskill and Lock Haven formations can be 1707 m (Inners, 1981) to 1219 m (Taylor, 1984) thick, respectively, compared to a median well depth in this area of 50.3 m (maximum reported well depth of 518 m). In NE Pennsylvania, 94% of the pre-drilling samples were from locations overlying the Catskill or Lock Haven formations.

For simplification in the Western Area, we evaluated bedrock at the Group level rather than the individual formation (see stratigraphic column in Fig. 3). For example, we combined the Greene, Washington, and Waynesburg formations into the Dunkard Group. In the Western Area, 90.5% of the pre-drilling samples were from the Allegheny/Pottsville formations and the Conemaugh Group. Our broad approach would lead to misidentification of the geology only where wells are extremely deep (and could intersect a different unit) or near the contact between formations; a very small fraction of the overall dataset. In areas where alluvial or glacial deposits were present, most domestic water wells were still completed into bedrock, and steel casing driven through the unconsolidated deposits.

#### 4.6. Analysis of data

We used Microsoft Access and Excel, and Stata® software, version 11.2 (StataCorp., 2009) to calculate summary statistics.

We provide a summary of the pre-drilling analytical data including key statistics as well as boxplots for analytes in the [Supplemental Information](#).

#### 4.7. Geochemical software and ion balance

We used the water-quality analysis program AquaChem® (Schlumberger, 2011) to generate Piper trilinear diagrams, calculate ion-charge balances, and determine primary ion water types. Piper diagrams visually characterize the broad geochemical evolution of shallow groundwater and can distinguish different groundwater types, mixtures of waters, and in some cases, major geochemical processes along groundwater flow paths (e.g. Hounslow, 1995).

We classified groundwater type or hydrochemical facies by selecting the single cation and anion present in each sample with highest percentage of cation or anion concentration (as milliequivalents per liter, meq/L). For convenience and the purpose of this broad synthesis, we did not address in detail mixed-water types, although the Piper diagrams show that geochemical mixing widely occurs. For example, mixed water types such as calcium-sodium bicarbonate (Ca-Na-HCO<sub>3</sub>) or Ca-Mg-HCO<sub>3</sub> do occur within our study areas. It is possible that subdividing our broad regions into smaller areas will lead to better understanding of specific geochemical processes conditional with natural water chemistry. We include results only with a charge balance within ±7% in all groundwater hydrochemical facies evaluations. However, the full dataset, regardless of charge balance, was used for comparisons to relevant drinking-water standards.

#### 4.8. Water treatment by water well owners

As part of our study, we determined whether water samples were collected prior to or after a home water-treatment system (a bag filter for sediment removal, water softener, and far less commonly, a home/tap reverse-osmosis unit). In cases where the water-treatment status was unknown, it was assumed that the water was sampled before any treatment in accordance with the sampling program protocols.

Table S6 (Supplemental Information) shows the breakdown of sample numbers that were treated, untreated, or unknown for each region. We excluded all treated and unknown samples, regardless of treatment system type, in hydrochemical facies classification by Piper diagrams. Although our approach excluded many samples for this specific evaluation, over 2600 untreated samples remained in NE Pennsylvania and over 3400 samples in the Western Area within a charge balance of ±7%. These are still large datasets by any measure and are sufficiently representative of the total population.

In our interpretation of arsenic, barium, cadmium, chromium, iron, lead, lithium, manganese, mercury, silver, sodium, and strontium water-quality patterns, we used all sample results regardless of charge balance or whether the water sample was collected after a treatment system. We used total results (unfiltered samples) since state regulatory agencies base their assessments on concentrations of total metals in water-supply systems, not dissolved solutes even if particulates occur in the water. We understand that from a regulatory standpoint, turbid water can lead to an inference of contamination. In turbid water, metals bound in the crystal lattices of clay-sized particulates become incorporated in the chemical analysis. These particulates are natural, and regulators may interpret these results as contamination presumed in the dissolved state when, in fact, it is not present. We wanted to assess how prevalent such false positives could be.

## 5. Results

### 5.1. Quality of analyses

In addition to evaluating ion charge balances for individual samples, the mean of all charge balances can provide an assessment of overall quality of a database (Fritz, 1994) and should be close to zero. The mean raw ion charge balance (with 1 standard deviation) was  $0.30 \pm 7.82\%$  and  $1.77 \pm 6.07\%$  for NE Pennsylvania and the Western Area, respectively. The overall mean of the absolute ion charge balance is  $4.53 \pm 6.37\%$  and  $3.86 \pm 5.01\%$ , for NE Pennsylvania and the Western Area, respectively.

If we only consider those data with ion balances within  $\pm 7\%$ , then for NE Pennsylvania the overall mean ion balance is  $0.49 \pm 3.17\%$  and the absolute mean balance is  $2.63 \pm 1.84\%$ . Similarly for the Western Area, the overall mean ion balance is  $1.41 \pm 2.97\%$ , and the absolute mean balance is  $2.74 \pm 1.82\%$ . These mean ion balances are considered excellent for both groups of samples. For NE Pennsylvania, 83% (9361) of the samples with analyzed major ions balanced within our  $\pm 7\%$  criterion and within the Western Area, 88% (7042) of the samples were within  $\pm 7\%$ . We present the summary statistics for the ion balances in Table S7 in the Supplemental Information for reference.

### 5.2. Pre-drilling groundwater-quality exceedances

#### 5.2.1. Northeast Pennsylvania

Table 1 shows the percent of exceedances of USEPA water-quality standards from the pre-drilling groundwater samples in NE Pennsylvania. Fig. 5a–k shows the spatial distribution of pre-drilling samples exceeding standards for arsenic, barium, iron, lead, manganese, sulfate, chloride, lithium, sodium, total dissolved solids (TDS), and turbidity, respectively.

Fig. S3 in the Supplemental Information shows the spatial distribution of all samples exceeding an MCL, SMCL, HAL, or RSL (excluding turbidity) for NE Pennsylvania. In NE Pennsylvania, 10.2% of the pre-drilling samples exceeded one or more MCLs (excluding turbidity). If turbidity is included then 22.1% of all pre-drilling samples exceed a MCL. Secondary MCLs were exceeded in 46.1% of the samples. Fig. 6 shows the percentage of water-quality exceedances per guideline or standard category for NE Pennsylvania.

The most common pre-drilling MCL exceedances were for turbidity (15.6%), arsenic (4.2%), lead (3.6%), and barium (3.3%). The most common SMCL exceedances were for manganese (33.9%) and iron (23.8%). Sodium concentrations exceeded the USEPA HAL of 20 mg/L in 38.5% of pre-drilling samples. Lithium exceeded the USEPA RSL for tap water in 19.1% of the well pre-drilling samples. Other metals also occasionally exceeded recommended standards. If all the MCL, SMCL, RSL, and HAL standards or guidelines are considered collectively (excluding turbidity), then approximately 63.0% of all pre-drilling samples collected from NE Pennsylvania exceeded at least one drinking-water standard, 34.3% exceeded two or more drinking-water standards, and approximately 15.4% exceed three or more standards. Including turbidity, 63.1% of pre-drilling samples exceeded one or more drinking-water standards.

Our results agree with historical water-quality studies in NE Pennsylvania (e.g. Ayotte et al., 2011; DeSimone, 2008; Williams et al., 1998; Taylor, 1984) that relied on much smaller datasets and before unconventional gas drilling began. For example, Williams et al. (1998) reported about 50% of 223 water wells sampled in a three county area within the pre-drilling sampling area of NE Pennsylvania had iron and manganese concentrations in groundwater exceeding SMCLs. Williams et al. (1998) also showed that chloride, arsenic, barium, cadmium, and lead concentrations

**Table 1**  
Pre-drilling sample results exceeding drinking water guidelines – Northeastern Pennsylvania.

Parameter <sup>a</sup>	Number of samples analyzed <sup>d</sup>	Drinking water guideline or standard (mg/L)	Drinking water guideline or standard type <sup>b</sup>	Number of Samples exceeding guideline or standard <sup>d</sup>	Percent of Samples exceeding guideline or standard <sup>d</sup>
Arsenic	11,034	0.010	MCL	462	4.2
Barium	11,074	2.0	MCL	362	3.3
Benzene	11,075	0.005	MCL	1	<0.1
Cadmium	11,034	0.005	MCL	4	<0.1
Chloride	11,073	250	SMCL	221	2.0
Chromium	11,034	0.1	MCL	0	0.0
Ethylbenzene	11,075	0.7	MCL	0	0.0
Iron	11,075	0.3	SMCL	2647	23.8
Lithium	1729	0.04	RSL	331	19.1
Lead	11,032	0.015	TTL	405	3.6
Manganese	11,074	0.05	SMCL	3752	33.9
Mercury	11,034	0.002	MCL	0	0.0
Surfactants <sup>c</sup>	11,071	0.5	SMCL	28	0.3
pH	11,073	6.5–8.5 S.U.	SMCL	645	5.8
Selenium	11,032	0.05	MCL	17	0.2
Silver	11,034	0.1	SMCL	0	0.0
Sodium	11,074	20	Advisory	4263	38.5
Strontium	3423	12	RSL	10	0.3
Sulfate	11,075	250	SMCL	82	0.7
TDS	11,075	500	SMCL	554	5.0
Toluene	11,075	1	MCL	0	0.0
Turbidity	11,076	5 NTU	MCL	1738	15.7
Xylenes	11,075	10	MCL	0	0.0

TDS: Total Dissolved Solids.

