Water Data to Answer Urgent Water Policy Questions:

Monitoring design, available data, and filling data gaps for determining whether shale gas development activities contaminate surface water or groundwater in the Susquehanna River Basin

The second in a series of three reports focused on water data needed to address water policy issues. The first report focuses on agricultural management practices in the Lake Erie drainage basin and the next report will provide an overview of existing water-quality data across the Northeast-Midwest region.

A report published by
The Northeast-Midwest Institute in collaboration with the U.S. Geological Survey

NORTHEAST-MIDWEST INSTITUTE

USGS
science for a changing world

For More Information

For more information about the Northeast-Midwest Institute please see www.nemw.org. Additional information about this report and associated companion reports is available at www.nemw.org.

Citation

Betanzo, E.A., Hagen, E.R., Wilson, J.T., Reckhow, K.H., Hayes, L., Argue, D.M., and Cangelosi, A.A., 2016, Water data to answer urgent water policy questions: Monitoring design, available data and filling data gaps for determining whether shale gas development activities contaminate surface water or groundwater in the Susquehanna River Basin, Northeast-Midwest Institute Report, 238 p., http://www.nemw.org
DOI: 10.13140/RG.2.1.4051.6885

Funding

This work was made possible by a grant from the U. S. Geological Survey.

ISBN: 978-0-9864448-2-1

Cover photos: USGS (hydraulic fracturing drill site and stream monitoring photos); Ostroff Law via Wikipedia Commons (hydraulic fracturing site in Warren Center); Daniel Soeder, U.S. Department of Energy (Marcellus outcrop background photo)

Project Team

Elin A. Betanzo, Northeast-Midwest Institute, Project Manager

Erik R. Hagen, Northeast-Midwest Institute, Lead Researcher

John T. Wilson, U.S. Geological Survey, Hydrologic Data Analyst

Kenneth H. Reckhow, Professor Emeritus at Duke University, Statistical Analyst

Laura Hayes, U.S. Geological Survey, GIS Specialist

Denise M. Argue, U.S. Geological Survey, Hydrologic Data Analyst

Allegra A. Cangelosi, Northeast-Midwest Institute, Senior Policy Analyst, Environmental Projects

Technical Advisory Committee

Kate Dunlap, Trout Unlimited

Andrew Gavin, Susquehanna River Basin Commission

Adam Goehner, Pembina Institute

Peter Murdoch, U.S. Geological Survey

James Saiers, Yale School of Forestry and Environmental Studies

Tony Shaw (retired), Pennsylvania Department of Environmental Protection

Bert Smith, formerly with Chesapeake Energy, now with EnviroClean Products and Services

Daniel Soeder, Department of Energy National Energy Technology Laboratory

John Wilson, U.S. Geological Survey

David Yoxtheimer, Penn State University

Technical Reviewers

Jeffrey Deacon, U.S. Geological Survey

William Kappel, U.S. Geological Survey

Dennis Risser, U.S. Geological Survey

Gary Rowe, U.S. Geological Survey

Blue Ribbon Project Steering Committee

Glenn Benoy, International Joint Commission

Paul A. Biedrzycki, City of Milwaukee Health Department

Suzanne Bricker, National Oceanic and Atmospheric Administration

Allegra Cangelosi, Northeast-Midwest Institute

Julius Ciaccia, Northeast Ohio Regional Sewer District

Thomas Davenport, Environmental Protection Agency, Region 5

Joseph Depinto, Limnotech

Blayne Diacont, Range Resources

Timothy Eder, Great Lakes Commission

Norman Grannemann, U.S. Geological Survey

Carlton Haywood, Interstate Commission on the Potomac River Basin

Jonathan Higgins, The Nature Conservancy

Jennifer Hoffman, Chesapeake Energy

Susy King, New England Interstate Water Pollution Control Commission

James Miller, Organic Valley

Jeffrey Myers, New York Department of Environmental Conservation

Laura Rubin, Huron River Watershed Council

Robert Tudor (retired), Delaware River Basin Commission

Mark Walbridge, U.S. Department of Agriculture, Agricultural Research Service

Susan Weaver, Pennsylvania Department of Environmental Protection

The Northeast-Midwest Institute President and CEO Foreword

No resource is more vital to the future of the Northeast-Midwest (NEMW) region—its population, industry, food production, and quality of life--than its naturally abundant fresh water. Yet is the region's capacity to monitor the state of this natural asset and changes to it over time commensurate with its value? Sufficient water monitoring capacity and information is particularly of concern as large-scale development activities such as shale gas extraction and agriculture dramatically expand freshwater use. As a Washington-based, nonprofit, and nonpartisan research organization dedicated to economic vitality, environmental quality, and regional equity for Northeast and Midwest states, the Northeast-Midwest Institute (NEMWI) could hardly ignore this critical question for the region and its environment. Both the economic benefits and potential water-quality impacts from these issues are high priorities for policy makers. To manage these activities responsibly, policy-makers need access to regional water-quality monitoring information that is timely, credible and useful. NEMWI teamed with the U.S. Geological Survey (USGS) to investigate the capacity of the NEMW region's water monitoring programs to support informed policy decisions.

NEMWI and the USGS undertook two case studies to illustrate the types and amounts of data needed to answer critical water policy questions, and determine how much of those water-quality data are currently available. This report summarizes one case study: our investigation of water-quality data relating to shale gas development in the Susquehanna River Basin, and is a companion to a similar case study investigation of water-quality data relating to nutrient enrichment in the Lake Erie drainage basin. The results of these case studies will inform a qualitative review of water data availability across the NEMW region in a subsequent State of the Region Report.

In December of 2014, the State of New York decided that high-volume hydraulic fracturing development should not proceed because the current scientific information is insufficient to determine the level of risk to public health and whether the risks can be adequately managed (New York State Department of Health, 2014). A recent draft report by the U.S. Environmental Protection Agency (USEPA) studying the impacts of hydraulic fracturing on drinking water resources nationwide did not find evidence of widespread systemic impacts on drinking water in the United States; however, the limited amount of data collected before and during hydraulic fracturing activities reduced their ability to determine whether hydraulic fracturing affected water quality in cases of alleged contamination (U.S. Environmental Protection Agency, 2015a). These two reports make clear the urgency and the need for reliable water-quality data to support policy decisions around shale gas development.

