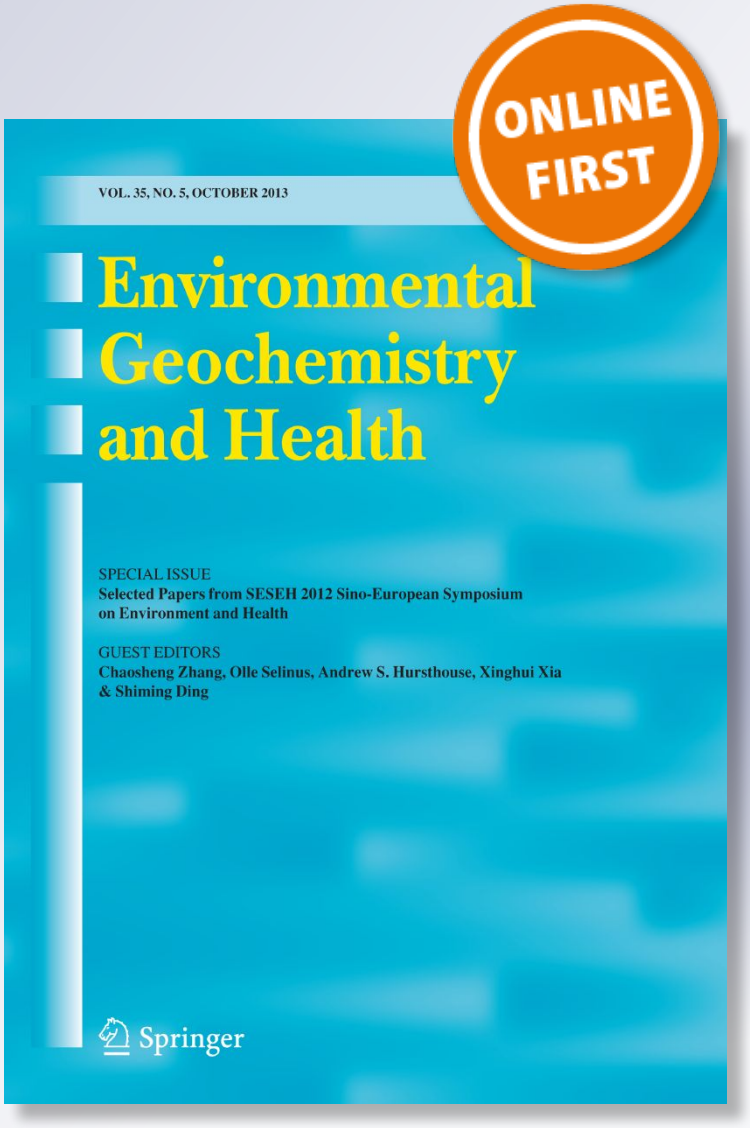
Detecting the effects of coal mining, acid rain, and natural gas extraction in Appalachian basin streams in Pennsylvania (USA) through analysis of barium and sulfate concentrations

Xianzeng Niu, Anna Wendt, Zhenhui Li, Amal Agarwal, Lingzhou Xue, Matthew Gonzales & Susan L. Brantley

Environmental Geochemistry and Health
Official Journal of the Society for Environmental Geochemistry and Health

ISSN 0269-4042

Environ Geochem Health
DOI 10.1007/s10653-017-0031-6



Springer

Environ Geochem Health
DOI 10.1007/s10653-017-0031-6

ORIGINAL PAPER

Detecting the effects of coal mining, acid rain, and natural gas extraction in Appalachian basin streams in Pennsylvania (USA) through analysis of barium and sulfate concentrations

Xianzeng Niu · Anna Wendt · Zhenhui Li · Amal Agarwal · Lingzhou Xue · Matthew Gonzales · Susan L. Brantley

Received: 7 December 2016 / Accepted: 30 September 2017

Abstract To understand how extraction of different energy sources impacts water resources requires assessment of how water chemistry has changed in comparison with the background values of pristine streams. With such understanding, we can develop better water quality standards and ecological interpretations. However, determination of pristine background chemistry is difficult in areas with heavy human impact. To learn to do this, we compiled a master dataset of sulfate and barium concentrations ($[SO_4]$, [Ba]) in Pennsylvania (PA, USA) streams from publically available sources. These elements were chosen because they can represent contamination related to oil/gas and coal, respectively. We applied changepoint analysis (i.e., likelihood ratio test) to identify pristine streams, which we defined as streams with a low variability in concentrations as measured over years. From these pristine streams, we estimated the baseline concentrations for major bedrock types in PA. Overall, we found that 48,471 data values are available for $[SO_4]$ from 1904 to 2014 and 3243 data for [Ba] from 1963 to 2014. Statewide $[SO_4]$ baseline was estimated to be 15.8 ± 9.6 mg/L, but values range from 12.4 to 26.7 mg/L for different bedrock types. The statewide [Ba] baseline is 27.7 ± 10.6 µg/L and values range from 25.8 to 38.7 µg/L. Results show that most increases in $[SO_4]$ from the baseline occurred in areas with intensive coal mining activities, confirming previous studies. Sulfate inputs from acid rain were also documented. Slight increases in [Ba] since 2007 and higher [Ba] in areas with higher densities of gas wells when compared to other areas could document impacts from shale gas development, the prevalence of basin brines, or decreases in acid rain and its coupled effects on [Ba] related to barite solubility. The largest impacts on PA stream [Ba] and $[SO_4]$ are related to releases from coal mining or burning rather than oil and gas development.

Keywords Water quality · Human impact · Shale gas · Historical data · Pristine river

Electronic supplementary material The online version of this article (doi:10.1007/s10653-017-0031-6) contains supplementary material, which is available to authorized users.

X. Niu (✉) · A. Wendt · M. Gonzales · S. L. Brantley
Earth and Environmental Systems Institute, Penn State University, EES Building, 2217, University Park, PA 16802, USA
e-mail: xzniu@psu.edu

Z. Li
College of Information Sciences and Technology, Penn State University, University Park, PA 16802, USA

A. Agarwal · L. Xue
Department of Statistics, Penn State University, University Park, PA 16802, USA

Published online: 13 October 2017

Introduction

When a new land use activity occurs in an already-contaminated environment, incidents of contamination can be difficult to document unless the pre-existing conditions have been analyzed. For example, waves of mineral or energy extraction activity have impacted water resources in Pennsylvania (PA), USA, along with agriculture and municipal development (Sams III and Beer 2000). In recent years, a new activity, natural gas extraction from shale, has grown in importance throughout PA and concerns have arisen over the use of high-volume hydraulic fracturing to extract shale gas from the Marcellus and other formations that may harm water quality (Brantley et al. 2014). To understand impacts on water resources requires assessment of the background values of pristine streams as measured prior to impact. With such an assessment, we are enabled to develop appropriate water quality standards and better ecological interpretations. However, determining such conditions is difficult because: (1) water data are spatially and temporally sparse; (2) climate conditions and rain chemistry are highly variable; (3) contaminants can derive from multiple sources at different times; (4) the distributions of bedrock, drainage patterns, and land use are highly heterogeneous; and (5) measurement protocols and capabilities have varied with time. Some of these issues, which can be particularly problematic for regional studies, have been previously discussed in the literature (Kirby et al. 2008; Olson and Hawkins 2012; Smith et al. 2003).

During the early 1960s, the US. Geological Survey (USGS) established the Hydrologic Benchmark Network (HBN) to monitor changes in the flow and water quality of "minimally disturbed" streams and rivers cross the USA (Alexander et al. 1998). The dataset collected from the HBN has been used as a reference for distinguishing natural from human-induced changes in river ecosystems or as baseline information for modeling natural background concentrations of nutrients and other chemical elements (Alexander et al. 1998; Olson and Hawkins 2012; Smith et al. 2003). However, because of budgetary restrictions, it is impossible to cover all climate and geological regions in the HBN. For example, only one observation station from the HBN is located in PA (i.e., Young Womans Creek). This lack of stations is problematic because water chemistry varies more spatially than temporally due to heterogeneity of minerals in the bedrock (Rhodes and Horton 2015).

To explore how to determine pristine background in a highly impacted area, we selected Pennsylvania (PA, USA) as a study region and analyzed two solutes, sulfate (SO_4) and barium (Ba), using publically available historical water quality data. Pennsylvania has been the site of development of charcoal, conventional natural gas and oil, and now shale gas, and each energy source has created different environmental issues that often mark the waterways of the state. In addition, Pennsylvania has highly heterogeneous bedrock and highly variable climatic conditions in addition to a variety of other land uses and thus represents a difficult case study for the determination of impacts on surface water quality. We sought to determine the baseline water quality in PA with respect to concentration of sulfate, [SO_4], and barium, [Ba], to learn what approaches can be used to establish the pristine chemistry in surface waters in a highly populated region with more than 100 years of intense and variable land use. We follow previous researchers in using a combination of spatial and statistical analysis to assess background chemistry of natural surface waters (e.g., Kim et al. 2015; Rhodes and Horton 2015; Voutchkova et al. 2014). With respect to sulfate, our investigation is similar to that of Raymond and Oh (2009). Those researchers explored the long-term impacts of acid mine drainage on stream chemistry of three heavily coal mining impacted watersheds of PA: here, we extended the analysis to the entire state.

The analytes chosen for focus, barium (Ba) and sulfate (SO_4), are indicators for activities related to oil/gas and coal, respectively. The most important use of Ba worldwide is in drilling muds for oil and gas development, and the use of Ba thus tends to track the oil and gas industry (Hanor 2000). In addition, Ba is often present in high concentrations in sedimentary basin brines and thus has been identified as a "fingerprinting" element that can document contamination by flowback or production waters that return to the surface during natural gas extraction (Brantley et al. 2014). For example, brines in the Appalachian basin of Pennsylvania contain high concentrations of Ba and distinctive ratios of concentrations of other ions including Na, Cl, Ca, Mg, Ba, Sr, and Br.

When collected at the surface during gas extraction, these brines are treated as waste waters that sometimes

contaminate natural waterways because of spills, leaks, or permitted discharges to streams or soils (Brantley et al. 2014). However, the brines also flow naturally to the surface in some parts of the state where they are highly diluted by surface waters (Adams 2011; Hladik et al. 2014; Llewellyn 2014; Poth 1962). Other natural sources of Ba also abound because Ba occurs in the Earth's continental crust as the 12th to 15th most common element, generally increasing with increasing Si, K, and Ca (Hanor 2000). Hanor (2000) reports that most of the Ba on earth is present in potassium feldspar or micas, and to a lesser extent in Ca silicates. These minerals thus naturally leach Ba into freshwaters (Hanor 2000). Turekian (1977) reports the average Ba concentration in freshwaters ([Ba]) equals 0.020 mg/L worldwide. In some ground or surface waters, [Ba] can be limited by the solubility of the most common Ba mineral, barite ($BaSO_4$) (Cravotta III 2008; Hanor 2000). [Ba] in natural waters can also be controlled by sorption equilibria with clays (Hanor and Chan 1977; Hanor 2000). The US. Environmental Protection Agency drinking water standard for Ba is 2 mg/L because it can cause muscular and gastrointestinal problems when ingested at high concentrations.