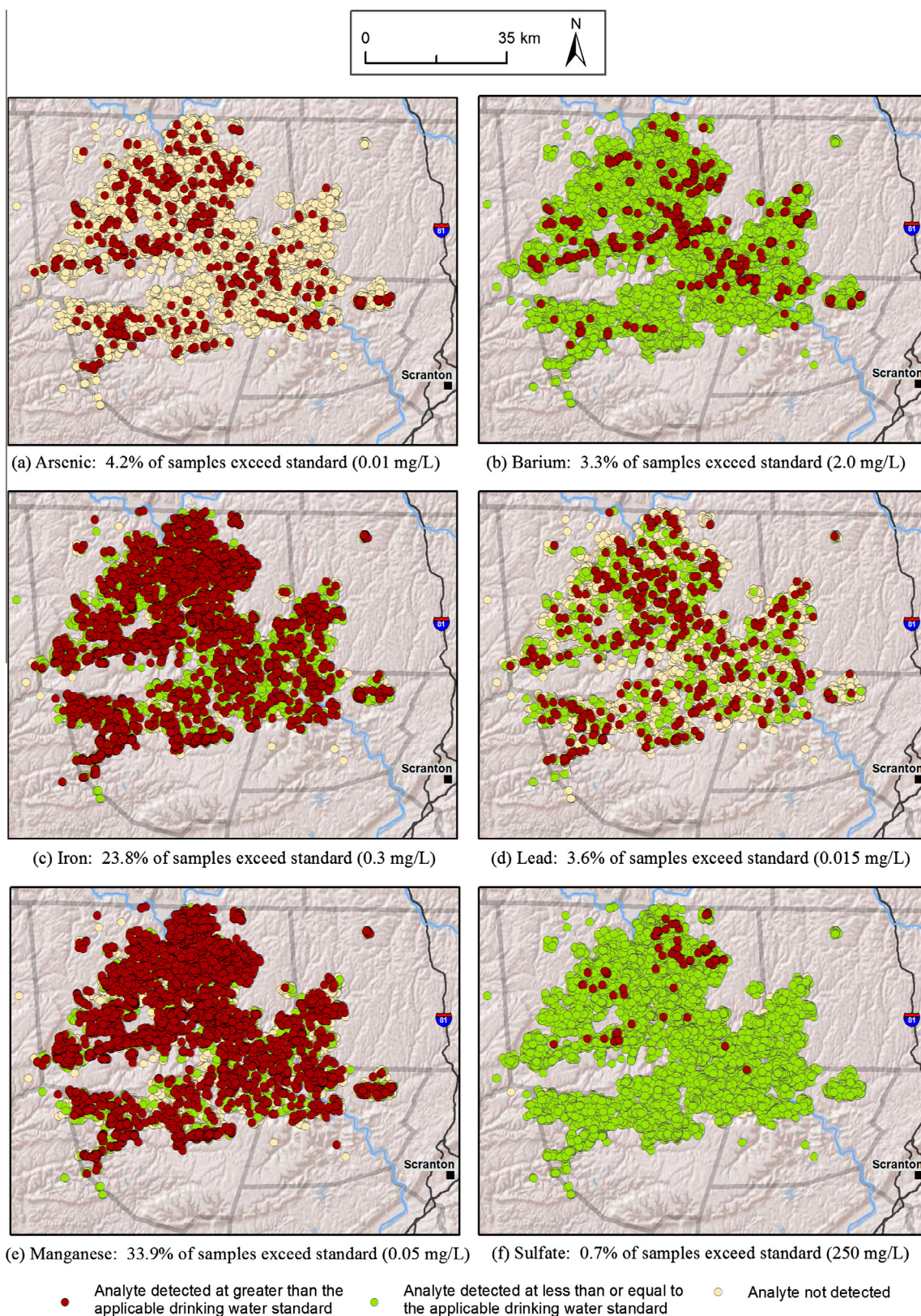
<sup>a</sup> Samples analyzed for metals were not filtered during sampling or analysis. Some of the results may represent samples collected after a treatment system (e.g. particulate filter, water softener, etc.) within the water distribution system.

<sup>b</sup> Guideline or Standard types (see Table 1): MCL –USEPA primary Maximum Contaminant Level for public drinking water supplies; SMCL –USEPA Secondary MCL; RSL –USEPA Regional Screening Level for tap water (January 2015); TTL – treatment technology action level for lead defined by USEPA for public drinking water supplies; Advisory –USEPA drinking water health advisory for individuals on a restricted sodium diet.

<sup>c</sup> Surfactants analyzed as methylene blue active substances (MBAS).

<sup>d</sup> For duplicate samples only sample with highest value used for each parameter.





**Fig. 5.** Distribution of trace metals and other constituents in pre-drilling groundwater samples in Northeastern Pennsylvania. Samples exceeding the applicable MCL or other drinking water standard are shown with a red symbol; such elevated concentrations occur commonly in pre-drilling samples across the study area. The applicable standard and percent of samples exceeding that standard are listed beneath each map. For wells sampled more than once, the sample with the highest value is shown. Layers are vertically stacked in order shown in legend and some “no exceeds” may be covered.



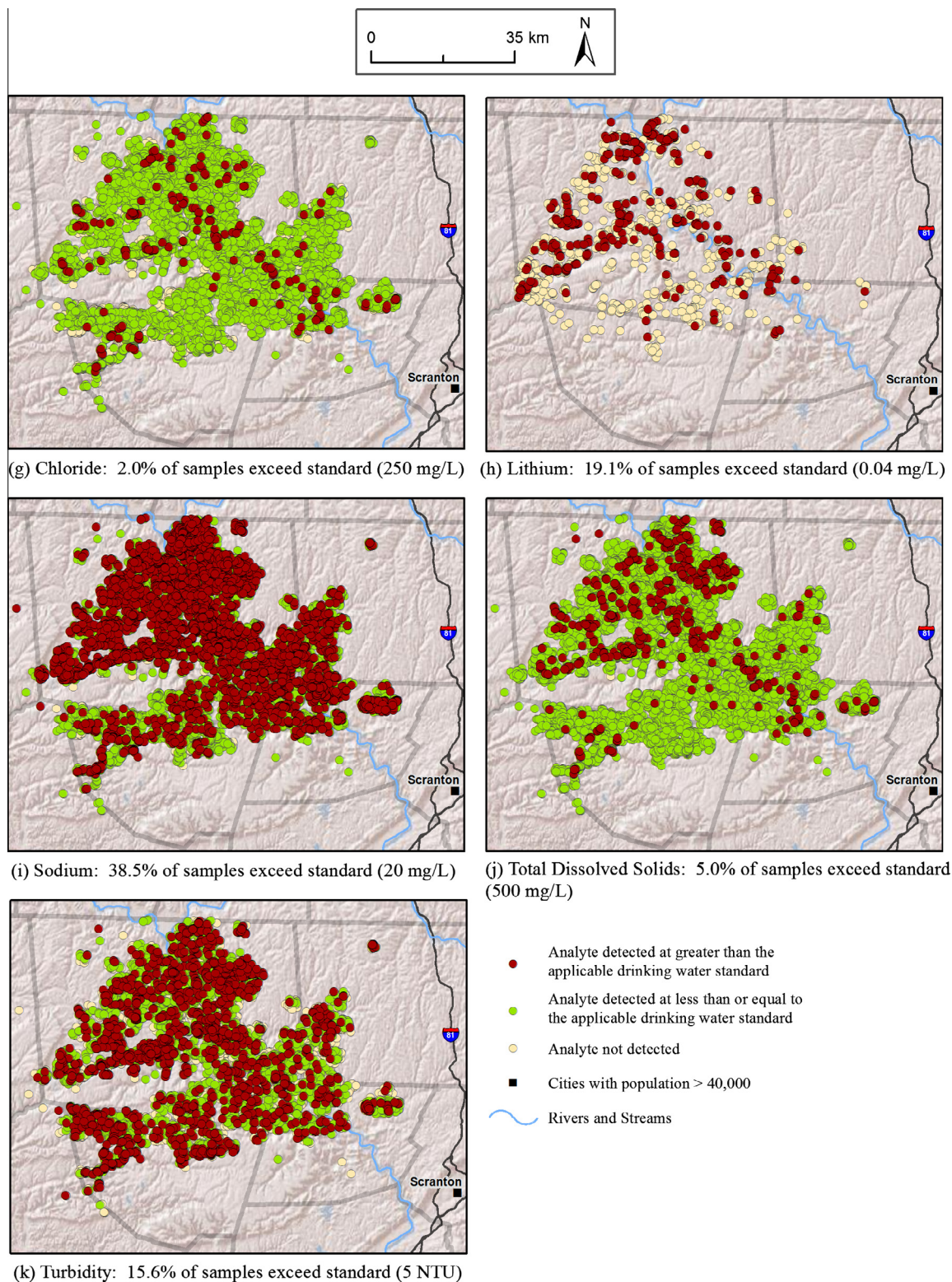


Fig. 5 (continued)

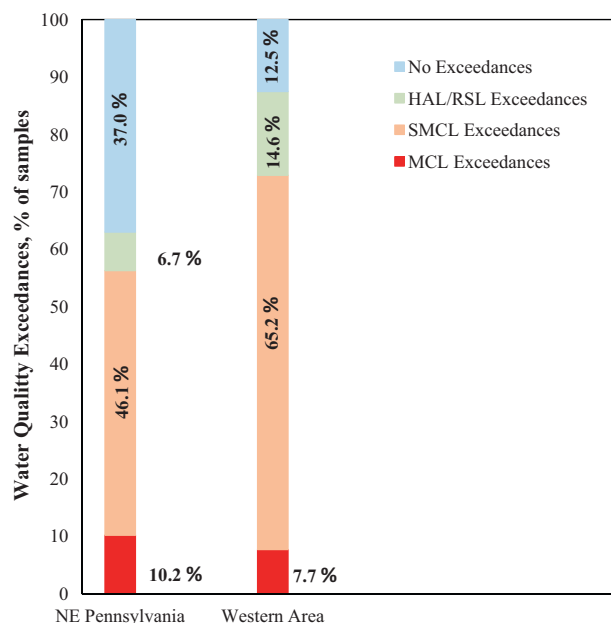
in groundwater can also exceed MCLs/SMCLs, especially in restricted-flow zones containing shallow naturally-saline groundwater, and about 40% of sodium concentrations in that study exceeded the HAL.

Table 2 compares the key water-quality parameters from the historical data provided in Williams et al. (1998) and Taylor (1984) to the 2009–2012 pre-drilling data for the Catskill and

Lock Haven formations. The median chloride, TDS, and sodium concentrations in the pre-drilling data, shown in Table 2, have similar ranges as the historical data (pre-2007) from the Catskill and Lock Haven formations. These major parameters defining groundwater tie closely with water-quality type.

The Williams et al. (1998) pre-1987 historical data from the Catskill and Lock Haven formations also show that collectively





**Fig. 6.** Figure showing the percentage of water-quality exceedances in NE Pennsylvania and the “Western Area” by guideline. MCLs are USEPA maximum contaminant levels. SMCLs are USEPA secondary maximum contaminant levels. HALs are USEPA health advisory levels and RSLs are USEPA screening levels for water consumed at the tap. Turbidity exceedances are excluded from this chart. Each sample included only once; if there was an exceedance of a primary MCL, then the sample was not considered in the evaluation of the remaining guidelines, and so forth.

22% of historical samples exceeded an MCL, compared to approximately 9.6% (Catskill and Lock Haven samples only) in the 2009–2012 pre-drilling data, with the most common MCL exceedances in the historical data caused by arsenic, barium, selenium, and lead.

Battelle (2013) evaluated pre-2007 historical groundwater data from about 500 water wells (undivided by geological formation) in Bradford and Susquehanna Counties of NE Pennsylvania. Table 3 shows a comparison of the historical groundwater data from the Battelle (2013) report (pre-2007) compared to the pre-drilling data. The median values for all parameters are very similar.