In this report, the NEMWI in cooperation with the USGS investigates the minimum quantity of water-quality data that would be necessary to detect a statistically significant change in water quality related to shale gas development in the Susquehanna River Basin, and analyzes whether those data are currently being generated.

Among the significant findings of the case study are the following:

- The existing surface-water quality data in the Susquehanna River Basin are insufficient to detect water-quality change related to shale gas development, but additional monitoring can build on available surface-water data. Key steps to generating the needed data include increased monitoring at a subset of priority monitoring sites that includes increased sampling frequency, sampling for additional priority parameters and streamflow, and commitment to long-term monitoring.

- The publicly available groundwater-quality data in the Susquehanna River Basin are not sufficient to detect whether shale gas development is contaminating groundwater, and the available data are not adequate to serve as the foundation of a new monitoring program.
- The needed groundwater data can be collected through the design and implementation of a systematic, long-term groundwater monitoring program for detecting groundwater quality change related to shale gas development in the Susquehanna River Basin, building on data collected by the shale gas industry, if possible.
- The most effective data collection can be achieved through a coordinating entity with representation from water monitoring organizations, shale gas industry, domestic well owners, and citizens to ensure all the necessary supporting information is incorporated into an effective monitoring design.

Ensuring quality fresh water, now and in the future, is a priority of the highest order for the Northeast-Midwest region. The research presented in this report documents that existing data are not sufficient to detect water-quality change related to shale gas development for surface water or groundwater in the Susquehanna River Basin, perhaps the most data-rich river basin with active shale gas development. Surely collecting and using the needed information to identify effective policy solutions for protecting water resources is a first step to achieving this critical environmental and economic goal for the Northeast-Midwest region in the years ahead.

Michael J. Goff, Ph.D.
President and CEO
Northeast-Midwest Institute

Table of Contents

List of Figures

List of Tables

Acronyms

ALLARM: Alliance of Aquatic Resource Monitoring

ArcGIS: Arc Geographic Information System

ARS: Agricultural Research Service

BTEX: Benzene, Toluene, Ethylbenzene, and Xylene

CUAHSI: Consortium of Universities for the Advancement of Hydrologic Sciences, Inc.

FY: Fiscal Year

GAGES: Geospatial Attributes of Gages for Evaluating Streamflow

GIS: Geographic Information System

HUC12: 12-digit hydrologic unit code

HVHF: High-volume hydraulic fracturing

NAWQA: National Water-Quality Assessment

NEMI: National Environmental Methods Index

NEMW: Northeast Midwest

NEMWI: Northeast-Midwest Institute

NGWMN: National Ground-Water Monitoring Network

NHDPlus: National Hydrography Dataset Plus

NORM: Naturally Occurring Radioactive Material

NURE: National Uranium Resource Evaluation

NWIS: National Water Information System

NYSDEC: New York State Department of Environmental Conservation

PADEP: Pennsylvania Department of Environmental Protection

PFBC: Pennsylvania Fish and Boat Commission

QA/QC: Quality Assurance/Quality Control

RSE: Residual Standard Error

RWQMN: Remote Water Quality Monitoring Network

SRBC: Susquehanna River Basin Commission

SSURGO: Soil Survey Geographic Database

STAC: Scientific and Technical Advisory Committee

STEWARDS: Sustaining The Earth’s Watersheds - Agricultural Research Database System

STORET: STOrage and RETrieval

TAC:	Technical Advisory Committee
TDP:	Total Dissolved Phosphorus
TDS:	Total Dissolved Solids
TP:	Total Phosphorus
TU:	Trout Unlimited
USDA:	U.S. Department of Agriculture
USEPA:	U.S. Environmental Protection Agency
USGS:	U.S. Geological Survey
WQN:	Water Quality Network
WQX:	Water Quality Exchange

Glossary

Baseflow: Baseflow is the portion of streamflow coming from groundwater sources.

Cement: Cement is used to grout well casings to the walls of the well bore in order to seal the formation.

Conventional wells: Conventional oil and gas wells are typically vertical and target relatively permeable formations like sandstone or carbonate rocks. Conventional wells typically use low volumes of hydraulic fracturing fluids (50,000-100,000 gallons) and target discrete pay zones (the rock in which oil or gas are recoverable) over relatively short intervals (measured in tens of feet). (See also “unconventional well.”)

Domestic wells: Wells that are used for drinking water supply in private homes.

Ecoregions: Ecoregions are used by USEPA to designate geographically defined regions in which the type, quality, and quantity of environmental resources are generally similar. Ecoregions may respond similarly to environmental stressors.

Fracturing fluids: Fluid and material injected into an HVHF well as part of the hydraulic fracturing process.

Gas storage wells: Depleted oil and gas fields can serve as reservoirs for natural gas. In Pennsylvania, the Oriskany formation is a commonly used natural gas storage reservoir in which natural gas is stored under pressure, often using depleted fields and retrofitted vertical gas wells.

Groundwater sampling wells: Groundwater sampling wells include all types of wells, such as domestic, public supply, industrial, irrigation, monitoring, etc. and comprise the sites (along with springs) from which data were obtained for the Susquehanna data set.

Groundwater monitoring site: A dedicated sampling site for monitoring groundwater quality. Not located at a domestic well.

Groundwater sampling site: Any site (well or spring) where groundwater is sampled.

Hydraulic Fracturing Well: Any well that uses hydraulic fracturing technology to increase well production, including conventional vertical and unconventional horizontal wells.

HVHF (High-volume hydraulic fracturing): In this report, horizontal unconventional wells are presumed to require high-volume hydraulic fracturing and are referred to as “HVHF” wells. HVHF wells employ hydraulic fracturing at multiple stages along horizontal sections of unconventional wells that can extend several thousand feet in the target formation. In the Susquehanna River Basin, HVHF wells typically use high volumes (2-10 millions of gallons) of hydraulic fracturing fluids in the Marcellus formation.

Network analysis: A monitoring method used by USGS's NAWQA program in order to determine if change in groundwater quality has occurred in a given aquifer or hydrogeologic/land-use setting. The network analysis method requires at least 25 to 30 groundwater sampling sites per aquifer and (or) land-use setting, and involves a minimum of a paired two-sample comparison between first and second sampling events sampled over approximately one decade apart.

Open hole: Any well that has not been cased (lined), or portion of a well that is not lined. In a water-supply well, open hole wells can be found in productive bedrock formations.