In contrast to Ba, sulfate (SO_4) is a much more ubiquitous and concentrated component of natural waters because sulfur is the 17th most common element on Earth and because the commonly formed sulfate-containing phases have generally high solubility. Sulfate is also a common constituent of pristine waters and is present in freshwaters worldwide at approximately 11 mg/L (Turekian 1977). Sulfate is a major contaminant due to acid mine drainage or acid rain (Rodhe et al. 2002), both of which have heavily impacted rivers in Pennsylvania (Kirby et al. 2008; Raymond and Oh 2009). As such, sulfate is a good indicator species related to coal mining and coal burning (Adams 2011; Hladik et al. 2014). The US. Environmental Protection Agency has set a secondary drinking water standard for sulfate as 250 mg/L.

One reason to study the two analytes, Ba and SO_4, together is because they are coupled in natural waters due to the low solubility of barite, $BaSO_4$ (Cravotta III 2008; Hanor 2000). As discussed by Hanor (2000), [Ba] and [SO_4] have been observed in some natural waters to occasionally be 2X supersaturated with respect to barite (Hanor and Chan 1977), but in general the mineral precipitates easily and higher supersaturation values are not expected. To understand long-term variability of these analytes thus requires analysis of the two solutes together.

The main goal of this study is to determine the baseline [Ba] and [SO_4] values in PA surface water using publically available historical water quality data and hence detect the effects of coal mining, acid rain, and natural gas extraction on PA surface water quality. Specifically, we attempted to: (1) compile a master dataset of [Ba] and [SO_4] from public databases for all PA; (2) discover temporal and spatial patterns of [Ba] and [SO_4]; (3) explore methods to determine "pristine" streams/rivers; (4) determine the baseline values of [Ba] and [SO_4] in PA pristine streams; and (5) assess the human impacts where possible—especially coal mining and gas/oil extraction—based on changes of [Ba] and [SO_4] from their respective baseline values. The impacts of coal mining have been previously documented for a few small PA watersheds (Raymond and Oh 2009). This approach is reproduced and extended here for the entire state. These efforts demonstrate the robustness of the approach and the utility of the publicly available data.

Materials and methods

Data sources

To study the background values of sulfate and barium concentrations in PA rivers, we created a master water chemistry dataset for Ba and SO_4 by compiling publically available water quality data for all time periods over PA. The data sources include the following online databases and publications: the USGS National Water Information System (http://waterdata.usgs.gov/nwis), the Susquehanna River Basin Commission database (http://www.srbc.net/), the EPA STORET Data Warehouse (https://www.epa.gov/waterdata/storage-and-retrieval-and-water-quality-exchange), and the Shale Network database (www.shalenetwork.org, doi:10.4211/his-data-shalenetwork). The Shale Network database is a newly developed data warehouse that contains water chemistry data contributed from six universities, eight government entities, 41 volunteer groups, and 11 private companies and can be accessed online at www.cuahsi.org.

Spatial data needed for this study, such as maps of PA major rivers, bedrock types, and coal mining areas were derived from the Pennsylvania spatial data access website (PASDA, http://www.pasda.psu.edu/). Locations of conventional and unconventional natural gas wells data were obtained from the Department of Environmental Protection reporting service (http://www.depreportingservices.state.pa.us/ReportServer/Pages/ReportViewer.aspx?/Oil_Gas/Permits_Issued_Detail). Coal production data were compiled from the US. Energy Information Administration (EIA, http://www.eia.gov/beta/) and the Pennsylvania mining and mineral resources information website (Penn State University library, http://psu.libguides.com/PAMinesandMining). Sulfate concentrations in precipitation were downloaded from the National Atmospheric Deposition Program website for sampling locations in the state (http://nadp.sws.uiuc.edu/data/NTN/).

In this study, the Ba and SO_4 data were only utilized from providers who maintain quality control and assurance protocols (see supplementary Table A.1 in Online Resources 1). Assessing error in long-term datasets derived from multiple providers can be difficult or impossible. For example, error in the measurements varies depending upon provider and is not always described appropriately. We have therefore assumed that an adequate estimate of error in measurement of [Ba] is ± 5% for all measurements above the censor (or reporting) limits. The censor limits in our compiled dataset include 500 (1987–1988), 100 (1972–1979), 50 (1981), 20 (2006), and 10 (1998). All units are in μg/L. Data measured in earlier time periods are likely to have larger error bars than more recent measurements since access to analytical instrumentation has improved. Data are therefore presented here assuming accuracy to ± 5% for data since 1985 (onset of standard use of inductively coupled plasma atomic emission or mass spectrometry) and ± 10% for data measured prior to 1985. For sulfate, we assume accuracy of ± 10% for all data. As shown later in this paper, the relative standard deviations estimated for baseline concentrations of the two analytes are larger than these estimated error terms, reflecting that most of the variation in our final baseline assessments are related to temporal variability of stream concentrations rather than measurement error.

Filtered versus unfiltered and censored versus uncensored data

In our compiled master dataset, two kinds of data are available: filtered and unfiltered. Unfiltered means that concentrations were analyzed from water samples directly collected from rivers, while filtered refers to water samples that were filtered to remove suspended fine particles. For example, unfiltered [Ba] represent total barium, including suspended and dissolved concentrations, while filtered represents only dissolved barium. In this study, only data measured from filtered water samples were used. The filtered concentrations account for 99% of the total data we collected for sulfate and 96% for barium. Where filtered and unfiltered samples were reported, we generally observed that concentrations from unfiltered samples were greater than or equal to that from filtered samples.

Censored and uncensored data are indicators of data quality control. Some laboratory results were reported as "censored data" because their values were either below the minimum or above the maximum detection limits. Generally, the "censored" data were reported as "greater than" or "less than" the detection limits instead of as measured values. In this study, only uncensored data were used for analysis. The uncensored data account for 99.6 and 92.7% of the total filtered data for SO_4 and Ba, respectively. Given the small fraction of censored data, exclusion of these censored data is not a problem for the analysis.

Delineating coal mining areas and natural gas/oil well drilling areas

Coal mining and natural gas extraction are two major energy production activities that have been known to impact surface water quality in some areas of PA (Brantley et al. 2014; Kirby et al. 2008; Raymond and Oh 2009; Vidic et al. 2013). We divided the state into areas with coal mining activities (COAL) and areas without coal mining activities (NON-COAL), areas with natural gas/oil extraction (i.e., well drilling, WELL) and areas without gas/oil well drilling activities (NON-WELL).

Given the fact that coal mining is a non-point source of pollution which could impact large surrounding areas, we chose storm-management watersheds as the basic unit in this study to delineate COAL

and NON-COAL areas. We first calculated the density of coal mining sites (including active, inactive, and abandoned coal mining sites) for each watershed (i.e., total number of coal mining sites divided by the area of each respective watershed). We did not take into account the actual area of coal mines because we did not have access to such data. COAL areas are therefore determined here as watersheds with a density of individual coal mines that is greater than 1 mine site per 100-km^2. This criterion was chosen to make sure that any watershed with relatively dense coal mining sites was categorized as COAL area. The rest were considered as NON-COAL areas (see Fig. A.1 in Online Resources 1).

Similarly, we calculated the density of gas/oil wells (including both conventional and unconventional wells) for each watershed, and defined the WELL area as watersheds with well density greater than 1 well per 100-km^2, and the rest as NON-WELL areas (see Fig. A.2 in Online Resources 1).

Determining the pristine rivers

A "pristine" river is defined in this study as any river that has maintained relatively constant barium and sulfate concentrations for a long period of time. In other words, the pristine river must meet two criteria: (1) chemical concentrations are stable (i.e., experiencing little to no human-induced impacts), and (2) the river remains "stable" for a long time. Based on the availability of data, we set a minimum of 10- and 5-year as the temporal criteria for sulfate and barium, respectively. Considering the fact that natural processes, such as bedrock weathering, ground water fluctuations, and changes in climate conditions also result in changes in chemical concentrations, we defined "stable" as varying in chemical concentration within a limited extent. An implicit assumption behind our approach is that riverine chemistry that is stable at the 5–10 year timescale is our best estimate of baseline chemistry not impacted by anthropogenic activity.

To find the relatively stable rivers, we first screened our dataset and selected rivers that have a relative standard deviation (ReStDev, i.e., the ratio of standard deviation normalized by the long-term mean) $\leq$ 30% and defined this as a potential pristine river pool. Rivers with ReStDev greater than 30% were assumed to be "obviously" contaminated at some point in time by human activities.

In the second step, with rivers from the potential pristine river pool (i.e., ReStDev $\leq$ 30%), we ranked the rivers by the observed standard deviations (StDev) for [SO_4] and [Ba] data, respectively. We then applied the changepoint analysis (i.e., likelihood ratio test) to detect the changepoint that divided the rivers into low variation (i.e., low StDev) in [SO_4] or [Ba] concentrations and high variation (i.e., high StDev) groups. Any rivers with a StDev smaller than the changepoint were considered to be "stable" in chemical concentrations and were therefore assumed to be pristine rivers. Detailed changepoint analysis methodology can be found in Online Resources 1.

Determining baseline values of [SO_4] and [Ba]

Here, baseline values are defined as the chemical contents of "pristine" rivers that were determined by the approach discussed in previous sections. There are many factors affecting the surface water baseline values. These factors include bedrock type, groundwater, land use, vegetation, topography, and climate conditions (Olson and Hawkins 2012). Previous studies suggested that, among other environmental factors, bedrock chemistry is the most important variable predicting stream chemistry (Olson and Hawkins 2012). To simplify the analysis, we therefore assumed that bedrock is the first-order factor controlling baseline chemical concentrations and attempted to determine the baseline [SO_4] and [Ba] values for the major bedrock types in PA. Using GIS (Geographic Information System) techniques, we first regrouped the PA bedrock types into six major bedrock groups based on the property of each bedrock type. We then overlaid the major bedrock type layers with the layer describing locations of the pristine rivers to determine the pristine rivers for each major bedrock type. The means (or medians) of [SO_4] and [Ba] for each major bedrock type were calculated and assigned to each major bedrock group as their first-order baseline estimates.

Human impact analysis

Assuming all other natural characteristics (e.g., bedrock, climate, topography, etc.) are equal, differences between the measurements of sulfate or barium

concentrations and their respective baseline values could be good indicators of the impacts of human activities such as coal mining and gas/oil extractions. Accordingly, we created maps of sulfate and barium changes from the respective baseline values to visualize spatial patterns of changes to determine potential correlations with human activities such as coal mining and natural gas-exaction in PA. Wilcoxon Signed Rank test (for medians) and *t* test (for means) were conducted to study temporal trends and spatial patterns of the historical sulfate/barium data and the correlation with energy production activities (e.g., coal production, gas extraction) and environmental policy changes such as the United States Clean Air Act of 1970 (Faoro and McMullen 1977).