### 5.2.2. Western area

Table 4 shows the distribution of major solutes and trace metals of interest compared to USEPA water-quality standards for

pre-drilling water well samples in the Western Area. Fig. 7a–k shows the spatial distribution of pre-drilling samples that exceed USEPA drinking-water MCLs, SMCLs, RSLs, or HALs for: arsenic, barium, iron, lead, manganese, sulfate, chloride, lithium, sodium, TDS, and turbidity, respectively.

Fig. S4 in the Supplemental Information shows the spatial distribution of pre-drilling sample sites that exceed one or more MCLs, excluding turbidity. In the Western Area 7.7% of all pre-drilling samples exceeded one or more MCLs (excluding turbidity). If turbidity is included then 39% of all pre-drilling samples exceed an MCL. Secondary MCLs were exceeded in 65.2% of the samples.

Excluding turbidity (36.2%), the most commonly exceeded MCLs in pre-drilling samples from the Western Area are lead (4.4%) and arsenic (3.1%). If all of the MCL, SMCL, RSL, and HAL water-quality standards are considered collectively (excluding turbidity) then 87.5% of pre-drilling samples exceed one or more drinking-water standards. About 63.8% of all pre-drilling samples exceeded two or more drinking-water standards, and approximately 31.1% exceeded three or more drinking-water standards. Including turbidity, 88.7% of pre-drilling samples exceeded one or more drinking-water standards. Fig. S4 in the Supplemental Information also shows the spatial distribution of those samples that exceed a SMCL, HAL, or RSL. The most commonly detected natural exceedances of SMCLs were manganese (53.5%), iron (50.0%), and TDS (22.8%). The sodium USEPA HAL of 20 mg/L was exceeded in 51.8% of samples collected from water wells in the Western Area.

Our findings in the Western Area are also consistent with local and national historical water-quality studies (e.g. Ohio USEPA, 2012; Gross and Low, 2012; Chambers et al., 2012; Ayotte et al., 2011; DeSimone et al., 2009; DeSimone, 2008; USGS, 2006; McAuley and Kozar, 2006; White and Mathes, 2006; Stoner et al., 1987; Razem and Sedam, 1985; Matisoff et al., 1982). For example, Razem and Sedam (1985) noted in a study of 100 groundwater samples collected in 1983 that SMCLs were commonly exceeded for TDS (38%), manganese (34%), iron (19%), and sulfate (17%) among other constituents, and that 62% of the samples analyzed exceeded the sodium HAL of 20 mg/L.

### 5.3. Sample turbidity influence on water-quality

The analysis of unfiltered samples (total analysis) for certain trace metals (e.g. iron and manganese) can produce false positives.

**Table 2**

Summary of historical analytical data (Williams et al., 1998 and Taylor, 1984) from the Catskill and Lock Haven formations compared to pre-drilling data (2009–2012) for Northeastern Pennsylvania (units are in mg/L). Historical data compared to full pre-drilling dataset for each formation, and also to the dataset where shallower wells were excluded (those <p75) in valleys.

Parameter	Catskill Formation Median concentrations (mg/L)				Lock Haven Formation Median concentrations (mg/L)			
	Historical Pre-1987 Williams et al., 1998 (Taylor, 1984)		Pre-drilling 2009–2012 All Data [Shallow valley wells excluded]		Historical Pre-1987 Williams et al., 1998 (Taylor, 1984)		Pre-drilling 2009–2012 All Data [Shallow valley wells excluded]	
	No.	Median (mg/L)	No.	Median (mg/L)	No.	Median (mg/L)	No.	Median (mg/L)
Calcium	40 (165)	26 (28.1)	5441 [4043]	33.1 [33.7]	54 (45)	39 (39)	4954 [3359]	43.2 [43.6]
Chloride	43 (165)	10 (4)	5441 [4043]	7.4 [7.3]	56 (45)	12 (5)	4954 [3359]	8.6 [8.1]
Iron	41 (165)	0.09 (0.07)	5441 [4043]	<0.05 [<0.05]	55 (45)	0.27 (0.27)	4954 [3359]	0.10 [0.09]
Manganese	36 (165)	0.03 (0.02)	5441 [4043]	<0.015 [<0.015]	52 (45)	0.05 (0.09)	4954 [3359]	0.04 [0.04]
Magnesium	40 (164)	5.5 (5.5)	5441 [4043]	4.9 [5.0]	55 (45)	10 (10.2)	4954 [3359]	10.5 [10.8]
Potassium	36 (164)	2.0 (1.02)	5441 [4043]	1.2 [1.2]	52 (45)	1.1 (2.3)	4954 [3359]	1.5 [1.6]
Sodium	37 (165)	11 (10.1)	5441 [4043]	12.5 [12.6]	54 (45)	28 (23)	4954 [3359]	21.2 [21.5]
Sulfate	41 (164)	10 (10)	5441 [4043]	13.0 [13.0]	55 (45)	16 (15)	4954 [3359]	18.4 [19.0]
TDS	38 (165)	160 (158)	5441 [4043]	165 [167]	51 (45)	300 (238)	4954 [3359]	240 [241]

Notes: Median concentrations for historical samples from Tables 10, 11, 16, 17 and 19 of Williams et al. (1998). Numbers in parenthesis are from Taylor (1984), Table 13 data. These tables also provide the numbers of samples. Calcium, magnesium, and potassium sample counts are from Table 20 of Williams et al. (1998). Historical samples analyzed for dissolved metals; pre-drilling samples analyzed for total metals. Numbers in brackets are pre-drilling data with shallow valley wells removed from dataset.

**Table 3**  
Summary of Battelle's historical (pre-2007) groundwater analytical data summary from Bradford and Susquehanna Counties compared to pre-drilling data (2009–2012) for Northeastern Pennsylvania (units are in mg/L).

Parameter	Historical data, pre-2007 (Battelle, 2013)		Pre-drilling data, 2009–2012	
	Number of samples	Median value (mg/L)	Number of samples	Median value (mg/L)
Calcium (as total)	70	39.7	11,074	35.9
Calcium (as dissolved)	159	31	–	–
Chloride	502	8	11,073	7.5
Iron (as total)	82	0.293	11,075	0.065
Iron (as dissolved)	141	0.110	136	<0.05
Magnesium (as total)	69	8.1	11,074	6.5
Magnesium (as dissolved)	197	6.69	–	–
Manganese (as total)	81	0.080	11,074	<0.015
Manganese (as dissolved)	344	0.0992	135	0.036
Potassium (as total)	82	1.1	11,074	1.3
Potassium (as dissolved)	135	1.05	–	–
Sodium (as total)	82	16.6	11,074	14.4
Sodium (as dissolved)	410	9.86	–	–
Sulfate	231	15	11,075	14.6
TDS	223	196	11,075	189

Notes:

–: No data.

Median concentrations are for historical data from Battelle (2013), Tables 3 and 4. Median concentrations calculated using only the highest parameter values for the duplicate samples.

**Table 4**  
Pre-drilling sample results exceeding drinking water guidelines – “Western Area”.

Parameter <sup>a</sup>	Number of samples analyzed <sup>d</sup>	Drinking water guideline or standard (mg/L)	Drinking water guideline or standard type <sup>b</sup>	Number of samples exceeding guideline or standard <sup>d</sup>	Percent of samples exceeding guideline or standard <sup>d</sup>
Arsenic	7933	0.010	MCL	249	3.1
Barium	7944	2.0	MCL	39	0.5
Benzene	7945	0.005	MCL	4	<0.1
Cadmium	7933	0.005	MCL	17	0.2
Chloride	7944	250	SMCL	210	2.6
Chromium	7933	0.1	MCL	1	<0.1
Ethylbenzene	7946	0.7	MCL	0	0.0
Iron	7944	0.3	SMCL	3967	50.0
Lithium	5020	0.04	RSL	113	2.2
Lead	7933	0.015	TTL	347	4.4
Manganese	7944	0.05	SMCL	4251	53.5
Mercury	7933	0.002	MCL	0	0.0
Surfactants <sup>c</sup>	7946	0.5	SMCL	8	0.1
pH	7945	6.5–8.5	SMCL	479	6.0
Selenium	7933	0.05	MCL	6	<0.1
Silver	7932	0.1	SMCL	0	0.0
Sodium	7944	20	Advisory	4113	51.8
Strontium	5931	12	RSL	0	0.0
Sulfate	7943	250	SMCL	387	4.9
TDS	7945	500	SMCL	1810	22.8
Toluene	7945	1	MCL	0	0.0
Turbidity	7946	5 NTU	MCL	2875	36.2
Xylenes	7945	10	MCL	0	0.0

<sup>a</sup> Samples analyzed for metals were not filtered during sampling or analysis. Some of the results may represent samples collected after a treatment system (e.g. particulate filter, water softener, etc.) within the water distribution system.

<sup>b</sup> Guideline or Standard types MCL – USEPA primary Maximum Contaminant Level for public drinking water supplies; SMCL – USEPA Secondary MCL; RSL – USEPA Regional Screening Level for tap water (January 2105); TTL – treatment technology action level for lead defined by USEPA for public drinking water supplies; Advisory – USEPA drinking water health advisory for individuals on a restricted sodium diet.