Parameter: In this document, a generic term referring to water-quality measurements, including field measurements and water-quality constituents analyzed in a laboratory.

Principal aquifer: A principal aquifer is "a regionally extensive aquifer or aquifer system that could be used as a source of potable water. An aquifer is a geologic formation, a group of formations, or part of a formation that contains sufficient saturated permeable material to yield significant quantities of water to wells and springs. Aquifers are often combined into aquifer systems (U.S. Geological Survey, 2014a)."

Produced water and "flowback" water: "Produced" water is defined in this report as any water that returns to the surface (others have used the term "shale gas wastewater") during shale gas development processes. This broad definition of produced water includes water that returns to the surface after hydraulic fracturing is completed and before well production commences—which some have defined as "flowback" water.

Shale Gas Development: The complete life-cycle of gas development activities for hydrocarbon extraction from shale formations, beginning with clearing and construction and including drilling, hydraulic fracturing, disposal of waste fluids, well production and depletion, well-closure and site remediation, and long-term post-closure monitoring.

Stray gas: For the purpose of this report, stray gas is defined as the migration of methane into groundwater.

Streamgages: An active, continuously-functioning measuring device in the field for which a daily mean streamflow is computed or estimated and quality assured for at least 355 days of a water year (U.S. Geological Survey, 2014a).

Thermogenic and microbial methane: Thermogenic methane is formed also from organic sources, but at high temperatures and pressures that are found deep in the earth. Microbial methane is formed in relatively shallow aquifers from the decomposition of organic material, as can occur in swamps or in landfills.

Tophole: After construction of the well pad, a small drilling rig will move in and drill the vertical section of the well known as the "tophole."

Type I Error: Detecting an effect that is not present (incorrectly rejecting the null hypothesis).

Type II Error: Failing to detect an effect that is present (incorrectly failing to reject the null hypothesis).

Unconsolidated and bedrock aquifers: Unconsolidated aquifers are composed of materials that are not cemented together, like sand or gravel. Bedrock aquifers are consolidated and can occur in either porous sedimentary rock like sandstone, limestone, dolomite, shale or siltstone, or in non-porous but fractured igneous or metamorphic rock.

Unconventional Well: These wells target low permeability source-rock formations such as shale, coal seams, tight sands, and others. They can be vertical or horizontal wells. Horizontal unconventional wells typically require high volumes of hydraulic fracturing fluids. Some unconventional vertical wells may be test wells. (See also "conventional well.")

Vertical Well: This is a well-drilling method used over the past 100 years. Single-stage hydraulic fracturing has been used to increase production in vertical conventional wells since the 1950's.

Water types: The dissolved constituents in water are primarily composed of positive and negative ions. Water can be classified into types based on the dominant dissolved cations and anions. Some contaminants are more closely associated with certain water types because of the geochemical conditions.

Conversion Factors

Conversion from English system to International System of Units

Multiply	By	To obtain
	Length	
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
	Area	
square mile (mi^2)	259.0	hectare (ha)
square mile (mi^2)	2.590	square kilometer (km^2)
	Volume	
cubic foot (ft^3)	0.02832	cubic meter (m^3)
	Flow rate	
foot per second (ft/s)	0.3048	meter per second (m/s)
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

Conversion from International System of Units to English system units

Multiply	By	To obtain
	Length	
centimeter (cm)	0.3937	inch (in.)
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)
	Area	
square meter (m^2)	0.0002471	acre
square kilometer (km^2)	247.1	acre
square meter (m^2)	10.76	square foot (ft^2)
square kilometer (km^2)	0.3861	square mile (mi^2)
	Volume	
liter (L)	33.82	ounce, fluid (fl. oz)
liter (L)	0.2642	gallon (gal)
liter (L)	61.02	cubic inch (in^3)
	Mass	
gram (g)	0.03527	ounce, avoirdupois (oz)
kilogram (kg)	2.205	pound avoirdupois (lb)

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:

°F=(1.8×°C)+32

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as follows:

°C=(°F-32)/1.8

Vertical coordinate information is referenced to the "North American Vertical Datum of 1988 (NAVD 88)"

Horizontal coordinate information is referenced to the "North American Datum of 1983 (NAD 83)"

Specific conductance is given in microsiemens per centimeter at 25 degrees Celsius (µS/cm at 25°C).

Concentrations of chemical constituents in water are given either in milligrams per liter (mg/L) or micrograms per liter (µg/L).

1. Executive Summary

Throughout its history, the United States has made major investments in assessing natural resources, such as soils, timber, oil and gas, and water. These investments allow policy makers, the private sector and the American public to make informed decisions about cultivating, harvesting or conserving these resources to maximize their value for public welfare, environmental conservation and the economy. As policy issues evolve, new priorities and challenges arise for natural resource assessment, and new approaches to monitoring are needed. For example, new technologies for oil and gas development or alternative energy sources may present new risks for water resources both above and below ground. There is a need to evaluate whether today's water monitoring programs are generating the information needed to answer questions surrounding these new policy priorities.

The Northeast-Midwest Institute (NEMWI), in cooperation with the U.S. Geological Survey (USGS) National Water-Quality Assessment (NAWQA) Program, initiated this project to explore the types and amounts of water data needed to address water-quality related policy questions of critical concern to today's policy makers and whether those data are currently available. The collaborating entities identified two urgent water policy questions and conducted case studies in the Northeast-Midwest region to determine the water data needed, water data available, and the best ways to fill the data gaps relative to those questions. This report details the output from one case study and focuses on the Susquehanna River Basin, a data-rich area expected to be a best-case scenario in terms of water data availability.

1.1 Case Study Question

The policy question that frames this case study evaluating water monitoring capacity is: ***Do shale gas development activities contaminate surface water or groundwater in the Susquehanna River Basin?*** This question is of urgent importance to decision-makers in the Marcellus Shale region. The number of unconventional wells in Pennsylvania grew from less than 200 in 2007 to greater than 9,300 as of August 2015 (Pennsylvania Department of Environmental Protection, 2015a). The Susquehanna River Basin has 63 percent forest cover, a population of more than 4 million people, and more than 49,000 miles of rivers, streams, creeks, and brooks (Susquehanna River Basin Commission, 2014a). The rapid growth of high-volume hydraulic fracturing (HVHF) in the Susquehanna River Basin has raised concerns about the potential for degraded surface-water quality and potential impacts on drinking water aquifers throughout the basin. Because hydraulic fracturing involves injection of high volumes of water and chemicals deep underground at high pressures, it has raised questions about potential contamination of surface water and groundwater and human and environmental health and safety implications. A divisive debate has emerged between citizens who support shale gas development and the economic development it brings, and others who oppose the practice based on water-quality and other environmental concerns. This conflict remains unresolved among the public due in part to a lack of information regarding water-quality and other potential environmental impacts.