Results and discussion

Sulfate data availability

The dataset for filtered surface water [SO_4] in PA is available from 1904 to 2014 (Fig. A.3 in Online Resources 1). We found a total of 48,471 reported measurements overall for 3315 observation sites from 1490 streams. We found only one measurement (in 1904) before 1920 and a few during the 1920s and 1930s. Measurements started to increase from the mid-1940s and peaked during the late 1960s and early 1980s. A second data peak appeared after year 2000. The temporal variation in data availability is likely related to phenomena such as (1) the public perception of water quality problems; (2) the availability of research funding; and/or (3) changes in environmental policies such as the Clean Air Act.

Temporal and spatial patterns of sulfate concentrations

Sulfate concentrations were analyzed for different time periods as shown in Fig. 1. Overall, the sulfate concentrations were low during the first four decades of the 1900s when median values ranged from 12 to 13 mg/L. The sulfate concentrations dramatically increased and peaked during the 1940s and 1950s (with a median value of 61 mg/L, $p < 0.05$), which corresponds to the peak time of coal production in PA (Raymond and Oh 2009). When the Clean Air Act was introduced in 1970, river [SO_4] values started to decline (the median equaled 39 mg/L for the time period of 1961–1982), and after the 1990s, the SO_4 concentration dropped to a median value of 15.2 mg/L during the first decade of the 2000s. More detailed statistics of SO_4 concentrations in different time periods can be found in supplementary Table A.2 in Online Resources 1.

Spatial patterns of sulfate data availability and concentration levels are shown in Fig. 2. Coal mining areas are shown in gray background. Sulfate observation sites are shown as (red) dots where the concentration levels are indicated by the size of the dot (larger symbols represent higher sulfate concentrations).

Only one analysis was found for 1900–1920, and this value of [SO_4] was observed in Lancaster County in the southeast part of PA (Fig. 2a). During the 1920s and 1930s, a total of 149 measurements were available from 119 observation sites across the entire state (Fig. 2b), with a more random pattern showing on the map. More than four thousand data from 143 observation sites were available during the 1940s and 1950s, and the [SO_4] levels were clearly higher in coal mining areas than other areas (Fig. 2c). More intense observations occurred during the 1960s and 1970s (Fig. 2d) with a total of 20,807 data available from 1652 observation sites across the State. During this time period, it is even more evident that [SO_4] were higher in gray-colored coal mining areas than in non-coal mining areas. During the 1980s and 1990s, observations were clustered in three "hot-spots" of coal mining and two non-coal mining areas in the northwestern and southeastern region of PA (Fig. 2e). The highest variations were also observed during this time period (165 $\pm$ 1557.2 mg/L, ranging from 0.02 to 90,100 mg/L, Table A.2). A total of 16,322 data values were available from 649 observation sites from 2002 to 2014. Overall, [SO_4] dropped dramatically during the time period of 2002–2014 (median value decreased by 55%, $p < 0.05$) when compared with previous time period (Figs. 1, 2f).

Coal production versus sulfate concentration

As discussed in the previous section, sulfate concentrations highly correlate with locations of coal mining as shown in Fig. 2. The correlation has also been observed in a previous study using data from a small watershed in Pennsylvania (Raymond and Oh 2009). Here, we attempted to compare the correlations in both

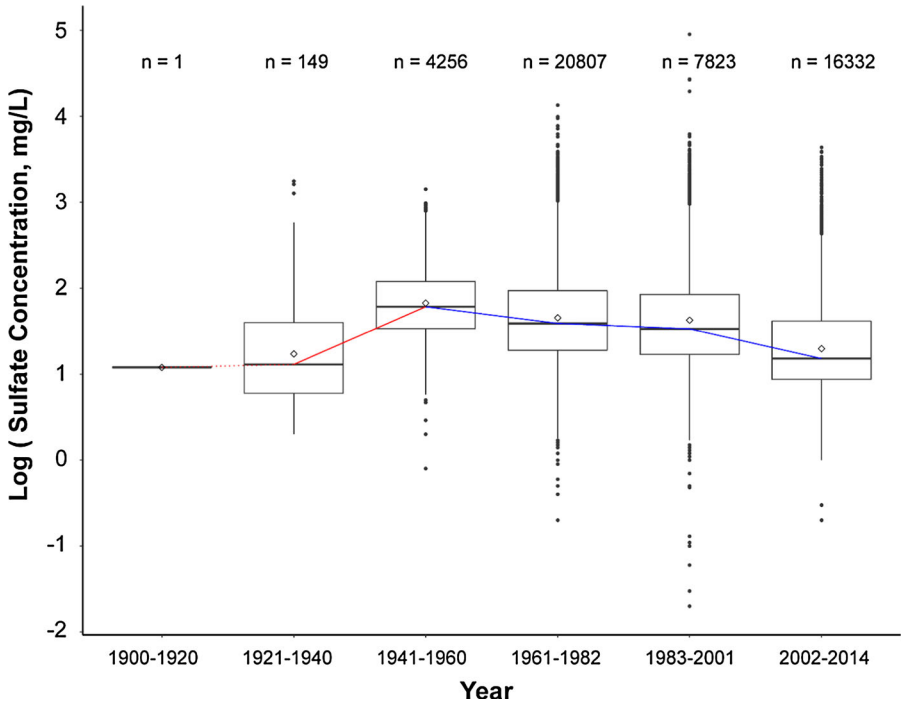


Fig. 1 Box plots summarizing all reported sulfate concentrations from Pennsylvania rivers for different time periods. Each box delineates the 1st (Q1) and 3rd (Q3) quartiles. The vertical lines beyond the box represent upper [Q1 − 1.5(Q3–Q1)] and lower inner fence [Q3 + 1.5(Q3–Q1)], and the solid dots beyond the fences represent outliers. Diamond symbols are means. Lines inside the box show the medians. Solid lines between boxes are drawn to show when differences in medians are statistically significant as determined using Wilcoxon Signed Rank test at level $\alpha = 0.05$. Red lines indicate increase between boxes and blue indicates decrease. When lines are drawn as dotted, the differences between boxes are statistically insignificant. Number of data points (n) are indicated above each time period. Means, medians, and quartiles were determined on data before log (10) transformation

COAL and NON-COAL regions of PA. To do that, we screened the master dataset to select sites where at least 10 years continuous data for *both* sulfate and barium were available. We chose sites with both sulfate and barium measured simultaneously because we also wanted to analyze the interaction between the two (see details in the section of "Sulfate and barium interaction" below). A total of eight sites, all from the USGS data source, met our criteria and this set is termed here the "continuous" set of sites (see supplemental Fig. A.4 in Online Resources 1 for the locations of the selected sites). Three sites from three streams were located in the western (West) region of PA, representing the intensive coal mining areas. Four sites from three streams are in the southeastern part of PA (East), representing a region with almost no coal mining nor oil/gas wells. One site is located in the northern central PA (Center) in Clinton County (i.e., Young Womans Creek). Young Womans creek is the only site included in the USGS Hydrological Benchmark Network as an "undisturbed" stream in PA.

The trends of regional means of sulfate concentrations for these continuous subsamples of sites are shown in Fig. 3. The results show that the sulfate concentrations dramatically decreased in the West region of PA, especially between 1965 and 1985. This trend is coincident with the decline of total PA coal production (Fig. 3). A strong positive linear relationship ($p < 0.001$) between coal production and river sulfate concentrations in the West (see Fig. A.5 in Online Resource 1) is in agreement with the work of Raymond and Oh (2009). In contrast to the West data, however, no significant correlation ($p > 0.05$) between coal production and [SO_4] were found in the East and Center regions, documenting the lack of impact of coal mining in those areas (figures not

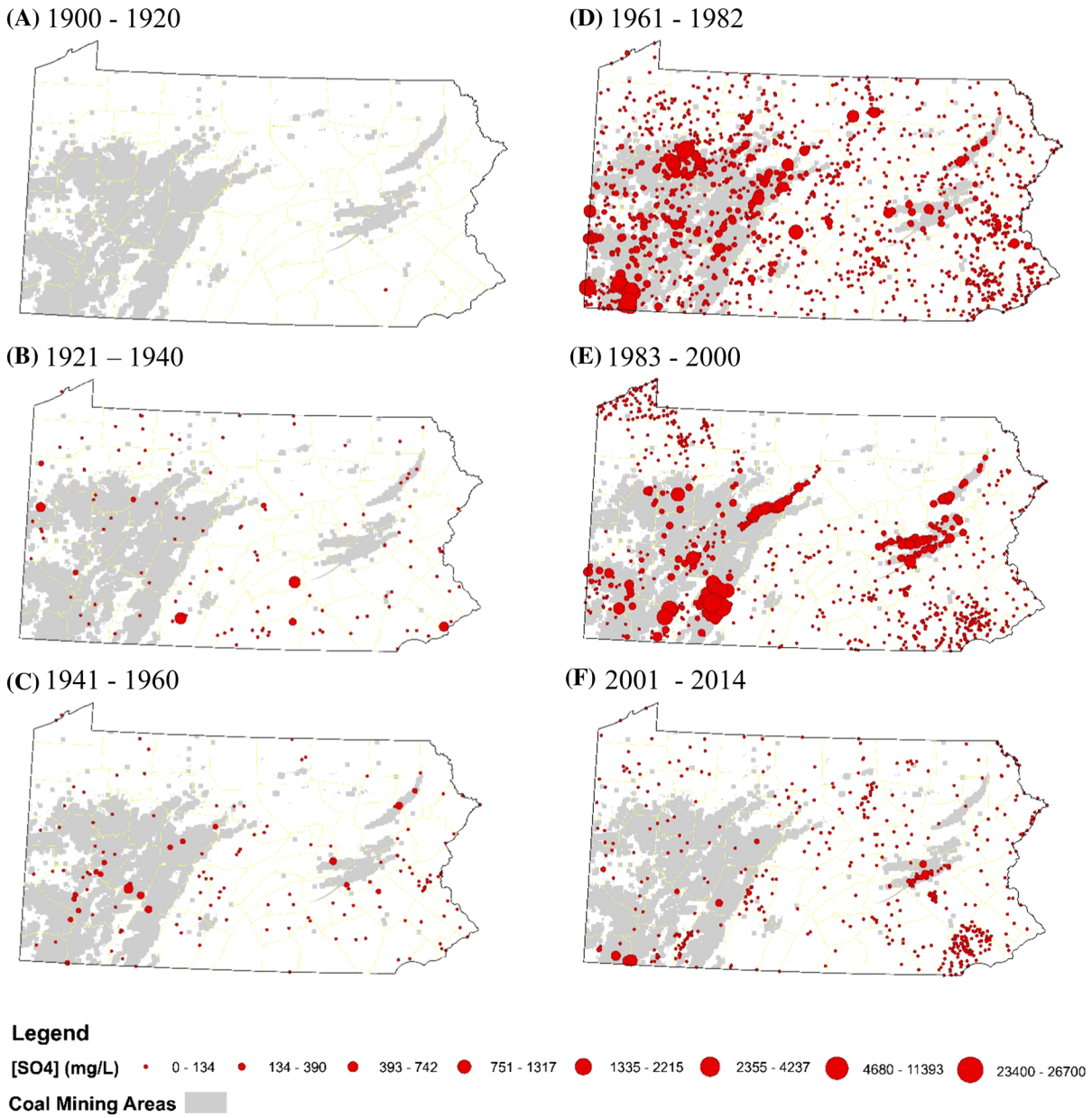


Fig. 2 Maps of all reported measurements of sulfate concentrations in filtered surface water samples in PA (mg/L) for time periods as indicated. Coal mining areas (gray color on the map, including all active and abandoned sites) are shown as observed today

shown). However, a slight decrease in sulfate concentrations in both East and Center regions was observed and is likely attributable to the impacts of acid rain, which will be discussed in following section.