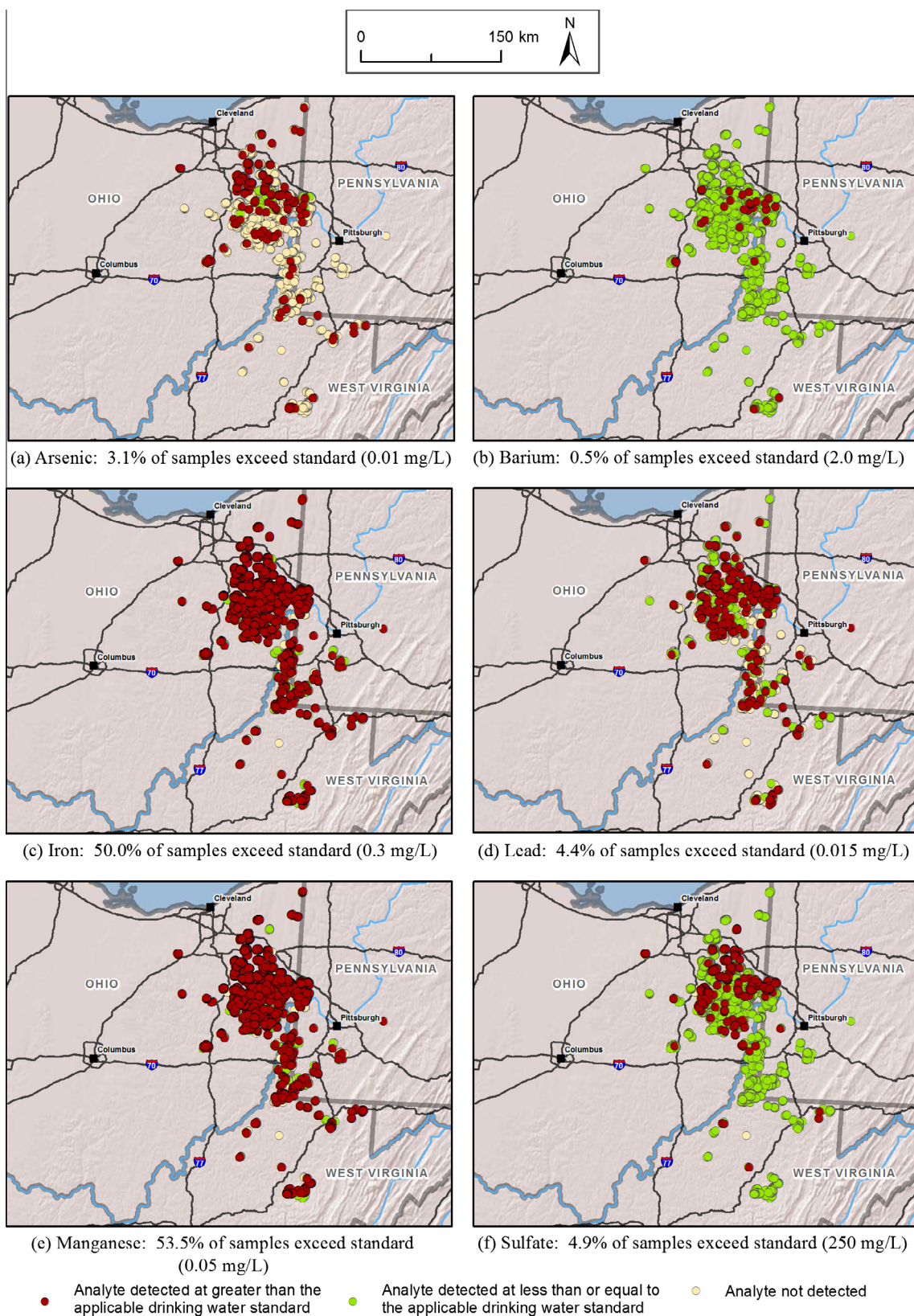
<sup>c</sup> Surfactants analyzed as methylene blue active substances (MBAS).

<sup>d</sup> For duplicate samples only sample with highest value used for each parameter.

Particulates induced during pumping of water wells are included in analysis of unfiltered samples. When water wells are pumped, the increased velocity of water moving through the formation and the well bore induces clay- to silt-sized sediment into the water. The USEPA recognizes this problem and advocates pumping at rates less than 100 milliliters per minute (mL/min) to minimize turbidity during sampling of monitoring wells (Puls and Barcelona, 1996). Domestic water wells pump at much higher rates, and consequently, turbidity can occur even under normal use. We found about 15.7% of the samples in NE Pennsylvania and 36.2% of

samples in the Western Area had turbidity exceeding the USEPA MCL of 5 nephelometric turbidity units (NTU). Many exceedances of total iron and manganese SMCLs relate to this turbidity, confounding true appraisal of trace metals concentrations from a scientific perspective or to characterize potential contamination. In rare cases, low pH in upland areas can react with domestic plumbing, and cause an occasional false-positive issue, but based upon our review, this is likely rare in both regions studied. We provide details on this part of our study in [Appendix A of the Supplemental Information](#).





**Fig. 7.** Distribution of trace metals and other constituents in pre-drilling groundwater samples in the “Western Area”. Samples exceeding the MCL or other groundwater standard are shown with a red symbol; such elevated concentrations occur commonly in pre-drilling samples across the study area. The applicable standard and percent of samples exceeding that standard are listed beneath each map. Layers are vertically stacked in order shown in legend and thus some “no exceeds” may be covered. For wells sampled more than once, the sample with the highest value is shown.

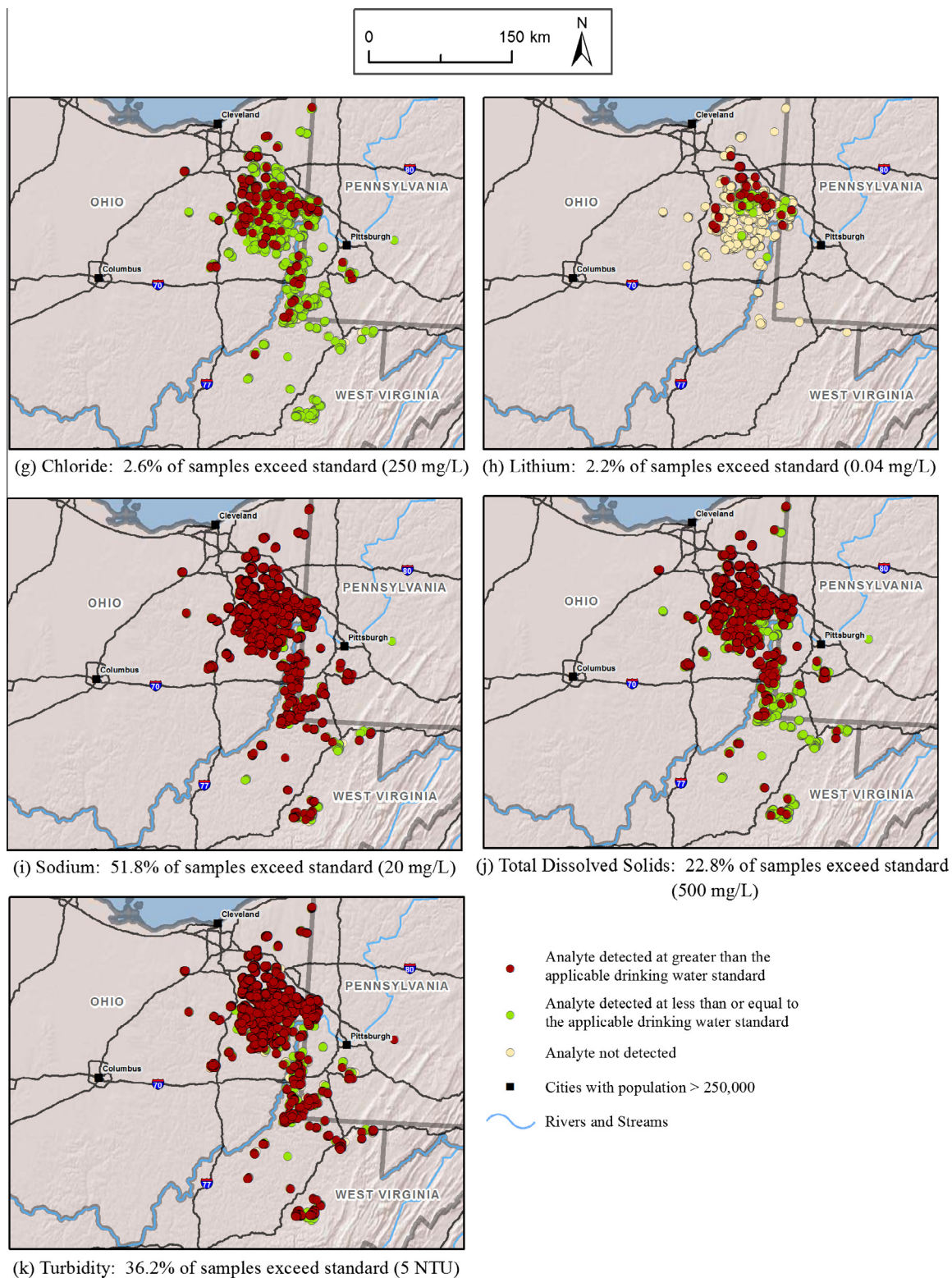


Fig. 7 (continued)

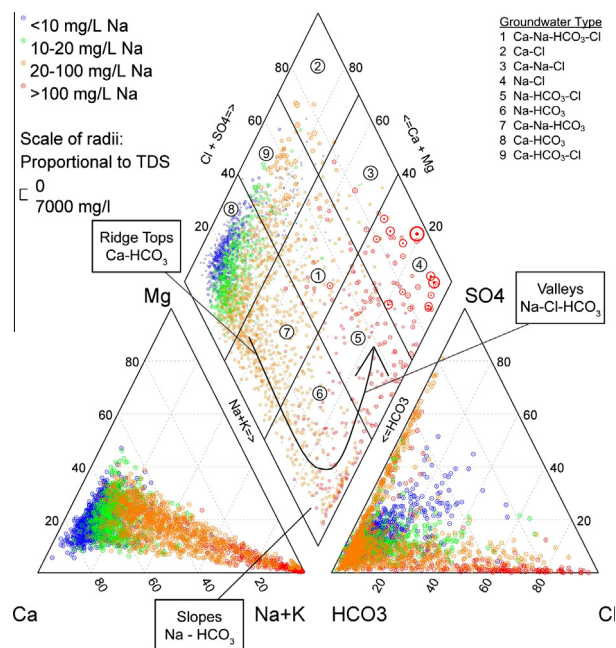
#### 5.4. Hydrogeological relationships, regulatory exceedances, and hydrochemical water type

Piper trilinear diagrams for NE Pennsylvania (Fig. 8) and the Western Area (Fig. 9) show a geochemical evolution along flow paths from a Ca-HCO<sub>3</sub> groundwater type on recharging hilltops to sodium enriched waters along valley flanks and in valleys,

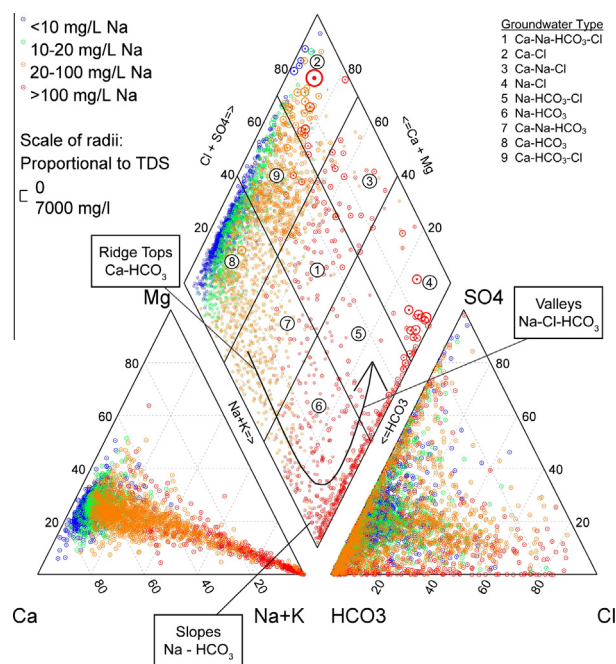
where discharging groundwater with higher salinities naturally occur. Sodium derives from a combination of natural calcium and magnesium ion exchange with sodium found in clays within the formation matrix and fracture surfaces.