Understanding the water-quality impacts of shale gas development is also a priority for the Northeast-Midwest region, generally, as evidenced in an informal project survey of Northeast-Midwest Congressional

Coalition member offices where many asked whether contaminants from shale gas development are reaching surface water or groundwater. Member offices want this information to inform appropriate decisions about regulation and management of shale gas development, and so they can inform their constituents about the impacts of development. It is clear that understanding the potential impact of unconventional oil and gas development and hydraulic fracturing on water quality is an emerging national concern. A recent draft report by the U.S. Environmental Protection Agency (USEPA) studying the impacts of hydraulic fracturing on drinking water resources nationwide generated significant interest and concern from both sides of the debate with its major conclusions: the draft report did not find evidence of widespread systemic impacts on drinking water in the United States; however, the limited amount of data collected before and during hydraulic fracturing activities reduced the USEPA's ability to determine whether hydraulic fracturing affected water quality in cases of alleged contamination (U.S. Environmental Protection Agency, 2015a).

Despite this high-profile interest, the question of whether shale gas development contaminates surface water or groundwater remains largely unresolved. Government agencies, academia, and volunteer groups have worked to collect water-quality data as they are able since shale gas development started in the Susquehanna River Basin in 2007, but it is not clear whether the systematic effort necessary to generate answers to questions surrounding shale gas development and water quality has been made. In summary, this case study topic represents an urgent environmental and public health priority for the Northeast-Midwest region, a national energy policy priority, and an issue for which a comprehensive and analytical look at water-quality data needs and data availability has the potential to create a consensus regarding the data needed to inform future management actions and policy solutions.

1.2 Case Study Approach

The case study approach consisted of three main tasks:

1. Describe the water-quality data, both types and amounts, needed to answer the case-study policy question,
2. Assess the extent to which those data are available and usable, and
3. Identify the additional data that would be needed to more effectively answer the case-study policy question, and estimate the level of effort to collect those data.

For the first task, the project team and technical advisory committee (TAC) identified the data types that were most critical for measuring water-quality change related to shale gas development for both surface water and groundwater, and analyzed available data to determine the quantity of data needed to detect statistically significant changes in water quality. The project team collected data available from agencies and organizations that monitor water quality within the Susquehanna River Basin, evaluated those data against the data needed, and reviewed the usability of those data to help answer the policy question. Once data availability and usability were determined, the project team assessed data gaps and made recommendations for collecting needed data to answer the policy question.

This report identifies the minimum amount of water data needed to detect statistically significant change in water quality related to shale gas development in the Marcellus Shale region, given certain study design assumptions; identifies available water data; and makes recommendations for filling the information gaps.

1.3 Case Study Findings

The case study findings relative to water data needed, water data available and usable, and approaches to filling the data gaps are summarized below.

1.3.1 Water data needed to answer the policy question

The following criteria were used to identify the most important water-quality data needed to answer the case-study policy question:

- **Water data must satisfy study design criteria to be used to answer the case-study policy question.**

It is not possible to identify water data needs for answering the case-study policy question without an initial discussion of an appropriate study design (Figure ES-1). The right water data must be available in the right locations with the right supporting information to detect water-quality change and identify the cause of that change.

The TAC decided to focus on the ability to detect the cumulative effects of shale gas development because water-quality monitoring is critical to identifying this type of contamination. Cumulative effects of shale gas development are defined in this study as the combination of the long-term impacts of the individual steps of shale gas development and the associated potential contamination pathways taken together, and the cumulative impacts of multiple well pads within a geographic area. Short-term spills of concentrated brines are likely to be ephemeral, with spikes in concentrations. Acute incidents are best detected through on-site monitoring, accident reporting with targeted response monitoring, and through continuous water-quality monitors in receiving streams. Cumulative effects of shale gas development were selected as the study focus because these are more subtle to detect and water-quality monitoring is the only path to identifying low level and long-term changes. In the absence of water-quality data, the long-term cumulative effects of shale gas development on water quality will be unknown.

The location of monitoring sites in the right places for answering the case-study policy question is critical. Surface water monitoring sites must be located in each of the four ecoregions with active or planned shale gas development, because stream chemistry in each ecoregion is unique and will respond differently to disturbances or changes in land use. Monitoring sites must be located in watersheds with HVHF wells and in reference watersheds in each ecoregion. In this context, reference watersheds are areas with no HVHF well development. Monitoring sites in these types of watersheds allow for the detection of water-quality changes in watersheds with HVHF well development and comparison with undeveloped watersheds to identify whether water-quality changes are resulting from HVHF development. Ideally, monitoring sites will measure water quality in watersheds smaller than 70 square miles. Small watersheds provide the best opportunities to identify pollutants that are primarily derived from a single source. Monitoring in larger watersheds can provide opportunities for nested monitoring with smaller watersheds that quantifies the transport of pollutants downstream and pollutant inputs from multiple small watersheds with active shale gas development.

Water-quality and streamflow data at these monitoring sites must be available with sufficient sampling frequency and duration to evaluate trends in concentration over time. Finally, data on shale gas development, geology, climate, and other changes in land use throughout the monitored watershed must be available to correlate water-quality change with shale gas development activity. Without this information, the relationship between shale gas development and water quality cannot be evaluated, even if shale gas development is causing water-quality change.