Acid rain versus sulfate concentration

To illustrate the impacts of acid rain on stream sulfate concentrations, we selected an observation site (site code: USGS01545600) near Young Womans Creek (YWC) in Clinton County, where no coal mining was reported upstream of the observation site (see supplementary Fig. A.6 in Online Resources 1). Less than

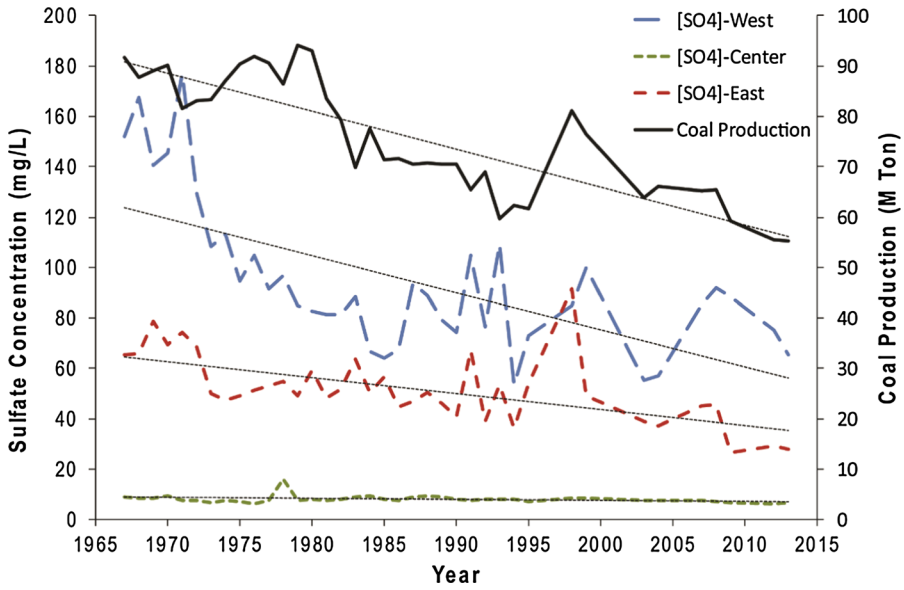


Fig. 3 Trends of regional means of sulfate concentrations and coal production (1965–2014) for the subset of streams with continuous data as discussed in text (dotted lines are linear trendlines for respective variables)

3 km from the YWC sulfate observation site, acid rain is monitored at site PA18.

[SO4] in YWC (i.e., Center) are consistently low compared to the other streams in the West and East regions over the last four decades (Fig. 3, with an overall mean of 7.9 mg/L, StDev of 1.5, and range between 6.3 and 16.1 mg/L). Nonetheless, a steady decrease in [SO_4] was observed in both creek and acid rain since the late 1990s, i.e., after implementation of the Clean Air Act amendment in 1990 (see supplementary Fig. A.7 in Online Resources 1). In fact, a significant positive linear relationship ($p = 0.0002$) was found between [SO_4] in the creek and that in acid rain (Fig. 4).

Barium data availability

The filtered surface water barium concentration data in PA are available from 1963 to 2014 (see supplemental Fig. A.8 in Online Resources 1). In contrast to sulfate, fewer barium measurements are publically available, and they were also measured on a smaller number of rivers. Specifically, 3243 barium data values are available overall that were measured at 471 observation sites from 283 streams. The first barium data became available in 1963 and a few more became available during the late 1960s. Measurements of barium concentrations started to increase from the late 1970s, and many more data were available for the time period of 1997 to 2006.

Temporal and spatial patterns of barium concentrations

Statistics for all the barium concentrations in PA rivers for different time periods are summarized in Fig. 5. Results show that barium concentrations were the highest in the 1960s (with a median value of 60 µg/L). A trend indicating decline in [Ba] started from the early 1970s through the early 2000s (median values decreased from 40 to 23 µg/L, $p < 0.05$). However, [Ba] values slightly increased during 2007 and 2014 (median value increased from 23 to 28 µg/L, $p < 0.05$). The slight increase in [Ba] in recent years might be an indicator of possible impacts from development of gas wells since the first high-volume hydraulically fractured well was emplaced in PA in 2004. It also might be related to decreases in sulfate concentrations because the mineral barite has very low solubility (see discussions in the following section).

Maps of mean barium concentrations for each time period are shown in Fig. 6. The (green) dots represent the locations of [Ba] observations, and the size of the dot represents the level of the [Ba] (larger symbols indicates higher concentrations). Both locations of so-called conventional gas wells (purple) and unconventional wells (orange) are also shown on the map. "Conventional wells" are generally vertical boreholes that were not completed with high-volume hydraulic fracturing. Unconventional wells are generally vertical boreholes with horizontal legs at depth that were

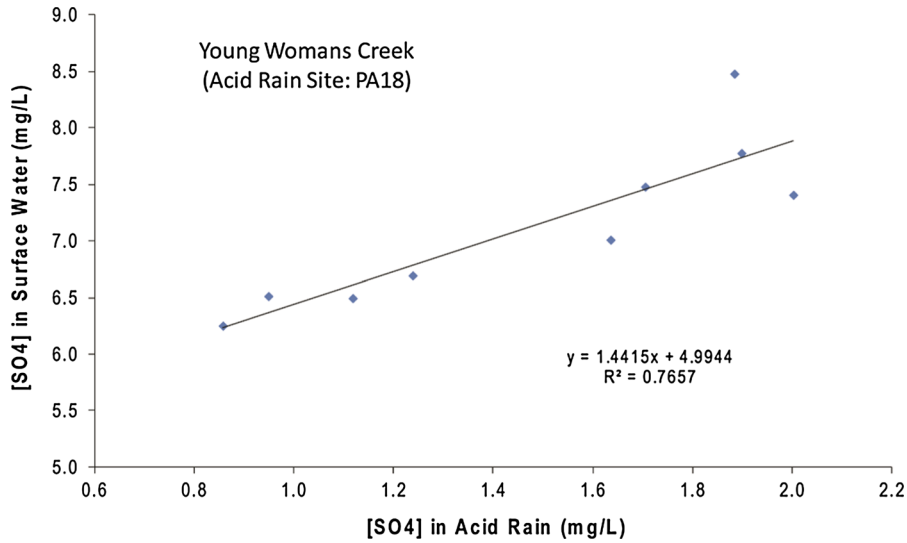


Fig. 4 A plot of sulfate concentrations in Young Womans Creek versus sulfate in rain observed from a nearby location (site PA18) as described in text ($p_{slope} = 0.002$)

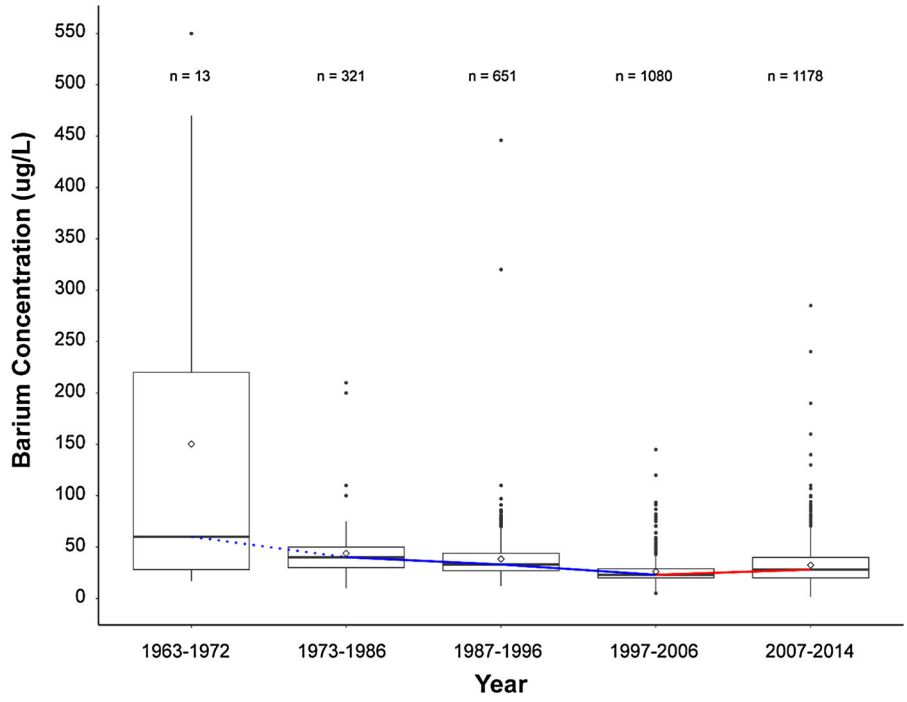


Fig. 5 Box plots summarizing all reported barium concentrations from Pennsylvania rivers for different time periods (All symbols and lines are the same as in Fig. 1)

completed with high-volume hydraulic fracturing (Vidic et al. 2013).