All major-ion water types occur in valley and upland settings in both NE Pennsylvania and the Western Area. But in upland areas in NE Pennsylvania (where deep freshwater groundwater circulation





**Fig. 8.** Piper trilinear plot of pre-drilling samples in Northeastern Pennsylvania in four categories of sodium concentration. Circles in the diamond field are scaled to the concentration of TDS in the sample. The samples included here are known to be untreated and are within the  $\pm 7\%$  criteria on the ion charge balance ( $n = 2610$ ). Water types are after Deutsch (1997).



**Fig. 9.** Piper trilinear plot of pre-drilling samples in the "Western Area" in four categories of sodium concentration. Circles in the diamond field are scaled to the concentration of TDS in the sample. The samples included here are known to be untreated and are within the  $\pm 7\%$  criteria on the ion charge balance ( $n = 3439$ ). Water types are after Deutsch (1997).

patterns have developed), most water wells are too shallow to encounter Na-HCO<sub>3</sub>/Na-Cl water types. In NE Pennsylvania about 88% of water from Carboniferous-age Burgoon Sandstone and Huntley Mountain formations are Ca-HCO<sub>3</sub> dominated water types. Most of these water wells occur in upland recharge areas. In contrast, the Devonian-age Lock Haven and Catskill formation samples

mostly occur in lowland areas (Shultz, 1999) and have more Na-Cl and Na-HCO<sub>3</sub> water types (combined at 22.1%) than the stratigraphically and topographically higher aquifers. The more frequent occurrence of Na-Cl groundwater type in the Lock Haven and Catskill formations is consistent with what Williams et al. (1998) found. The Lock Haven and Catskill formations produce groundwater of a Na-Cl type from naturally saline intervals that Williams et al. (1998) termed "restricted flow zones" that are not being continuously flushed by active circulation of fresh groundwater (Risser et al., 2013). In the Western Area freshwater flow systems are shallower due to smaller topographic relief, and wells drilled in both valley and upland settings are equally likely to encounter all water types. There, water type and water quality do not associate with geological formation (except where carbonates and coal formations occur, which add sulfate).

In NE Pennsylvania, lithium and sodium exceeding water-quality standards are more likely to be found in water wells drilled into the Catskill and Lock Haven formations than in others. Likewise, arsenic and lead exceeding standards are more likely to be found in water wells drilled into the Burgoon Sandstone and "Other" formations. Table 5 shows common water-quality exceedances related to samples collected from water wells completed into these formations in NE Pennsylvania. Detailed summary of the sample counts (and exceedances) for each parameter for each geological unit is provided in Table S8 in the Supplemental Information.

Table 6 provides a comparison of common water-quality guideline exceedances from water wells to their geological completion formations in the Western Area. Overall, water wells drilled into the Allegheny-Pottsville Formation have the poorest water quality followed by the Other formations and the Monongahela Group. A detailed summary of the sample counts (and exceedances) for each parameter for each geological unit is provided in Table S9 in the Supplemental Information.

The groundwater type (hydrochemical facies) is closely associated with water-quality exceedances. For purposes of this study we have categorized the water types based upon the dominant single major anion and cation. However, many of the water types actually consist of mixtures of anions and cations which often show as mixed water types such as calcium-sodium-bicarbonate (Ca-Na-HCO<sub>3</sub>) or Ca-Mg-HCO<sub>3</sub>.

Exceedances of water quality criteria also relate to specific geochemical water types. In NE Pennsylvania more than 50% of the pre-drilling samples with Na-Cl type groundwater exceed drinking-water guidelines for manganese, TDS, and sodium (Table 7). For calcium sulfate (Ca-SO<sub>4</sub>) type waters, 50% of samples exceed water-quality guidelines for iron, manganese, sulfate, sodium, and TDS. Table S10 in the Supplemental Information provides detailed sample counts (and exceedances) for each parameters compared to groundwater type.

All but one SMCL and RSL exceedance for chloride and strontium, respectively, occurred in either Na-Cl or Ca-Cl type groundwater. The majority of sulfate SMCL exceedances occur in CaSO<sub>4</sub> or Other water types. The Other water types (more than one sample) were magnesium bicarbonate (Mg-HCO<sub>3</sub>), magnesium sulfate (Mg-SO<sub>4</sub>), and sodium sulfate (Na-SO<sub>4</sub>). The Mg-HCO<sub>3</sub> is consistent with shallow recharging groundwater in carbonates, and the latter two types could be reflective of unusual local geochemical processes involving ion exchange, carbonate geochemistry, and/or sulfur oxidation.

In the Western Area (Table 8), iron, manganese, sulfate, and TDS exceed guidelines in over 50% of Ca-SO<sub>4</sub> type groundwater. Over 75% of the samples exceed guidelines for sodium in all water types except Ca-HCO<sub>3</sub> and Ca-Cl. The Ca-SO<sub>4</sub> water type is associated with groundwater sourced from coal-containing formations. The Other water types consist of Mg-HCO<sub>3</sub>, Mg-SO<sub>4</sub>, and Na-SO<sub>4</sub>. The



**Table 5**

Water-quality exceedances based upon water wells drilled into each geological unit – Northeastern Pennsylvania.

Percent of sample count that exceeds drinking water guideline or standard (%)	Burgoon Sandstone Formation (184)	Huntley Mountain Formation (387)	Catskill Formation (4983)	Lock Haven Formation (4780)	"Other" formations (33)
>1–5	Sodium	Arsenic, lead, sodium	Arsenic, barium, chloride, lead, pH < 6.5, pH > 8.5, TDS	Arsenic, barium, chloride, lead, pH < 6.5, pH > 8.5, sulfate	Lead
>5–10	Arsenic	–	–	TDS	Arsenic, lithium, sodium
>10–25	Lead, pH < 6.5	Manganese, pH < 6.5	Iron, lithium, manganese, turbidity	Lithium, turbidity	
>25–50	Turbidity, iron	Turbidity, iron	Sodium	Iron, manganese	Iron, pH < 6.5, turbidity
>50–75	Manganese	–	–	Sodium	Manganese
>75–100	–	–	–	–	–

Notes:

–: No parameters in this category.

Only parameters with exceedances greater than 1% of sample count included.

"Other" formations category includes: Mauch Chunk Formation, Timmers Rock Formation, Bloomsburg and Mifflintown formations, and Allegheny and Pottsville formations. Number in parenthesis after geological formation are total sample counts from each geological unit.

**Table 6**

Water-quality exceedances based upon water wells drilled into each geological unit – "Western Area".

Percent of sample count that exceeds drinking water guideline or standard (%)	Allegheny and Pottsville Formation (2952)	Conemaugh Group (2527)	Dunkard Group (440)	Monongahela Group (88)	"Other" formations (83)
>1–5	Arsenic, chloride, lead, lithium, pH < 6.5, pH > 8.5	Arsenic, chloride, lead, pH < 6.5, pH > 8.5, sulfate	Arsenic, chloride, pH < 6.5	Cadmium, chloride, pH < 6.5, pH > 8.5	Chloride, lead, pH < 6.5
>5–10	Sulfate	–	Lead, pH > 8.5, TDS	Lead, sulfate	–
>10–25	–	TDS, turbidity	–	Manganese, turbidity	Arsenic, sulfate
>25–50	TDS, turbidity,	Iron, manganese, sodium	Iron, manganese, sodium, turbidity	Iron, TDS	TDS
>50–75	Iron, manganese, sodium	–	–	Sodium	Iron, manganese, sodium, turbidity
>75–100	–	–	–	–	–

Notes:

–: No parameters in this category.

Only parameters with exceedances greater than 1% of sample count included.

"Other" formations category includes: Kanawha Formation; Ohio Shale; Berea Sandstone and Bedford Shale, undivided; and Maxville Limestone and Rushville, Logan, and Cuyhoga formations, undivided. Number in parenthesis after geological formation are total sample counts from each geological unit.

**Table 7**

Water-quality exceedances by groundwater type – Northeastern Pennsylvania.

Percent of sample counts in water type that exceeds drinking water guideline or standard (%)	Ca-HCO <sub>3</sub> (4582)	Na-HCO <sub>3</sub> (1591)	Na-Cl (413)	Ca-Cl (152)	Ca-SO <sub>4</sub> (98)	"Other" (29)
1–5	Arsenic, lead, pH < 6.5, TDS	Lead, TDS	Lead, pH < 6.5, strontium	Arsenic	Selenium	Chloride, pH < 6.5
>5–10	Barium	Arsenic, barium, pH > 8.5	pH > 8.5	Barium, chloride, lead	Arsenic, lead	–
>10–25	Iron, sodium, turbidity	Iron, turbidity	Arsenic	Turbidity, pH < 6.5, TDS	pH < 6.5	–
>25–50	Manganese	Manganese	Barium, chloride, iron, turbidity	Iron, manganese	Turbidity	Iron, manganese, sulfate, turbidity
>50–75	–	–	Manganese, TDS	Sodium	Iron, manganese, sulfate, TDS	Sodium, TDS
>75–100	–	Sodium	Sodium	–	Sodium	–

Notes:

–: No parameters in this category. Number in parenthesis is sample count in each category.

Only parameters with exceedances at least 1% of sample count included.

Only samples with ion charge balance within ±7% included.

"Other" groundwater types include: Mg-HCO<sub>3</sub>, Mg-Cl, Mg-SO<sub>4</sub>, and Na-SO<sub>4</sub>.

**Table 8**

Water-quality exceedances by groundwater type – “Western Area”.

Percent of sample count in water type that exceeds drinking water guideline or standard (%)	Ca-HCO <sub>3</sub> (4582)	Na-HCO <sub>3</sub> (1674)	Na-Cl (143)	Ca-Cl (236)	Ca-SO <sub>4</sub> (311)	“Other” (96)
1–5	Arsenic, lead, sulfate	Arsenic, lead, sulfate	Sulfate	Arsenic, lead, sulfate, lithium, pH < 6.5	Arsenic, lead	Arsenic, cadmium, chloride, pH > 8.5
>5–10	–	–	Barium, lead, pH < 6.5, pH > 8.5	–	Lithium	Lead
>10–25	TDS	Iron, manganese, pH > 8.5, TDS, turbidity	Lithium, turbidity	–	pH < 6.5	Lithium, pH < 6.5
>25–50	Turbidity, sodium	–	Iron, manganese	Chloride	Sodium	Iron, manganese, turbidity
>50–75	Iron, manganese	–	Chloride	Iron, manganese, turbidity	Sulfate, TDS, turbidity	Sulfate, TDS
>75–100	–	Sodium	Sodium, TDS	Sodium, TDS	Iron, manganese	Sodium

**Notes:**

–: No parameters in this category. Number in parenthesis is sample count in each category.

Only parameters with exceedances at least 1% of sample count included.

Only samples with ion charge balance within ±7% included.