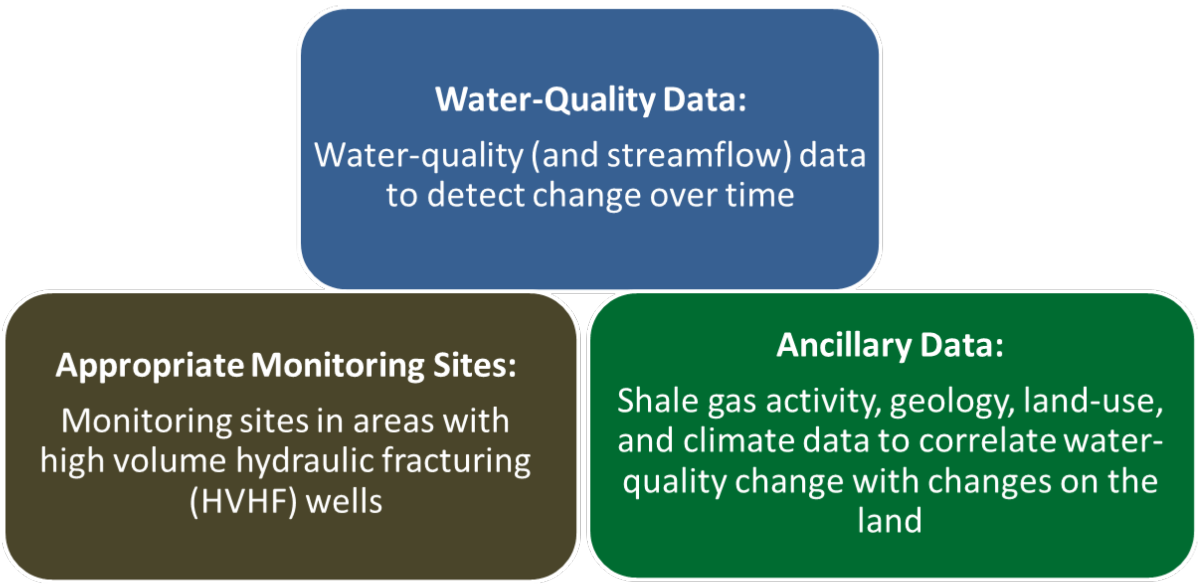


Figure ES-1. Study design needed to answer "Do shale gas development activities contaminate surface water or groundwater?"

The study design for groundwater again requires monitoring sites in the right places: networks of groundwater sampling sites are needed with each sampling site located within 1 mile of an HVHF well. Water-quality data collected before and after shale gas development at the nearby HVHF well sites are necessary to identify groundwater quality change. Again, information on the shale gas development, geology, other changes in land use, and climate near those sampling sites must be available to correlate water-quality change with shale gas development activity.

- **Multiple lines of evidence, using a suite of priority monitoring parameters, are needed to identify shale gas development as the source of water-quality change.**

No one parameter can identify whether shale gas development is the source of contamination if a change in that parameter concentration is detected. By monitoring a suite of parameters, more information is available to identify the likely source of contamination. Each shale gas development activity, from initial well pad development to production of gas from a completed well is associated with different pathways that contribute different potential contaminants, so monitoring for just one of the parameters would miss multiple types of potential contamination.

> ***Suite of priority parameters***
>
> **Parameters recommended for both surface water and groundwater:** alkalinity, dissolved barium, bromide, calcium, chloride, dissolved oxygen, gross alpha, gross beta, lithium, magnesium, nitrate, pH, potassium, radium-226 and -228 (but only if there are changes in gross alpha and gross beta), sodium, specific conductance, strontium, sulfate, total dissolved solids, uranium, and water temperature.
>
> **Additional parameters recommended for surface water only:** total barium, suspended sediment concentration, total organic carbon, total phosphorus, turbidity, and streamflow.
>
> **Additional parameters recommended for groundwater only:** benzene, toluene, ethylbenzene, xylene, and methane.

Consequently, a suite of water-quality parameters is needed to determine if contamination from the cumulative impact of shale gas development activities has occurred in the Susquehanna River Basin. The suites of priority parameters for surface water and groundwater are based on the specific hydrology, geology, past and current land use, and other environmental concerns expressed in the Susquehanna River Basin.

- **The sampling frequency and duration of monitoring must meet minimum requirements to adequately characterize and detect changes in priority monitoring parameters related to shale gas development.**

Assuming that monitoring sites located in watersheds with HVHF wells and in reference watersheds and the needed ancillary data are available (Figure ES-1), surface-water quality and streamflow data requirements can be characterized as shown in Table ES-1. As mentioned above, the entire suite of surface-water parameters is needed to identify whether shale gas development is the source of water-quality change. Monthly sampling frequency is needed to detect changes in water quality year round and to minimize the time needed to detect statistically significant water-quality change at each monitoring site (see section 6.4). A minimum of eight surface-water monitoring sites are needed: one monitoring site in a watershed with HVHF wells and one reference watershed monitoring site is needed in each of the four ecoregions with active or predicted HVHF development. Additional monitoring sites will provide critical information regarding the scope and magnitude of potential water-quality change associated with shale gas development, especially in watersheds with the highest density of HVHF wells and nested watershed monitoring sites.

The minimum monitoring duration must be sufficient to characterize background surface-water concentrations for each parameter and detect statistically significant change over normal background fluctuations. The minimum monitoring duration to detect change varies by ecoregion and is discussed in detail in section 6.6. Because the purpose of the monitoring described in Table ES-1 is to detect whether the cumulative effects of shale gas development are resulting in water-quality change, monitoring should continue at selected monitoring sites for the long-term, as long as shale gas development activities continue in the Susquehanna River Basin. The magnitude of water-quality change that could occur from contamination related to shale gas development is unknown, but it would take 3-6 years of monthly monitoring to detect a 20-percent change in median specific conductance or total barium in the Susquehanna River Basin.

Table ES-1. Summary of surface-water data needed to detect water-quality change resulting from cumulative shale gas development activities in the Susquehanna River Basin. Analysis supporting these findings is presented in Chapter 6.
[Abbreviations: HVHF, High-volume hydraulic fracturing]

Criteria	Surface-water data needed
Monitoring parameters	• Suite of priority surface-water parameters from Table 3 and streamflow at each monitoring site
Sampling frequency	• Monthly
Locations of monitoring sites	• Monitoring sites in each of the ecoregions with active or predicted HVHF activity, including: o Northern Allegheny Plateau, o North Central Appalachians, o Central Appalachians, and o Ridge and Valley.
Watershed characteristics	• Watersheds smaller than 70 square miles. • Medium and high density and reference watersheds: o Watersheds with greater than 0.5 HVHF wells per square mile, and o Watersheds with 0 HVHF wells per square mile and no significant shale gas development expected. • Watersheds larger than 70 square miles that offer opportunities for nested monitoring (i.e. one or more small watersheds that are being monitored for change are nested within the larger watershed)
Number of monitoring sites	• Minimum of 1 monitoring site in a high density watershed per ecoregion. • Minimum 1 reference monitoring site per ecoregion.
Duration and timing of monitoring	• At least 36 samples collected at monthly or longer intervals over 3-4 years including data collected after shale gas development (post-2007) • Minimum duration of monitoring to detect water-quality change varies by ecoregion • Ideal monitoring sites will have: o Data collected before shale gas development (pre- 2007), o An uninterrupted data record, o Current/ongoing data collection (2009 or later), and o Plans for long-term monitoring.