During 1963–1972, only 13 data values were reported from 12 observation sites (Fig. 6a), with most of these data coming from the western region of PA in the area with intensive conventional gas well drilling and coal mining. During the next time period, 1973–1986, a majority of observations were made in areas without oil/gas wells and without significant coal mining (southeastern PA, Fig. 6b). From 1987 to 1996 (Fig. 6c), three data clusters were found in northwestern (Erie county), central (Indiana county), and southeastern PA (Chester county). During the next time period, 1997–2006, the majority of [Ba] observations were made in two counties in eastern PA without oil/gas wells, Schuylkill and Northumberland counties (Fig. 6d). For the time period 2007–2014 (Fig. 6e), the data are clustered in four locations

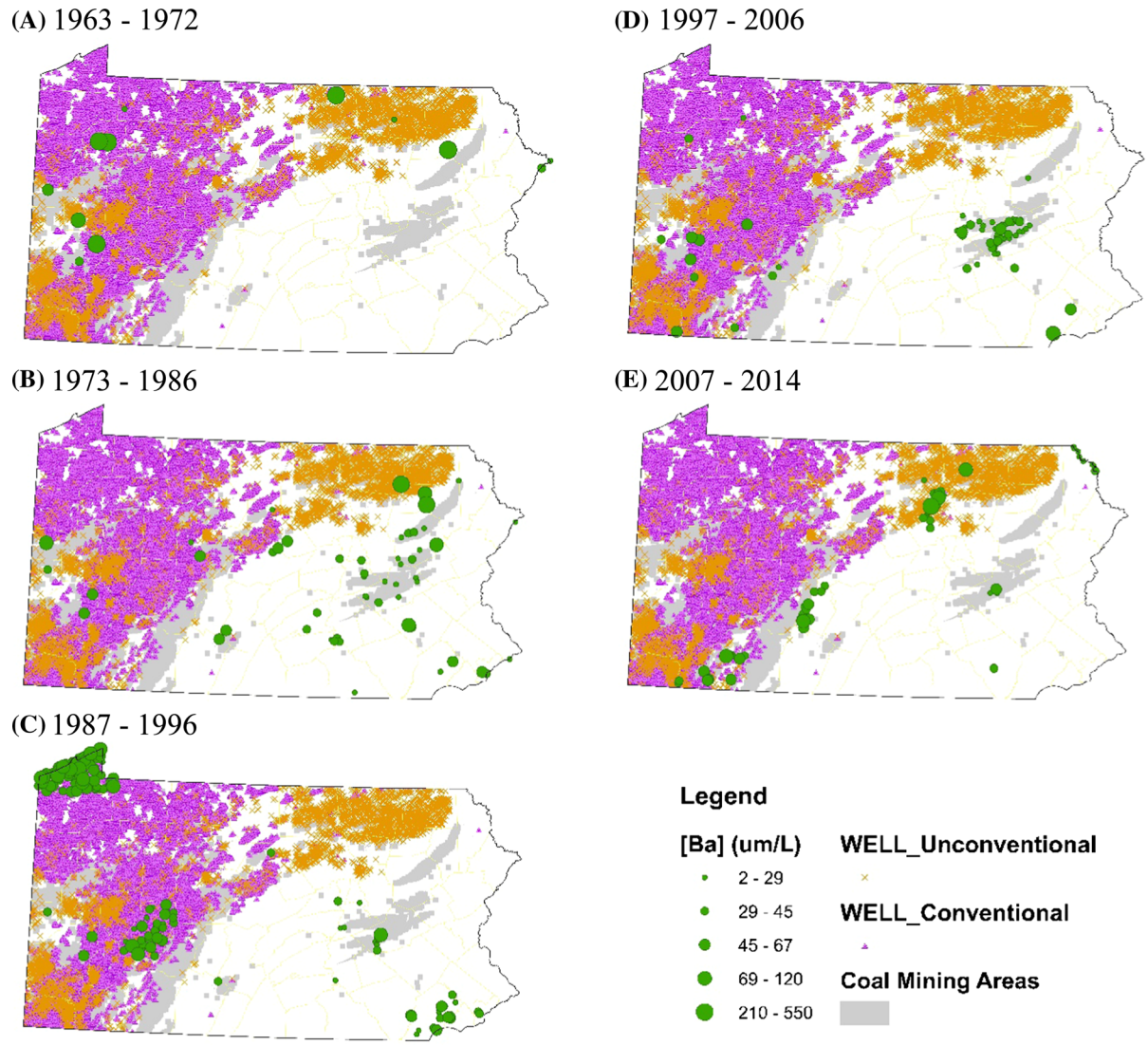


Fig. 6 Maps of barium concentrations (μg/L) in filtered surface waters reported for time periods as indicated. Coal mining regions (gray color on the map) are shown as observed today. Conventional gas wells are shown as background in purple and unconventional wells are in yellow

centered at Wayne, Lycoming, Blair, and Fayette counties, respectively. In general, the relatively small numbers of measurements, the large variations in repeat observation times, and the highly spatially clustered data make it very difficult to interpret the overall temporal and spatial trends of [Ba] concentrations in PA.

Barium concentrations versus coal mining and gas extraction (well)

Statistics for barium concentrations in PA rivers in areas with/without coal mining (COAL/NON-COAL) and areas with/without gas well drilling (WELL/NON-WELL) is summarized in Table 1. Overall, rivers in areas with wells (WELL areas) have the highest barium concentrations (38–40 μg/L), followed by the rivers in NON-WELL and NON-COAL areas (34.4 ± 19.0 μg/L, the error term is one standard

Table 1 Summary of statistics and *t* test *p* values of barium concentrations in PA surface water for areas with/without gas wells (WELL/NON-WELL) and areas with/without coal mining (COAL/NON-COAL)

			Treatment	NON-COAL	COAL	NON-COAL	COAL
				NON-WELL	NON-WELL	WELL	WELL
Treatment				A	B	C	D
[Ba]		Mean (μg/L)		34.4	26.4	38.3	40.0
		StDev (μg/L)		19.0	15.0	38.9	17.0
		Observations (n)		742	1306	450	450
p value	NON-COAL	NON-WELL	A				
	COAL	NON-WELL	B	0.000***			
	NON-COAL	WELL	C	0.029*	0.000***		
	COAL	WELL	D	0.000***	0.000***	0.337	

Two sample two-sided *t* test for difference in means were conducted between treatments. *p* values were calculated for differences between different combinations of WELL and COAL treatments. For example, the first column of *p* values are for comparison of treatment B versus A, C versus A, and D versus A, respectively

* Significantly different at level $\alpha = 0.05$. *** Significant at level $\alpha = 0.001$

deviation). In WELL areas, no significant difference was found in [Ba] in rivers between COAL and NON-COAL areas ($p = 0.337$). Rivers in areas with COAL but NON-WELL have the lowest barium concentrations (26.3 ± 15.0 μg/L, $p < 0.001$). The low [Ba] in the COAL only areas is attributed to the fact that coal mining discharges increase the sulfate concentrations in rivers, which in turn leads to a decrease in barium concentrations due to the low solubility of barite.

Direct comparisons between barium concentrations in areas with and without gas well drilling (WELL vs. NON-WELL) for different time periods are plotted in Fig. 7. Overall (1963–2014), the results show that [Ba] values in WELL areas (37 ± 30 μg/L) are significantly higher ($p < 0.05$) than in NON-WELL areas (29 ± 17 μg/L). This statement is also true for each of the time periods. This might lead to the conclusion that Ba leakage or discharge into PA streams or the use of Ba-rich waste waters as road salt or other such amendments has created the statistically different concentrations. As of 2013, more than a billion gallons of salty Ba-containing wastewater had returned to the land surface in PA due to gas extraction at Marcellus wells (Rahm et al. 2013). The largest anthropogenic use of barium is in drilling muds, and its annual production in the USA tracks oil and gas production (Hanor 2000). Recent reports showed that large volumes of brines produced from well drilling waste have been used for road spreading for dust control and for deicing roads during the winter in PA (Maloney and Yoxtheimer 2012; Skalak et al. 2014). However, limited study suggested that brine road spreading did not lead to significant [Ba] increases in sentiments (Skalak et al. 2014). On the other hand, leakage of Ba could be natural—leakage of natural Ba-containing brines (Adams 2011; Brantley et al. 2014; PADEP 2015; Poth 1962; Skalak et al. 2014).

Given the fact that our current dataset of available measurements of [Ba] is highly clustered in a few locations and the timing of gas/oil well drilling activities are not identified and included in this study, it is impossible to attribute the finding of higher [Ba] to well leakage or development practices. In addition, coal mining may also explain some of the observations. In this regard, [Ba] was also higher in both the northwest and northeast where conventional well drilling activities were intensive but little coal was mined. To clarify the source of the Ba, other analytes such as sodium, calcium, magnesium, and chloride could also be investigated since barium is not the only salt in the oil and gas waste waters (Brantley et al. 2014). Strontium and bromide are also good indicators of Marcellus brines (Vidic et al. 2013). Investigations of these other constituents in the future might help to distinguish if the source of the Ba is drilling muds or natural sources.

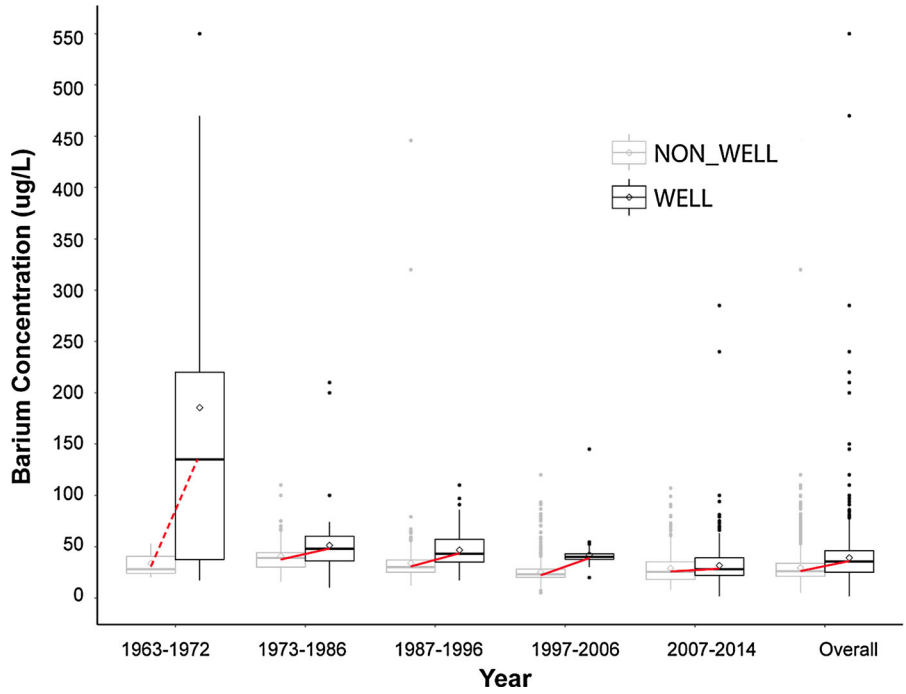


Fig. 7 Comparison of barium concentrations in surface waters between regions in PA with and without natural gas wells (WELL vs. NON_WELL) (all symbols and lines are the same as described for Fig. 1)

Sulfate and barium interaction

We also inspected the behavior of both [Ba] and [SO_4] together for the subset of "continuous" rivers. Barite has very low solubility: the solubility product constant, K_{sp}, equals 1.07×10^{-10} at 25 °C under ambient conditions (Johnson et al. 1992). Because we did not harvest all concentrations from the master datasets, we could not calculate thermodynamic activities for barium and sulfate using activity coefficients. However, in these dilute waters, activity coefficients will be near unity and thus will represent a small correction, especially in comparison with the correction to the solubility products (discussed below) for temperature. To nonetheless explore for indications of barite solubility control, we converted barium and sulfate concentrations to molar concentrations. The product of barium and sulfate molar concentrations (i.e., labeled here as Ba*SO_4) was then calculated using [Ba] and [SO_4] data measured from the same sites at the same time. Regional annual means of the [Ba]*[SO_4] are shown in Fig. 8. A series of K_{sp} values at a range of temperature from 25 to 5 °C are plotted as a gray gradient and as labeled lines for reference. The subset of eastern (East) and western (West) rivers in the figure shows a decrease in the measured concentration products with time, while the central river (Central, i.e., Young Womans creek) shows no temporal change.