“Other” groundwater types include: Mg-HCO<sub>3</sub>, Mg-Cl, Mg-SO<sub>4</sub>, and Na-SO<sub>4</sub>.

Mg-HCO<sub>3</sub> is consistent with shallow recharging groundwater in carbonates, and the latter two types could be reflective of unusual local geochemical processes involving ion exchange, carbonate geochemistry, and sulfur oxidation.

The specific number and percentages of water-quality exceedances for each groundwater type are provided in [Table S-11 in the Supplemental Information](#) for the Western Area.

### 5.5. Topographic position relationship to groundwater quality

#### 5.5.1. Northeast Pennsylvania

[Williams et al. \(1998\)](#) found that groundwater with higher concentrations of sodium, chloride, barium, and strontium was associated with water wells located in valley settings in NE Pennsylvania. [Table S12 in the Supplemental Information](#) provides a summary of groundwater type per topographic position for NE Pennsylvania and the Western Area.

In NE Pennsylvania, we found the large preponderance (85.7%) of the pre-drilling samples with Na-Cl water type occurring in water wells completed in valley settings with the next greatest water type percentage in valleys being Na-HCO<sub>3</sub> water type (53%). However, depending upon location, well depth, and specific fractures intersected, it is possible to intersect younger and local/intermediate flow systems that contain Ca-HCO<sub>3</sub> or Na-HCO<sub>3</sub> water types.

The pre-drilling samples in NE Pennsylvania show that samples exceeding water-quality guidelines for barium, chloride, and strontium (87.1%, 86.3%, and 100%, respectively) typically occur in water wells found in valley settings. In contrast, most sulfate exceedances (91.4%) occur in water wells completed in upland settings where coal beds are more prevalent and waters are better oxygenated.

Approximately 83% of the water wells drilled to depths greater than 61 m occurred in upland areas. [Table S-13 in the Supplemental Information](#) provides a summary of individual water-quality exceedances by topographic setting.

#### 5.5.2. Western area

For the Western Area, as noted in [Table S-12 in the Supplemental Information](#), there seems to be little relationship between water type and topographic position. Potable groundwater circulation in the Western Area extends only a few tens of

meters deep and most local to intermediate scale groundwater circulation paths are a few kilometers in length. Most water types occur at a percentage close to the overall percentage of samples in each topographic category (69.9% uplands, 30.1% valleys), but do not necessarily correlate. Na-Cl type groundwater type seems to be more likely in valley settings (37.4%), but the relationship between water quality and topography remains unclear, even with this large dataset.

There also is no substantial relationship between water-quality exceedances and topographic position for the Western Area, with the possible exception of barium and lithium. Barium exceedances of water-quality standards are higher in valley settings (63.0%) compared to upland settings (37.0%). Lithium exceedances are higher in upland settings (82.8%) than valley settings (17.2%). In the Western Area lower topography and dissected terrain coupled to coal seams leads to an inherently more complicated hydrochemical setting compared to NE Pennsylvania, and finding little relationship between water quality types with topographic settings is expected.

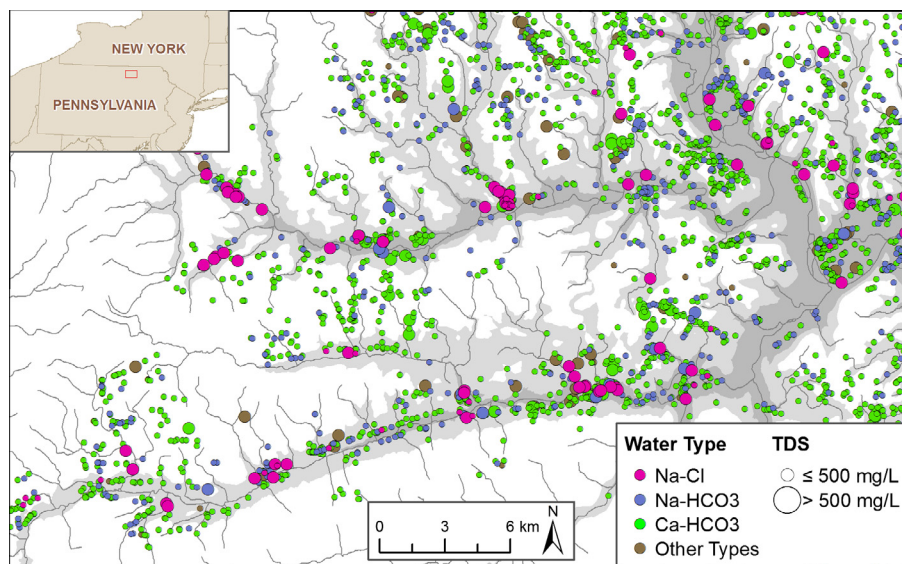
Approximately 91% of the water wells drilled to depths greater than 61 m occurred in upland areas. The specific number and percentages of water-quality exceedances for each parameter per topographic setting (upland versus valley) are provided in [Table S-14 in the Supplemental Information](#).

## 6. Discussion

Our results, based on a dataset that is orders of magnitude greater than previous studies on quality of shallow groundwater of Appalachia, agree broadly with the results of most prior studies. The general pattern of geochemical evolution with their commensurate exceedances of water-quality standards is clear. For example, Na-Cl water in NE Pennsylvania is associated with water-quality standard exceedances for many parameters. These exceedances occur in valley settings and low-lying areas.

[Fig. 10](#), as an example, shows that most (88.8%) of the Na-Cl groundwater type (pink dots) with higher total dissolved solids occur in major river valleys and associated tributaries.

In the Western Area, high salinity may occur in wells located both on valley slopes and even in upland areas where coal seams occur with oxidizable pyrite ([Razem and Sedam, 1985](#)). The combination of sulfide oxidation, ion-exchange, carbonate dissolution

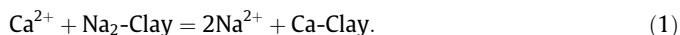


**Fig. 10.** Relationship between total dissolved solids (TDS) concentration, water type, and topography shown for a detail area in Bradford County, Northeastern Pennsylvania. TDS concentration is represented by symbol size. Water type (based on a simple classification using the dominant major anion and cation present in the sample) is represented by the color of the symbol as described in the legend. Streams and generalized topographic elevations are shown in gray for reference. Samples of Na-Cl and Na-HCO<sub>3</sub> water types are more prevalent in valleys and are associated with higher TDS concentrations.

and the intersection of deeper saline waters produce a broad range of water types. These processes are well known in the Appalachian Basin (Razem and Sedam, 1985). However, the same overall processes of chemical evolution should apply, as water types shift from Ca- to Na-dominated with increased age and residence time.

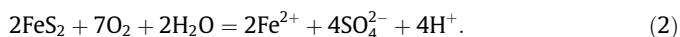
In both regions, the following reaction largely control the geochemical evolution of the groundwater.

Ion exchangeable clays in the aquifer rocks replace calcium and magnesium with sodium (Eq. (1)):

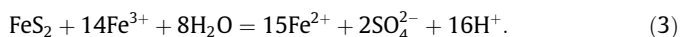


This reaction, when driven to completion, result in Na-HCO<sub>3</sub> geochemical facies waters. When these waters mix with saline water near the bottom of deep wells (or restricted flow zones) or in discharge zones in valleys, the water becomes enriched even more in sodium and chloride, typically resulting in a Na-Cl type water.

In the Western Area, in particular, the oxidation of pyrite in coal leads to the generation of sulfate and hydrogen ions from sulfuric acid (e.g. Eq. (2)):



This in turn leads to oxidation of the ferrous iron species to ferric iron in solution or as a solid iron hydroxide or oxyhydroxide (e.g. Eq. (3)):



The acid generated then dissolves more carbonate and other minerals in the aquifer, increasing the calcium and other metals in solution, causing higher dissolved solids consisting of sulfate from the pyrite oxidation and calcium and alkalinity from the carbonate minerals.

Our study confirms prior understanding of the natural evolution of water quality in shallow aquifers in the Appalachian Basin. Groundwater quality is affected by a number of factors resulting in local variability related to different combinations of geologic formation and topographic position of the water well which, in turn, relates to the underlying position of the well in local, intermediate, and regional flow systems. Most water wells are completed as

open holes and so may obtain water from multiple fractures, changing with hydrologic season and recharge events, which add to the variability of the groundwater chemistry.

In the future it may be possible to parse this large dataset into smaller regions to evaluate these relationships in more detail, either statistically or through geochemical modeling approaches. However, these exercises are beyond the scope of this paper, which was intended to re-evaluate previously identified water-quality variability using a much larger dataset. Overall, we see no broad changes in variability of chemical quality in this large dataset to suggest any unusual salinization caused by possible release of produced waters from oil and gas operations, even after thousands of gas wells have been drilled among tens of thousands of domestic wells within the two areas studied. What we find confirms what others have historically reported: there is wide variability in water quality in potable groundwater in the Appalachian Basin, and much of this water naturally exceeds regulatory standards. Those previous reports used much smaller datasets, and were published many decades earlier, thus confirming and supporting our conclusions.

## 7. Conclusions

The results of our study, from the analysis of 21,044 pre-drilling groundwater samples, have broad implications regarding the characterization of background chemistry in shallow groundwater supplies of the Appalachian Plateau, especially when assessing potential impacts from oil and gas operations. This very large dataset, in combination with previous studies, shows we can broadly conclude that in both NE Pennsylvania and the Western Area the groundwater commonly exceeds drinking-water guidelines, by means of natural processes. These exceedances are not random, but are related to factors such as geological formation/lithology, sample turbidity, age or residence time of the groundwater, and where within the groundwater flow system the sample was collected. There is also an association with groundwater type and TDS/salinity concentrations. In NE Pennsylvania there appears to be an association with water quality and topographic position (valleys versus uplands), whereas, in the Western Area this



relationship is not as strong. Based on our study, 63.1% of the water wells sampled in NE Pennsylvania and 88.7% in the Western Area had pre-drilling exceedances for drinking-water guidelines for one or more parameters (including turbidity). In NE Pennsylvania, 10.2% of the samples exceeded one or more of USEPA's MCLs set for drinking-water supplies, 46.1% of the samples exceeded one or more of USEPA's SMCLs, and 6.7% exceeded one or more of USEPA's HALs or RSLs for tap water. In the Western Area 7.7% of samples exceeded one or more MCLs, 65.2% exceeded one or more SMCLs, and 14.6% exceeded one or more health advisory or regional screening levels for tap water.

Our results are neither unusual nor surprising and are consistent with previous results in both areas using much smaller datasets both before and after unconventional oil and gas development (e.g. Molofsky et al., 2013; Gross and Low, 2012; Chambers et al., 2012; Low and Galeone, 2006; Williams et al., 1998; Stoner et al., 1987; Razem and Sedam, 1985; Taylor, 1984; Matisoff et al., 1982).