Groundwater data requirements can be characterized as shown in Table ES-2. Again, the entire suite of groundwater parameters is needed to identify whether shale gas development is the source of groundwater quality change. The study design for groundwater that is most applicable for existing groundwater data requires networks of 25-30 groundwater sampling sites in each of the primary drinking water aquifers in the Susquehanna River Basin, each site located within 1 mile of an HVHF well. Five priority spatial networks were identified for groundwater monitoring. Water-quality data should be collected at those sampling sites before and after shale gas development, with sampling events separated by approximately 10 years. Long-term monitoring would require that additional sampling be done every 10 years into the future. A subset of 5 sampling sites in each network should be sampled every 2 years to identify interim water-quality changes.

Table ES-2. Summary of groundwater data needed to detect water-quality change resulting from cumulative shale gas development activities in the Susquehanna River Basin. Analysis supporting these findings is presented in Chapter 9.

Criteria	Groundwater data needed
Monitoring parameters	• Suite of priority groundwater parameters from Table 3 at each monitoring site.
Spatial networks	• Minimum of 5 networks in each of the major drinking water aquifers with shale gas development, distinguished by topography: o Upper Devonian Lock Haven aquifer with upland topography, o Upper Devonian Lock Haven aquifer with valley topography, o Upper Devonian Catskill aquifer with upland topography, o Upper Devonian Catskill aquifer with valley topography, and o Pleistocene deposits aquifer.
Number and location of sampling sites	• For each network: o 25-30 sampling sites o Each site within 1 mile of a HVHF well
Duration, frequency, and timing of monitoring	• Two samples at each site, separated by approximately 10 years and taken: o before shale gas development, and o after shale gas development • Additional long-term monitoring, in subsequent 10 year increments, • A subset of 5 sites per network sampled every 2 years.

1.3.2 Availability and usability of existing water data to answer the policy question

This investigation found more than 960,000 surface-water records collected at about 14,700 monitoring sites over the last 85 years in the Susquehanna River Basin (Figure ES-2 (A)). However, Figure ES-2(B) shows there are only 10 monitoring sites in the Susquehanna data set compiled for this study that meet most of the criteria summarized in Table ES-1 for barium, an indicator closely associated with HVHF development, and none of these sites are located in watersheds with a high density of HVHF wells. There are no monitoring sites that meet all the criteria in Table ES-1.