In the eastern rivers where the mean annual temperature was reported to be around 15 °C, the mean annual concentration products [$(1.48 \pm 0.46) \times 10^{-10}$] approach the barite *K*sp at 25 °C (1.07×10^{-10}) since approximately 1986 and may have stabilized at this value. Again, however, without activity corrections, we cannot make a conclusion as to whether barite is expected to be precipitating. Nonetheless, it appears that these rivers have been supersaturated with respect to barite in the past and could be equilibrating with respect to barite now. In contrast, the concentration products for the western rivers, where the mean annual temperature is slightly lower (approximately 13.3–14.6 °C), are a factor of 2 times higher than the *K*sp with respect to 25 °C ([Ba][SO4] = $(3.03 \pm 0.96) \times 10^{-10}$). The concentration products for the central river [Young Womans Creek (4.38×10^{-11})], where the mean temperature in the dataset is reported as 8.9 °C, shows no temporal change and is only higher than the *K*sp at 5 °C. These data are suggestive that the western and eastern rivers have been oversaturated with respect to barite, while the central river has generally been

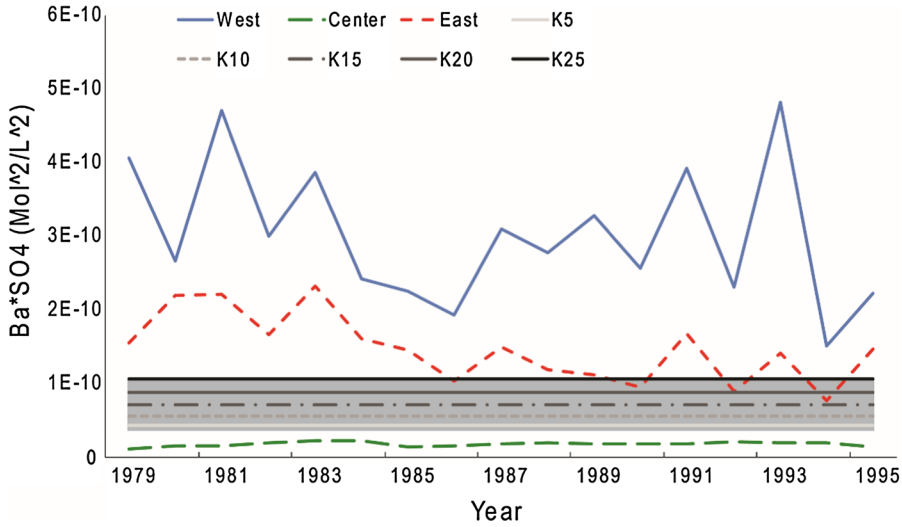


Fig. 8 Products of barium and sulfate concentrations plotted versus time (1976–1998). The thermodynamic solubility products for barite, K_{sp} from 5 to 25 °C (gray area, also lines labeled K_T, where T is the temperature) are plotted for comparison. Although the calculations of ion activities for these river data were beyond the scope of this paper, the plotted comparison shows that some rivers are likely oversaturated with respect to chemical equilibrium for barite in the west and east, while Young Woman's creek in central PA is likely undersaturated. See text for a description of the rivers compiled in "West," "Center," and "East" datasets

undersaturated. The extremely high concentration products in the western rivers are attributed to the large influence of coal mining on both Ba and SO_4 concentrations. The possible effect of barite solubility control has been noted in a previous finding (Cravotta III 2008). The observation that eastern and western rivers could be as much as 2X oversaturated with respect to barite is also similar to other observations in the literature where some natural waters influenced by seawater are up to 2X supersaturated with respect to barite (Hanor and Chan 1977).

Supersaturation can result when nucleation of crystals is inhibited. Another explanation for apparent supersaturation might be that small barite nuclei precipitate but then pass through the filter papers and are thus interpreted as solutes instead of particles. To test the latter interpretation, we attempted to inspect the unfiltered versus filtered data. Unfortunately, it was impossible to do so because all the data for unfiltered waters were reported by the USGS as 100 ppb, and we inferred that these values actually should have been noted as < 100 ppb (i.e., censored data). Nevertheless, an important implication of this analysis of stream chemistry is that streams in PA could be increasing in [Ba] not only because of human activities directly but also because of decreasing $[SO_4]$. If this is true in some rivers, [Ba] could be increasing in PA streams because the $[SO_4]$ in acid rain is decreasing.

Determining the sulfate and barium baseline concentration values

Determining pristine rivers

As described above, pristine rivers in this study were determined as any rivers that have StDev values less than the changepoint value derived from changepoint analysis. The results of changepoint analysis are shown in Fig. 9a for sulfate and in Fig. 9b for barium, respectively. The changepoint for sulfate is at a site-index of 80 and the corresponding critical StDev value is 11.06 mg/L. For barium, the changepoint is at a site-index of 44 and the corresponding critical StDev value is 6.5 µg/L. With this analysis, a total of 87 streams (or 121 observation sites) that met our pristine river criteria were found for sulfate and 37 streams (or 41 observation sites) for barium, respectively. The locations of pristine rivers can be found in the supplementary Fig. A.9 in Online Resources 1.

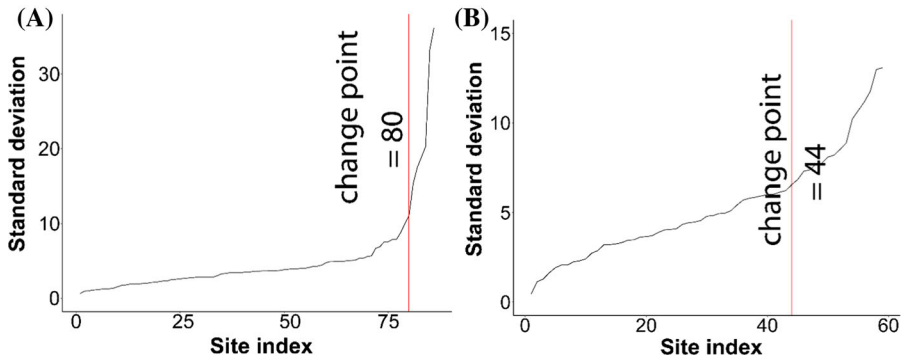


Fig. 9 Application of the likelihood ratio test for sulfate (**a**) and barium (**b**). The vertical line identifies the changepoint in each set of data

Major bedrock types in PA

Since bedrock chemistry is known as one of the most important factors affecting river concentrations (Olson and Hawkins 2012), we attempted to determine baseline values for each major bedrock types in PA as the first step in this analysis. In other words, we assumed that bedrock is a dominant factor controlling background concentrations, ignoring land use, weather variations, and other factors. Bedrock types of PA were derived from the Pennsylvania geologic map data (https://mrdata.usgs.gov/geology/state/state.php?state=PA). Based on the characteristics of each bedrock type, we regrouped the categories into six major bedrock types (Table 2). These six rock types will be used to determine the baseline of the barium and sulfate concentrations in this study. A map of major bedrock types in PA is included as supplementary Fig. A.10 in the Online Resources 1.

Determining baseline values

Baseline values are defined in this study as the mean stream concentrations measured before any "major human impacts." The baseline values and statistics for sulfate and barium are calculated from concentrations measured from the respective "pristine rivers," which were determined based on changepoint analysis described in previous sections, and summarized in Table 3. Overall the baseline sulfate concentration in PA is 15.8 ± 9.6 mg/L, which is comparable (within 95% of confidence interval) with a value observed in 1904 from a presumably "pre-contamination" river in Lancaster County in PA (12 mg/L, from the USGS dataset) and the estimated worldwide average value of 11 mg/L (Turekian 1977). Statewide, the baseline barium concentration is 27.7 ± 10.6 µg/L, which is comparable with a value (27 ± 32 µg/L, mean concentrations in PA surface water before 2003) determined in a previous study by Vidic et al. (2013) based on all USGS data for counties in PA where shale gas development was occurring. The estimate is also

Table 2 List of bins used for major bedrock types in PA

Major bedrock type	Bedrock sub-types[a]
Conglomerate	Conglomerate, gravel
Igneous rocks	Andesite, anorthosite, diabase, pegmatite
Limestone	Dolostone/dolomite, limestone
Metamorphic rocks	Felsic gneiss, gneiss, granitic gneiss, mafic gneiss, mafic metavolcanic rock, marble, meta-basalt, meta-rhyolite, mica schist, phyllite, schist, serpentinite, slate
Siltstone	Black shale, clay or mud, mudstone, shale, siltstone
Sandstone	Argillite, arkose, graywacke, quartzite, sand, sandstone

[a]Bedrock (sub-) types are derived from the Pennsylvania geologic map data

Table 3 Summary of baseline concentrations and statistical parameters for major bedrock types

Rock type	Igneous rocks	Limestone	Metamorphic rocks	Siltstone	Sandstone	Overall (statewide)
SO_4 (mg/L)						
Mean	17.7	26.7	20.9	12.6	12.4	15.8
StDev	11.2	10.2	8.6	6.8	8.2	9.6
Min	6.4	3.2	0.6	0.3	1.3	0.3
Max	51.7	114.0	84.0	98.3	140.0	140.0
Median	14.0	25.2	19.2	11.0	9.8	12.9
DataCount	120	938	1475	2769	2612	7914
Sites/streams	5/3	13/10	41/25	33/24	29/25	121/87
Ba (μg/L)						
Mean		33.8	38.7	29.5	25.8	27.7
StDev		12.7	12.6	10.5	9.7	10.6
Min		24.0	21.0	8.0	8.0	8.0
Max		54.0	64.0	66.0	62.0	66.0
Median		27.0	39.0	26.0	23.0	25.0
DataCount		15	40	282	523	860
Sites/streams	0/0	3/2	8/8	11/10	19/17	41/37

comparable to the average value of 20 μg/L found by Turekian (1977) for rivers worldwide.

Based on the data shown in Table 3, the mean sulfate concentrations in waters interacting with crystalline rocks (igneous and metamorphic) are on average 5–10 mg/L higher than waters interacting with sedimentary rocks in PA ($p < 0.05$). However, the maximum [SO_4] values were observed on sedimentary rocks (Table 3). Likewise, the mean [Ba] values were somewhat higher on crystalline as compared to sedimentary rocks.

Comparison with a USGS HBN site in PA (i.e., Young Womans Creek)

Comparison of [SO_4] and [Ba] baseline values with the respective concentrations in Young Womans Creek (YWC) is presented in Fig. 10. YWC is the only stream in PA that is part of the USGS Hydrological Benchmark Network (HBN). Rivers in the HBN often serve as references for modeling background values and for studies of impacts of human activities (Olson and Hawkins 2012; Smith et al. 2003). YWC, located in northern central PA, is underlain mostly by sandstone. The mean baseline [SO_4] for sandstone in PA (12.4 ± 8.2 mg/L) is slightly higher ($p < 0.05$) than that in YWC (7.7 ± 4.2 mg/L). However, the overall statewide baseline value for [SO_4] (15.8 ± 9.6 mg/L) is almost double the mean concentration of YWC because of the higher concentrations from other bedrock types (Table 3). This result suggests that using YWC alone to represent all rivers in PA will likely overestimate the impacts of human activities on [SO_4] concentrations. For example, the prevalence of crystalline rocks in southeastern PA would be expected to host rivers with higher sulfate concentrations.