Comparison of historical groundwater data from NE Pennsylvania that mostly pre-dates unconventional shale gas development (pre-2007) to sampling results from Chesapeake's pre-drill baseline program (2009–2012) shows that the current groundwater quality conditions are similar to those shown in historical data. This comparison thus provides a dataset representative of regional groundwater conditions that pre-dated unconventional shale gas drilling, and provides a comprehensive understanding of aquifer systems and groundwater chemistry in the region.

Fundamentally, water-quality data from domestic water wells need to be understood within the natural geochemical context when evaluating suspected water-quality changes from shale-gas methane production or any other source. The hydrogeological setting of the water well also needs to be fully understood within the context of the mineralogy in the aquifer, the topographic position of the well within the associated flow path, the natural variability of water-quality parameters, and the naturally expected hydrochemical facies. Basic principles of groundwater flow and the hydrochemical evolution of groundwater systems in Appalachia (e.g. Piper, 1933; Lohman, 1937, 1939; Poth, 1963) indicate that saline groundwater underlying fresher groundwater is connate. This connate groundwater was present in the sediments during deposition, thus the saline waters have not migrated long vertical distances from formations hundreds and thousands of meters deeper.

We see no broad changes in variability of chemical quality in this large dataset to suggest any unusual salinization caused by possible release of produced waters from oil and gas operations, even after thousands of gas wells have been drilled among tens of thousands of domestic wells within the two areas studied. What we do find confirms what others previously found with much smaller datasets decades earlier regarding the wide variability of water quality in domestic groundwater in the Appalachian Basin, and that much of this water naturally exceeds regulatory standards.

Finally, caution should be exercised when using total metals analysis to characterize water-quality changes of any type, given the common problems associated with sampling water from domestic wells, which may include turbidity. Samples should be filtered to remove suspended particles and dissolved metals should be analyzed to provide a more accurate analysis of metals concentrations.

## Acknowledgements

The authors would like to thank Chesapeake Energy Corporation for providing funding for this study and for access to their pre-drilling groundwater dataset for Ohio, West Virginia,

and Pennsylvania. Special thanks goes to several anonymous reviewers who provided many useful comments to the manuscript. We also wish to thank the landowners who allowed their water wells to be sampled. Additional relevant information is provided below:

- Chesapeake Energy Corporation provided funding for the authors of this paper through their organizations of employment, and, in the case of the senior author (Don Siegel), privately, to do basic research to explore and evaluate this very large dataset, and for the preparation of the paper. Data used in this paper were collected on behalf of Chesapeake by paid third party consultants to comply with regulatory programs. The analysis and interpretations, and report writing, were conducted by the named authors of the paper. The decision to submit the paper was that of the authors. The opinions and conclusions expressed in this paper are those of the authors and do not necessarily reflect those of Chesapeake Energy.
- During the preparation of this specific paper, all authors worked for the organizations whose affiliations are noted in authorship. Mark Hollingsworth is a current employee of Chesapeake Energy Company, having worked at Chesapeake from February 2011 to present. Prior to Mr. Hollingsworth's employment by Chesapeake, he worked for TestAmerica Laboratories, Inc. which provided laboratory analytical consulting services to Chesapeake. Bert Smith is a former employee of Chesapeake Energy having worked there from May 2012 to September 2013, and has been employed by Enviro Clean Products and Services from November, 2013 to the present. Enviro Clean P&S also does consulting work for Chesapeake. Prior to May, 2013 Mr. Smith worked for Science Applications International Corporation (SAIC), which did consulting work for Chesapeake Energy. AECOM provides architecture and engineering services to government and private industry around the world, including the energy sector and Chesapeake. Rikka Bothun worked for AECOM during most of the time this paper was under preparation, but left AECOM in December, 2014 and now works for a private consulting company that does not do consulting work for Chesapeake.
- None of the following authors (Don Siegel, Bert Smith, Elizabeth Perry, or Rikka Bothun) have competing corporate financial interests exceeding guidelines presented by Applied Geochemistry. Mark Hollingsworth is a current employee of Chesapeake Energy and owns stock in the company in an amount in excess of \$5000.
- Donald Siegel is the lead author and contributed to the paper's preparation, technical interpretations, and review of these data and paper. Bert Smith is the second named author and contributed to the paper preparation, technical interpretations, and review of these data and paper. Elizabeth Perry is the third named author and contributed to the paper preparation, technical interpretations, and review of these data and paper. Rikka Bothun is the fourth named author and contributed to the paper preparation, technical interpretations, and review of these data and paper. Mark Hollingsworth is the fifth named author, oversees the Chesapeake baseline dataset, contributed to the paper preparation, and in review of these data and paper.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.apgeochem.2015.06.013>.