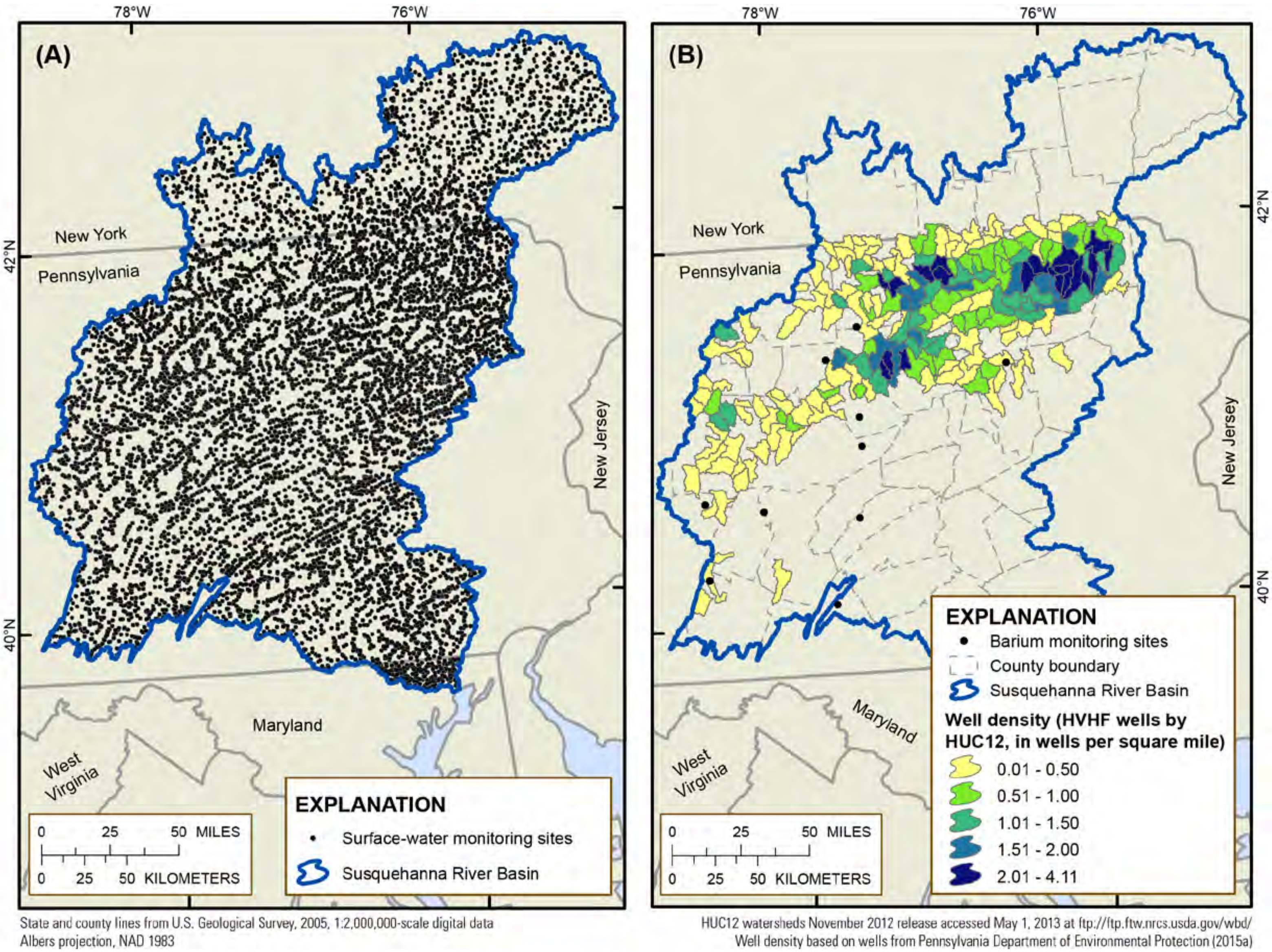


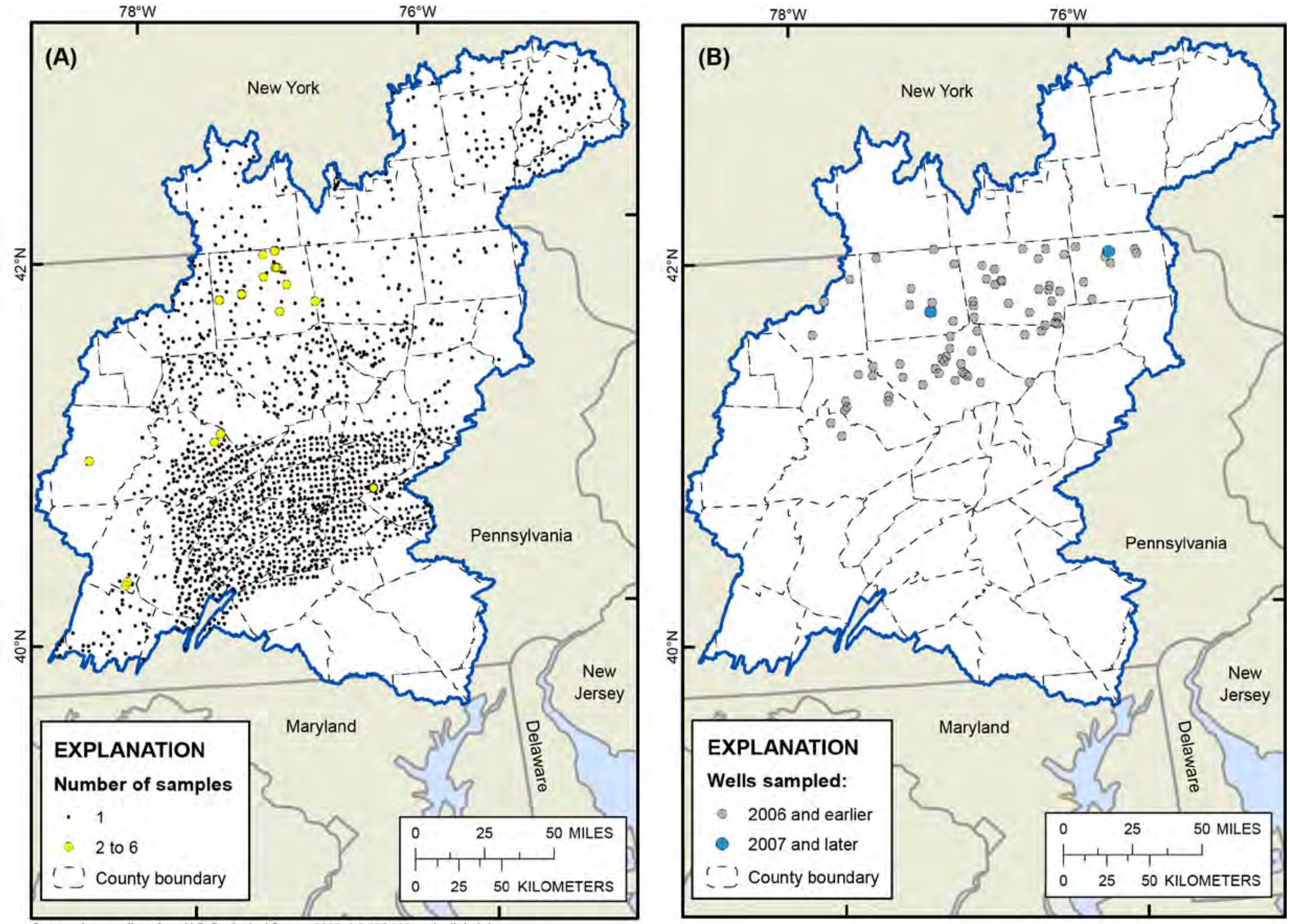
Figure ES-2. (A) Surface-water monitoring sites in the Susquehanna River Basin with water-quality records for at least one of the comprehensive list of surface-water parameters selected by the Technical Advisory Committee (n=14,730), and (B) Surface-water monitoring sites where the minimum data for detecting changes in barium concentration (total or dissolved) have been collected (n=10). None of the 10 monitoring sites are located in a watershed with a high density of HVHF wells.

- **The surface-water data needed for answering the case-study policy question are not currently available in the Susquehanna River Basin.**

While there are some applicable surface-water data available, each of the existing monitoring sites does not meet at least one of the criteria for location, parameters analyzed, frequency of monitoring, or duration of monitoring to detect statistically significant change associated with cumulative effects of shale gas development. Monitoring sites that have used an appropriate sampling plan for answering the case-study policy question are not in the right locations for detecting water-quality change related to shale gas development. Only 4 of 22 surface-water monitoring sites in the Susquehanna River Basin with enough existing data for a water-quality trend analysis for barium or specific conductance are located in watersheds with active HVHF wells, and few of the 26 recommended surface-water monitoring parameters are available for those sites. Only one of those monitoring sites is in a watershed with an HVHF well density greater than 0.5 HVHF wells per square mile. The existing surface-water data in the Susquehanna data set are not sufficient to detect whether the cumulative effects of shale gas development are resulting in water-quality change.

- **The groundwater data needed for answering the case-study policy question are not being collected.**

There is no systematic, large-scale, long-term monitoring effort underway to assess the effects of shale gas development on groundwater quality in the Susquehanna River Basin, and from the data sources that do exist, Figure ES-3 shows that limited groundwater data are publicly available to answer the policy question. The groundwater sampling sites with existing data are rarely located within 1 mile of an HVHF well, but even when they are in the right locations the sites lack data for most of the priority groundwater parameters. The available groundwater data lack the sampling frequency needed for a water-quality trend analysis and lack the number and location of sampling sites needed for a spatial water-quality network analysis. Selecting appropriate groundwater sampling sites is a major challenge for monitoring agencies because they do not have access to information on the location of future HVHF wells. Sampling sites within 1 mile of an HVHF well are due more to coincidence than due to planning.