On the other hand, [Ba] concentrations in YWC (29.4 ± 4.4 μg/L) are only slightly higher ($p < 0.05$) than the statewide average for sandstone (25.8 ± 9.7 μg/L) and the statewide overall values (27.7 ± 10.6 μg/L).

Spatial patterns of sulfate and barium baseline

Spatial patterns of the sulfate and barium baseline concentrations in rivers are shown in Fig. 11. As expected, geological features are a first-order control on river location and ion concentrations. Specifically, the map of [SO_4] documents high values largely in the southeastern part of the state, where crystalline rocks dominate, in the southwest where acid mine drainage is common, and in some sections of the central valley and ridge region of PA. It is likely that these latter

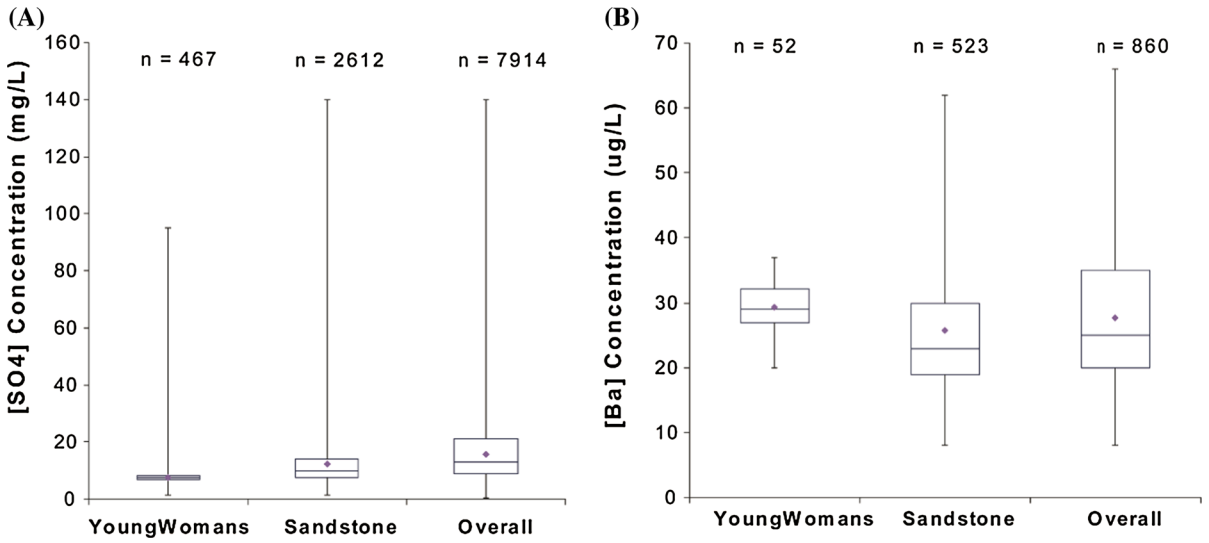


Fig. 10 Comparison of mean (and median) riverine values of [SO_4] (**a**) and [Ba] (**b**) between Young Womans Creek and PA rivers on sandstones/PA rivers on all rock types (labeled, Overall). n indicates the total data number for each category. The differences in medians are all statistically significant as determined using Wilcoxon Signed Rank test at level $\alpha = 0.05$

areas may be influenced by black shale or sandstone outcrops that expose sedimentary or mineralized pyrite to oxidative weathering, causing relatively high [SO_4].

Impacts of human activities

We next assessed changes in chemical concentrations from the baseline to detect evidence for the impact of human activities such as coal mining and shale gas development. Using publically available historical data, we mapped the changes in sulfate (Fig. 12a) and barium (Fig. 12b) concentrations from their respective baseline values. The results show that the biggest changes in sulfate concentrations are observed mainly in the coal mining areas. This observation is consistent with previous studies that demonstrated the impacts of coal mining on sulfate concentrations in surface waters in parts of PA (Raymond and Oh 2009). For barium, the largest changes are observed in spots in the northeastern region of PA. This is a part of the state that has a very high density of unconventional wells but also is the location of natural brine seeps (Brantley et al. 2014). However, as discussed above, the decreasing [SO_4] concentrations may be another factor affecting [Ba] in these sites. Furthermore, given that our analysis included all data and the sampling of individual rivers has changed over time, it is possible that the changing set of rivers explains why [Ba] has increased with time. Therefore, our analysis documents an increase in [Ba] in northeastern PA but cannot distinguish the cause. Local studies are needed to track down the causes of increased [Ba] in some rivers.

Summary and conclusions

In the master water quality database compiled from publically available data sources, sulfate data for PA streams are available from 1904 to 2014. Data availability varies greatly over time as monitoring programs change. The median surface water values of [SO_4] increased as coal production increased to become the highest during the 1940s and 50 s. These concentrations then dramatically decreased as coal production decreased. Decreases were also dramatic after the Clean Air Act in 1970, and amendments were implemented in the 1990s because of amelioration of acid rain. Our assessed statewide [SO_4] baseline, 15.8 ± 9.6 mg/L, represents the best estimate of sulfate concentrations in streams in PA before human impacts. This value varies among different bedrock types (ranging from 12.4 to 26.7 mg/L). Across the state, the changes in [SO_4] were evaluated for given locations as a function of time. Larger changes in

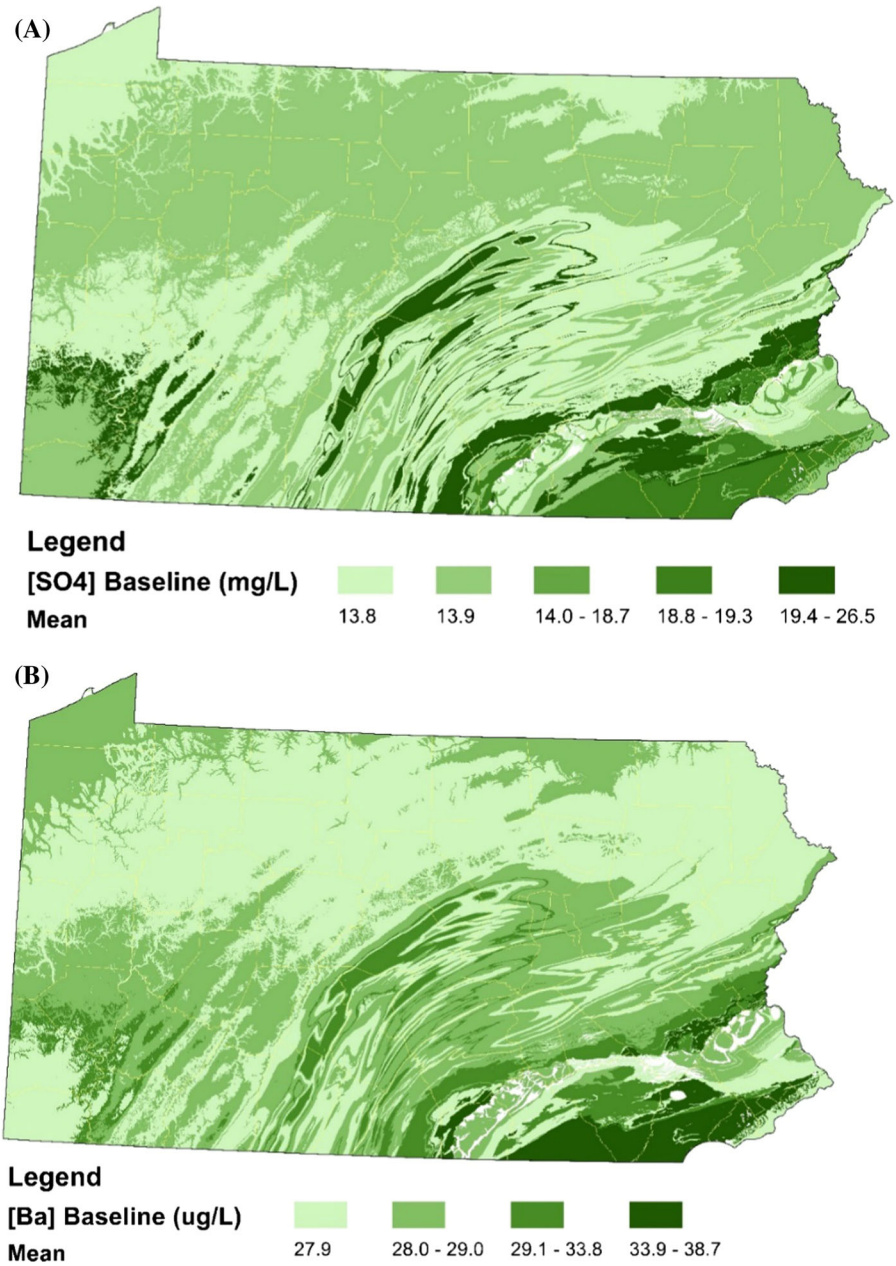


Fig. 11 Map of sulfate (**a**) and barium (**b**) baseline concentrations in PA surface waters as determined by major rock types. The color scheme of these maps is generated based on a natural breaks classification of each dataset, which divides the dataset into classes whose boundaries are set where there are relatively big differences in the data values

[SO_4] from the baseline value were observed in coal mining areas as compared to other areas. This clearly documents the impacts of coal mining and burning on [SO_4] concentrations in PA streams.

Barium concentrations in PA rivers are publically available from 1963 to 2014 but are much more limited in number than [SO_4] values. Data availability varies over time, and the data are highly spatially clustered in a few locations. The [Ba] values were the highest during the 1960s and gradually decreased until the early 2000s. The statewide baseline [Ba] is 27.2 ± 10.6 μg/L and ranges from 25.8 to 38.7 μg/L for different bedrock types. No clear spatial patterns were found for changes of [Ba] with respect to the baseline value. However, we documented higher [Ba] in areas with higher density of gas wells when compared to areas with lower densities. We also demonstrated a slight increase in [Ba] in recent years