## References

- ATSDR, 12/28/2011. Health Consultation and Data Review. Dimock Pennsylvania <[http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&ved=0CEcQFjAF&url=http%3A%2F%2Fwww2.epa.gov%2Fsites%2Fproduction%2Ffiles%2Fdocuments%2Fdimock-atsdr.pdf&ei=5YM-VamJBXYQtQXqs4CIBA&usq=AFQjCNH4WgTJ6sVIDqzk\\_Z3DrDLEGtUNc&bvm=bv.91665533,d.b2w](http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&ved=0CEcQFjAF&url=http%3A%2F%2Fwww2.epa.gov%2Fsites%2Fproduction%2Ffiles%2Fdocuments%2Fdimock-atsdr.pdf&ei=5YM-VamJBXYQtQXqs4CIBA&usq=AFQjCNH4WgTJ6sVIDqzk_Z3DrDLEGtUNc&bvm=bv.91665533,d.b2w)> (accessed April 2015).
- Ayotte, J.S., Gronberg, J.A.M., Apodaca, L.E., 2011. Trace Elements and Radon in groundwater Across the United States, 1992–2003. USGS Scientific Investigations Report 2011–5059.
- Battelle, 2013. Bradford and Susquehanna County, Pennsylvania Retrospective Case Study Characterization Report. Technical Memo 4, Contract No. CON00011206. Report Submitted to the American Petroleum Institute and America's Natural Gas Alliance (February).
- Boyer, E.W., Swistock, M.S., Clark, J., Madden, M., Rizzo, D.E., 2012. Impact of Marcellus Gas Drilling on Rural Drinking Water Supplies. The Center for Rural Pennsylvania.
- Callaghan, T., Fleeger, G.M., Barnes, S., Dalberto, A., 1998. Groundwater flow on the Appalachian Plateau of Pennsylvania. Coal Mine Drainage Prediction and Pollution Prevention in Pennsylvania. Pennsylvania Department of Environmental Protection, pp. 2–1 to 2–39 (Chapter 2).
- Carswell, L.D., Bennett, G.D., 1963. Geology and Hydrology of the Neshannock Quadrangle, Mercer and Lawrence Counties, Pennsylvania. Pennsylvania Geological Survey, 4th Series, Bulletin W15.
- Castle, J.W., 2000. Recognition of facies bounding surfaces and stratigraphic patterns in foreland ramp successions: an example from the Upper Devonian, Appalachian Basin, USA. *J. Sediment. Res.* 70 (4), 896–912.
- Chambers, D.B., Kozar, M.D., White, J.S., Paybins, K.S., 2012. Groundwater Quality in West Virginia, 1993–2008. USGS Scientific Investigations Report 2012–5186.
- Chapman, E.C., Capo, R.C., Stewart, B.W., Kirby, C.S., Hammack, R.W., Schroeder, K.T., Edenborn, H.M., 2012. Geochemical and strontium isotope characterization of produced waters from Marcellus shale natural gas extraction. *J. Environ. Sci. Technol.*
- DeSimone, L.A., 2008. Quality of Water from Domestic Wells in Principal Aquifers of the United States, 1991–2004. USGS Scientific Investigations Report 2008–5227.
- DeSimone, L.A., Hamilton, P.A., Gillion, R.L., 2009. Quality of ground water from private domestic wells. National Groundwater Association. Water Well J.
- Deutsch, W.J., 1997. Groundwater Geochemistry: Fundamentals and Applications to Contamination. CRC Press, Boca Raton, FL.
- Engelder, 2012. Capillary Tension and Imbibition Sequester Frack Fluid in Marcellus Gas Shale. Letter to PNAS objecting to Publication of Duke Study on Brine Migration. Department of Geosciences, The Pennsylvania State University.
- Feth, J.H., 1970. Saline groundwater resources of the conterminous United States. *Water Resour. Res.* 6 (5), 1454–1457.
- Feth, J.H., 1981. Chloride in Natural Continental Water – A Review. U.S. Geological Survey Water-Supply Paper 2176, 30 p.
- Fleeger, G.M., 1999. The Geology of Pennsylvania's Groundwater. Pennsylvania Geological Survey, 4th Series, Educational Series 3.
- Fritz, S.J., 1994. A survey of charge-balance errors on published analyses of potable ground and surface waters. *National Groundwater Association. Groundwater J.* 32 (4), 539–546.
- Gross, E.L., Low, D.J., 2012. Arsenic Concentration, related Environmental Factors, and the Predicted Probability of Elevated Arsenic in Groundwater in Pennsylvania. USGS Scientific Investigations Report 2012–5257.
- Heisig, P.M., Scott, T.-M., 2013. Occurrence of Methane in Groundwater of South-Central New York State, 2012–Systematic Evaluation of a Glaciated Region by Hydrogeologic Setting. USGS Scientific Investigations Report 2013–5190.
- Hounslow, A., 1995. Water Quality Data: Analysis and Interpretation. CRC press, p. 416.
- Inners, J.D., 1981. Geology and Mineral Resources of the Bloomsburg and Mifflinville Quadrangles and Part of the Catawissa Quadrangle, Columbia County, Pennsylvania. Pennsylvania Geological Survey, 4th Series, Atlas 164cd, 152 p.
- Jenness, Jeff, Brian Brost, Paul Beier, 2013. Land Facet Corridor Designer Documentation. <[http://www.jennessent.com/downloads/Land\\_Facet\\_Tools.pdf](http://www.jennessent.com/downloads/Land_Facet_Tools.pdf)> (accessed 13.10.13).
- Kramer, D., 2011. Shale-gas extraction faces growing public and regulatory challenges. *Phys. Today* 64, 23–25.
- Lloyd Jr., Orville B., Lyke, William L., 1995. Ground Water Atlas of the United States: Illinois, Indiana, Kentucky, Ohio, Tennessee. USGS Hydrologic Atlas HA 730-K. <[http://pubs.usgs.gov/ha/ha730/ch\\_k/index.html](http://pubs.usgs.gov/ha/ha730/ch_k/index.html)>.
- Lohman, S.W., 1937. Groundwater in North-Eastern Pennsylvania. Pennsylvania Geological Survey, 4th Series, Water Resources Report 4, 312 p.
- Lohman, S.W., 1939. Groundwater in North-Central Pennsylvania. Pennsylvania Geological Survey, 4th Series, Water Resources Report 6, 219 p.
- Low, D.J., Galeone, D.G., 2006. Reconnaissance of Arsenic Concentrations in Groundwater from Bedrock and Unconsolidated Aquifers in Eight Northern-Tier Counties of Pennsylvania. USGS Open File Report 2006–1376.
- Matisoff, G., Khoury, C.J., Hall, J.F., Varnes, A.W., Strain, W.H., 1982. The nature and source of arsenic in northeast Ohio groundwater. *Groundwater J.* 20 (4).
- McAuley, S.D., Kozar, M.D., 2006. Ground-Water Quality in Unmined Areas and Near Reclaimed Surface Coal mines in the Northern and Central Appalachian Coal Regions, Pennsylvania and West Virginia. USGS Scientific Investigations Report 2006–5059.
- Molofsky, L.J., Connor, J.A., Farhat, S.K., Wylie Jr., A.S., Wagner, T., 2011. Methane in Pennsylvania water wells unrelated to Marcellus shale fracturing. *Oil Gas J.*
- Molofsky, L.J., Connor, J.A., Wylie, A.S., Wagner, T., Farhat, S.K., 2013. Evaluation of methane sources in groundwater in Northeastern Pennsylvania. National Groundwater Association. *Groundwater J.*
- Nicholson, Suzanne W., Dicken, Connie L., Horton, John D., Labay, Keith A., Foose, Michael P., Mueller, Julia A.L., 2007. Preliminary integrated Geologic Map Databases for the United States: Kentucky, Ohio, Tennessee, and West Virginia. USGS Open-File Report 2005–1324. GIS Shapefile available on line at <<http://pubs.usgs.gov/of/2005/1324/>>.
- Ohio Geologic Survey, 1998. Physiographic Regions of Ohio, Ohio Department of Natural Resources, Division of Geological Survey, paged-sized map with text, 2 p., scale 1:2,100,000.
- Ohio USEPA, 2012. Ohio Integrated Water Quality Monitoring and Assessment Report.
- Pennsylvania Department of Environmental Protection (PADEP), 2015. <[http://www.depreportingservices.state.pa.us/ReportServer/Pages/ReportViewer.aspx?/Oil\\_Gas/Wells\\_Drilled\\_By\\_Counties](http://www.depreportingservices.state.pa.us/ReportServer/Pages/ReportViewer.aspx?/Oil_Gas/Wells_Drilled_By_Counties)> (accessed March/April 2015).
- Pennsylvania Bureau of Topographic and Geologic Survey. Department of Conservation and Natural Resources, 2001. Bedrock Geology of Pennsylvania. GIS shapefile available on line at <<http://www.dcnr.state.pa.us/topogeo/map1/bedmap.aspx>>.
- Pennsylvania Department of Conservation and Natural Resources. PaGWIS Records Database. <<http://www.dcnr.pa.gov/topogeo/groundwater/pagwis/records/index.htm>> (accessed July 2014, October 2014, and April 2015).
- Piotrowski, R.G., Harper, J.A., 1979. Black Shale and Sandstone Facies of the Devonian Catskill Clastic Wedge in the Subsurface of Western Pennsylvania: Pennsylvania Department of Environmental Resources Bureau of Topographic and Geologic Survey under contract with the US Department of Energy, Contract Number EY-76-S05-5198.
- Piper, A.M., 1933. Groundwater in Southwestern Pennsylvania. Pennsylvania Geological Survey, 4th series, Bulletin W-1, 406 p.
- Poth, C.W., 1963. Geology and Hydrology of the Mercer Quadrangle, Mercer, Lawrence, and Butler Counties, Pennsylvania. Pennsylvania Geological Survey, 4th Series, Water Resource Report 16, 149 p.
- Poth, C.W., 1973a. Ground-Water Resources of Butler County, Pennsylvania. Pennsylvania Geological Survey, 4th Series, Water Resource Report 36, 49 p.
- Poth, C.W., 1973b. Ground-Water Resources of Washington County, Pennsylvania. Pennsylvania Geological Survey, 4th Series, Water Resource Report 38, 32 p.
- Poth, C.W., 1973c. Ground-Water Resources of Beaver County, Pennsylvania. Pennsylvania Geological Survey, 4th Series, Water Resource Report 39, 39 p.
- Puls, R.W., Barcelona, M.J., April 1996. Low-Flow (Minimal Drawdown) Ground-Water Sampling Procedures. USEPA/540/S-95/504.
- Razem, Allan C., Sedam, Alan C., 1985. Ground-water Quality and Geochemistry of Aquifers Associated with Coal in the Allegheny and Monongahela Formations, Southeastern Ohio. U.S. Geological Survey, Water-Resources Investigations Report 85–4034.
- Risser, D.W., Williams, J.H., Hand, K.L., Behr, R.A., Markowski, A.K., 2013. Geohydrologic and Water-quality Characterization of a Fractured-Bedrock Test Hole in an Area of Marcellus Shale Gas Development, Bradford Count, Pennsylvania. Pennsylvania Geological Survey, 4th Series, Open-File Miscellaneous Investigation 13–01.1, 49 p.
- Schlumberger, 2011. AquaChem Management Software for Water Quality and Groundwater Sampling Data.
- Seaber, P.R., Brahana, J.V., Hollyday, E.F., 1988. Region 20, Appalachian Plateaus and Valley and Ridge. In: Back, William, Rosenshein, J.S., Seaber, P.R. (Eds.), *Hydrogeology: The Geology of North America*, vol. 02. Geological Society of America, p. 189200.
- Shultz, C.H. (Ed.), 1999. The Geology of Pennsylvania. Pennsylvania Geological Survey Special Publication 1, 888 p.
- Siegel, D.I., Chamberlain, S.C., Dossert, W.P., 1987. The isotopic and chemical evolution of mineralization in seiparian concentrations: evidence of episodic paleohydrogeologic methanogenesis. *Geol. Soc. Am.* 99, 386–394.
- Siegel, D.I., Szustakowski, R., Frape, S., 1991. A regional evaluation of brine mixing in the Albion Group (Silurian) sandstones of New York, Pennsylvania, and Ohio. *Bull. Petrol. Geochim. Explor.* 6, 66–78.
- Siegel, D.I., Azzolina, N.A., Smith, B.J., Perry, A.E., Bothun, R.L., 2015. Methane concentrations in water wells unrelated to proximity to existing oil and gas wells in Northeastern Pennsylvania. *Environ. Sci. Technol. J. March*, 4106–4112.
- StataCorp., 2009. Stata Statistical Software: Release 11. StataCorp LP, College Station, TX.
- Stoner, J.D., Williams, D.R., Buckwalter, T.F., Felbinger, J.K., Pattison, K.L., 1987. Water Resources and the Effect of Coal Mining, Greene County, Pennsylvania. Pennsylvania Geological Survey Water Resource Report 63.
- Taylor, L.E., 1984. Groundwater Resources of the Upper Susquehanna River Basin, Pennsylvania. Pennsylvania Geological Survey and Susquehanna River Basin Commission, Water Resource Report 58.
- Trapp Jr., H., Horn, M.A., 1997. Groundwater Atlas of the United States, Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, Virginia, West Virginia. U.S. Geological Survey Hydrologic Atlas HA 730-L. <<http://pubs.usgs.gov/ha/ha730/gwa.html>> (accessed 07.06.13).
- U.S. Environmental Protection Agency (USEPA, 2014). <<http://water.USEPA.gov/scitech/methods/cwa/>> and <<http://www.USEPA.gov/osw/hazard/testmethods/sw846/online/index.htm>> (accessed April 2014).

- U.S. Environmental Protection Agency (USEPA), 2015a. Drinking Water Contaminants. <<http://water.USEPA.gov/drink/contaminants/#List>> (accessed January 2015).
- U.S. Environmental Protection Agency (USEPA), 2015b. Regional Screening Levels. <[http://www.USEPA.gov/reg3hwmd/risk/human/rb-concentration\\_table/Generic\\_Tables/docs/restap\\_sl\\_table\\_run\\_MAY2014.pdf](http://www.USEPA.gov/reg3hwmd/risk/human/rb-concentration_table/Generic_Tables/docs/restap_sl_table_run_MAY2014.pdf)> (accessed January 2015).
- U.S. Geological Survey (USGS), January 2006. Methane in West Virginia Groundwater. USGS Fact Sheet 2006-3011.
- U.S. Geological Survey (USGS), 2012. National Geological Map Database. <[http://ngmdb.usgs.gov/Geolex/geolex\\_qs.html](http://ngmdb.usgs.gov/Geolex/geolex_qs.html)> (accessed 06.06.13).
- U.S. Geological Survey (USGS), 2014. Appalachian Plateaus Groundwater Availability Study. Information Sheet. <<http://va.water.usgs.gov/appalachianplateaus/concept.html>> (accessed April 2015).
- Vidic, R.D., Brantley, S.L., Vandenbossche, J.M., Yoxheimer, D., Abad, J.D., 2013. Impact of shale gas development on regional water quality. *Science* 340 (6134).
- Warner, N.R., Jackson, R.B., Darrah, T.H., Osborn, S.G., Down, A., Zhao, K., White, A., Vengosh, A., 2012. Geochemical evidence of possible natural migration of Marcellus Formation brine to shallow aquifers in Pennsylvania. *Proc. Natl. Acad. Sci. Environ. Sci.*
- Weiss, Andrew, 2001. Topographic Position and Landforms Analysis. Poster presentation, ESRI User Conference, San Diego, CA. Available, by permission from the author, at <[http://www.jennessent.com/arcview/TPI\\_Weiss\\_poster.htm](http://www.jennessent.com/arcview/TPI_Weiss_poster.htm)>.
- White, J.S., Mathes, M.V., 2006. Dissolved-Gas Concentrations in Groundwater in West Virginia, 1997–2005. USGS Data Series 15.
- Williams, J.H., 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. USGS Scientific Investigations Report 2010-5224.
- Williams, J.H., Taylor, L.E., Low, D.J., 1998. Hydrogeology and Groundwater Quality of the glaciated Valleys of Bradford, Tioga, and Potter Counties, Pennsylvania. USGS and Pennsylvania Department of Conservation and Natural Resources, Water Resource Report 68.
- Wilson, B., 2014. Geological and baseline groundwater evidence for naturally occurring, shallowly sourced, thermogenic gas in northeastern Pennsylvania. *AAPG Bull.* 98 (2), 373–394.
- Wunsch, D.R., 1995. Hydrochemical facies model for dissected, coal-bearing strata in the Appalachian Coal Field [abs.]: GSA Abstracts with Programs, vol. 27, no. 6, p. A-97, New Orleans, LA.