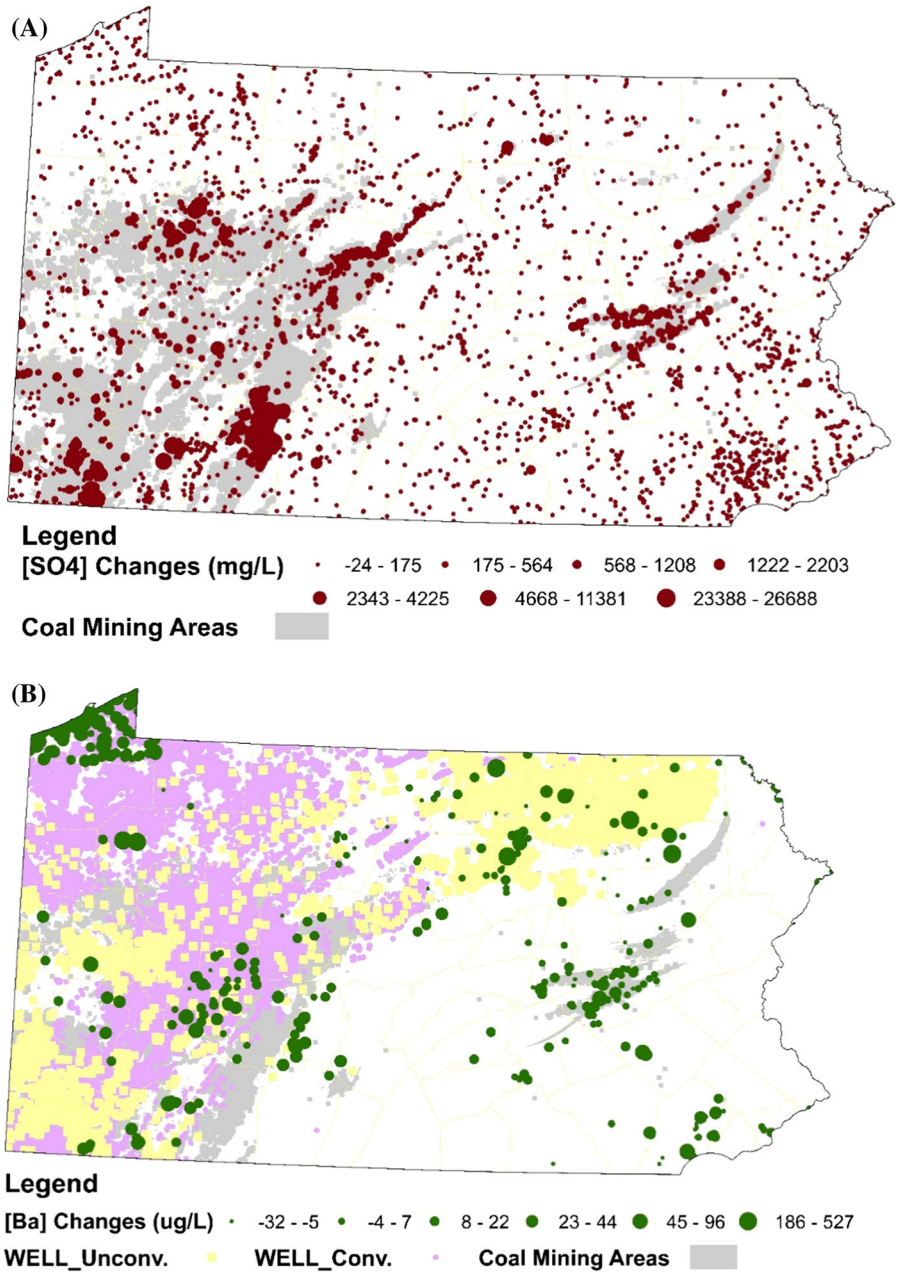


Fig. 12 Map of differences between historically observed concentrations and the presumed baseline values of sulfate (**a**) and barium (**b**) concentrations in Pennsylvania rivers

in northeastern PA where many shale gas wells were drilled. Our [Ba] observations could indicate leakage of brines during disposal of wastewater or leakage of barium-containing muds during drilling. An equally viable interpretation is that the same counties experiencing gas development also host natural brine seeps that bring Ba into streams. However, the most compelling explanation is that [Ba] is increasing because barite is staying equilibrated with the water, while [SO_4] is decreasing from the amelioration of acid rain and acid mine drainage. Further analysis of the Ba behavior is warranted.

This work demonstrates a new method to determine sulfate and barium baseline values using publically available data. This method can be applied to other analytes. Since the major bedrock type is the only factor considered in this study, more factors, such as multiple land use activities and varying climate

conditions, could be included in future studies to explore more detailed spatial variation of baseline values. Furthermore, data mining and data assimilation techniques could be applied to fill temporal and spatial data gaps and simulate the trends of chemical concentrations in PA surface water and hence to quantify the impacts of shale gas extractions. Overall, our approach shows that the use of coal has impacted Pennsylvania streams to a much greater extent than use of shale gas, as documented in barium and sulfate stream chemistry throughout the state.

Acknowledgements This work was funded from a gift to Penn State for the Pennsylvania State University General Electric Fund for the Center for Collaborative Research on Intelligent Natural Gas Supply Systems. The Shale Network database has been funded by the National Science Foundation RCN-SEES funding (OCE-11-40159), by Penn State, and by the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI). This work is also partially funded by National Science Foundation Grants DMS-1505256 and IIS-1639150.

References

Adams, M. B. (2011). Land application of hydrofracturing fluids damages a deciduous forest stand in West Virginia. *Journal of Environmental Quality, 40*(4), 1340–1344. doi:10.2134/jeq2010.0504.

Alexander, R. B., Slack, J. R., Ludtke, A. S., Fitzgerald, K. K., & Schertz, T. L. (1998). Data from selected US. Geological Survey National stream water quality monitoring networks. *Water Resources Research, 34*(9), 2401–2405. doi:10.1029/98wr01530.

Brantley, S. L., Yoxtheimer, D., Arjmand, S., Grieve, P., Vidic, R. D., Pollak, J., et al. (2014). Water resource impacts during unconventional shale gas development: The Pennsylvania experience. *International Journal of Coal Geology, 126,* 140–156. doi:110.1016/j.coal.2013.1012.1017.

Cravotta Iii, C. A. (2008). Dissolved metals and associated constituents in abandoned coal-mine discharges, Pennsylvania, USA. Part 2: Geochemical controls on constituent concentrations. *Applied Geochemistry, 23*(2), 203–226. doi:10.1016/j.apgeochem.2007.10.003.

Faoro, R. B., & McMullen, T. B. (1977). *National trends in trace metals in ambient air, 1965–1974* (pp. 28). Environmental Protection Agency, Research Triangle Park, NC (USA). Monitoring and Data Analysis Div.

Hanor, J. S. (2000). Barite–celestine geochemistry and environments of formation. *Reviews in Mineralogy and Geochemistry, 40*(1), 193–275. doi:10.2138/rmg.2000.40.4.

Hanor, J. S., & Chan, L.-H. (1977). Non-conservative behavior of barium during mixing of Mississippi River and Gulf of Mexico waters. *Earth and Planetary Science Letters, 37*(2), 242–250. doi:10.1016/0012-821X(77)90169-8.

Hladik, M. L., Focazio, M. J., & Engle, M. (2014). Discharges of produced waters from oil and gas extraction via wastewater treatment plants are sources of disinfection by-products to receiving streams. *Science of the Total Environment, 466–467,* 1085–1093. doi:10.1016/j.scitotenv.2013.08.008.

Johnson, J. W., Oelkers, E. H., & Helgeson, H. C. (1992). Supcrt92. *Computational Geosciences, 18,* 899–947.

Kim, K.-H., Yun, S.-T., Kim, H.-K., & Kim, J.-W. (2015). Determination of natural backgrounds and thresholds of nitrate in South Korean groundwater using model-based statistical approaches. *Journal of Geochemical Exploration, 148,* 196–205. doi:10.1016/j.gexplo.2014.10.001.

Kirby, C. S., McInerney, B., & Turner, M. D. (2008). Groundtruthing and potential for predicting acid deposition impacts in headwater streams using bedrock geology, GIS, angling, and stream chemistry. *Science of the Total Environment, 393*(2–3), 249–261. doi:10.1016/j.scitotenv.2007.12.026.

Llewellyn, G. T. (2014). Evidence and mechanisms for Appalachian Basin brine migration into shallow aquifers in NE Pennsylvania, USA. *Hydrogeology Journal, 22*(5), 1055–1066. doi:10.1007/s10040-014-1125-1.

Maloney, K. O., & Yoxtheimer, D. A. (2012). Production and disposal of waste materials from gas and oil extraction from the Marcellus Shale Play in Pennsylvania. *Environmental Practice, 14*(4), 278–287. doi:10.1017/s146604661200035x.

Olson, J. R., & Hawkins, C. P. (2012). Predicting natural base-flow stream water chemistry in the western United States. *Water Resources Research, 48*(2), W02504. doi:10.1029/2011wr011088.

PADEP. (2015). *Roadspreading of brine for dust control and road stabilization (March 2015)*. New York: Department of Environmental Protection.

Poth, C. W. (1962). The occurrence of brine in western Pennsylvania. *Pennsylvania Geological Survey Bulletin, M47,* 1–53.

Rahm, B. G., Bates, J. T., Bertoia, L. R., Galford, A. E., Yoxtheimer, D. A., & Riha, S. J. (2013). Wastewater management and Marcellus Shale gas development: Trends, drivers, and planning implications. *Journal of Environmental Management, 120,* 105–113. doi:10.1016/j.jenvman.2013.02.029.

Raymond, P. A., & Oh, N.-H. (2009). Long term changes of chemical weathering products in rivers heavily impacted from acid mine drainage: Insights on the impact of coal mining on regional and global carbon and sulfur budgets. *Earth and Planetary Science Letters, 284*(1–2), 50–56. doi:10.1016/j.epsl.2009.04.006.

Rhodes, A. L., & Horton, N. J. (2015). Establishing baseline water quality for household wells within the Marcellus Shale gas region, Susquehanna County, Pennsylvania, USA. *Applied Geochemistry, 60,* 14–28. doi:10.1016/j.apgeochem.2015.03.004.

Rodhe, H., Dentener, F., & Schulz, M. (2002). The global distribution of acidifying wet deposition. *Environmental Science and Technology, 36*(20), 4382–4388. doi:10.1021/es020057g.

Sams III, J. I., & Beer, K. M. (2000). Effects of coal-mine drainage on stream water quality in the Allegheny and Monongahela River Basins: Sulfate transport and trends

(U. S. D. o. t. Interior, & U. S. G. Survey, Trans.). *Water-Resources Investigations Report 99-4208*.

Skalak, K. J., Engle, M. A., Rowan, E. L., Jolly, G. D., Conko, K. M., Benthem, A. J., et al. (2014). Surface disposal of produced waters in western and southwestern Pennsylvania: Potential for accumulation of alkali-earth elements in sediments. *International Journal of Coal Geology, 126,* 162–170. doi:10.1016/j.coal.2013.12.001.

Smith, R. A., Alexander, R. B., & Schwarz, G. E. (2003). natural background concentrations of nutrients in streams and rivers of the conterminous United States. *Environmental Science and Technology, 37*(14), 3039–3047. doi:10.1021/es020663b.

Turekian, K. (1977). Geochemical distribution of elements. In: *Encyclopedia of science and technology* (4th ed., pp. 627–630). New York: McGraw-Hill.

Vidic, R. D., Brantley, S. L., Vandenbossche, J. M., Yoxtheimer, D., & Abad, J. D. (2013). Impact of shale gas development on regional water quality. *Science, 340*(6134), 1235009. doi:10.1126/science.1235009.

Voutchkova, D. D., Kristiansen, S. M., Hansen, B., Ernstsen, V., Sørensen, B. L., & Esbensen, K. H. (2014). Iodine concentrations in Danish groundwater: Historical data assessment 1933–2011. *Environmental Geochemistry and Health, 36*(6), 1151–1164. doi:10.1007/s10653-014-9625-4.