State and county lines from U.S. Geological Survey, 2005, 1:2,000,000-scale digital data
Albers projection, NAD 1983

Figure ES-3. (A) Groundwater sampling sites with bromide data in the Marcellus and Utica Shale area of the Susquehanna River Basin (n=1,686), and (B) Groundwater sampling sites located within 1 mile of an HVHF well (n=74). None of the 74 sampling sites have data available both before and after shale gas development, and few of the suite of priority groundwater parameters were measured before shale gas development.

- **Current water data usability for answering the policy question is limited by insufficient data documentation and availability.**

Water-quality monitoring programs are usually designed to meet a stated objective or follow a historical precedent. Data collected for one monitoring objective may not be directly applicable to another objective, due to the location of monitoring sites, frequency of monitoring, parameters measured, and analytical methods used. The surface-water data identified through this case study were generated by 35 organizations, and groundwater data were collected by 10 organizations that collect water-quality data for parameters related to shale gas development in the Susquehanna River Basin. Insufficient and inconsistent documentation of available data limited the utility of these existing data sets. Substantial project time and effort over the course of this multi-year project were required to locate, obtain, and consistently format data. Missing information that is particularly important for this case study includes specification of whether water-quality samples were filtered or unfiltered, and information on the aquifer from which groundwater samples were taken.

Data sharing and data accessibility were also limiting factors in data availability in this case study. It is possible that despite the work completed for this case study, additional relevant data that are not being shared or are not available in electronic format may exist. The Water Quality Portal (National Water Quality Monitoring Council, 2014a), a cooperative service that provides publicly available water-quality data from Federal databases, including data collected by more than 400 State, Federal, Tribal, and local organizations, was established to facilitate water data sharing. Data collected at only 19 percent of the surface-water monitoring sites identified through this case study in the Susquehanna River Basin are available through the Water Quality Portal, but 85 percent of the water-quality data records are available through the Portal. This finding indicates that the monitoring sites from the Susquehanna data set that are available through the Water Quality Portal are sites with longer data records, monitoring sites that are more likely to have the data necessary to identify a water-quality trend. Still, important data sets are missing from this collection including data collected by volunteer organizations, local governments, and academia. The Water Quality Portal includes data collected at 45 percent of the groundwater monitoring sites identified through this case study, and 61 percent of the groundwater quality records are available through the Portal. The most important groundwater data sets, those collected by industry, are not available through data sharing systems and access to those data sets is limited.

1.3.3 Approaches for filling data gaps to answer the case-study policy question

- **Increase monitoring at a minimum of 8 targeted surface-water monitoring sites; additional monitoring sites are highly recommended. The water data identified in Table ES-1 must be collected for each of these sites.**

A recently established surface-water monitoring program at the Susquehanna River Basin Commission (SRBC) called the Remote Water Quality Monitoring Network (RWQMN) is designed to collect surface-waterquality data related to shale gas development, and recent updates to a monitoring program at the Pennsylvania Department of Environmental Protection (PADEP) collect data more closely associated with shale gas development than previous monitoring efforts. Many of the monitoring sites for these programs are in the right locations, but additional sampling frequency, parameters, and streamflow data are needed. For SRBC monitoring sites, a minimum of 4 additional years of monitoring are needed at current sampling frequencies before water-quality trends can begin to be detected. Increased monitoring at a subset of priority monitoring sites would start generating answers to the case-study policy question in a time-frame that is useful to decision makers and would provide the water data needed to identify whether shale gas development is contaminating surface water in the monitored watersheds.

A minimum of 8 monitoring sites are needed, 2 in each of the ecoregions with active or planned HVHF development. Additional monitoring sites will provide critical information regarding the scope and magnitude of potential water-quality change associated with shale gas development, especially in watersheds with the highest density of HVHF wells and nested watershed monitoring sites. An example set of 17 priority surface-water monitoring sites is presented in Chapter 8. The following list includes several site characteristics that could make an individual monitoring site a priority site for increased monitoring:

- History of water-quality data available at the site
- Variety of priority surface-water parameter data available at the monitoring site
- Density of HVHF wells in the monitored watershed
- Availability of existing continuous monitors or streamgages at the monitoring site
- Availability of nested monitoring sites
- Current or long-term funding source available for the monitoring site
- History of cooperation with shale gas development companies that are active within the monitored watershed

The increased monitoring must include analysis for the full suite of priority surface-water parameters, increased frequency to monthly sampling, and addition of streamgages where needed. An example of data needs is presented in Table 19.

- **Maintain data collection and analysis at enhanced surface-water monitoring sites for a minimum of 10 years and as long as shale gas development activities continue in the Susquehanna River Basin.**

Monitoring at enhanced priority sites should continue for more than 10 years to determine whether cumulative shale gas development activities are resulting in water-quality change over the long-term. Monitoring to fill the data gaps should be implemented using an adaptive-monitoring approach and coordinated among monitoring agencies. Care should be taken to coordinate closely among participating monitoring programs so that data collection can be planned for compatibility, sharing, and easy analysis. The data should be evaluated on a regular basis so the monitoring program can be adjusted as necessary to adapt to a changing understanding of shale gas development and water quality.

- **Design and implement a systematic, long-term groundwater monitoring program for detecting groundwater quality change related to shale gas development in the Susquehanna River Basin, building on data collected by the shale gas industry, if appropriate.**

There is no groundwater monitoring equivalent to the SRBC RWQMN that is investigating the potential for shale gas development to change groundwater quality across the Susquehanna River Basin, and the groundwater data compiled for this study cannot be used as the foundation of a new groundwater monitoring program as described above. The groundwater data summarized in Table ES-2 must be collected through a systematic monitoring program to be able to answer the case-study policy question.

The shale gas industry has collected the most comprehensive set of groundwater data that pre-dates development at HVHF wells in the Susquehanna River Basin. The use of these existing data sets as the foundation of a new long-term groundwater monitoring program would result in the most cost-effective and most timely approach for collecting groundwater-quality data to answer the case-study policy question if the data sets meet minimum requirements for statistical analysis. Most of the priority groundwater parameters would need to be available for each sampling site (Table 3), the source aquifer identified, and the filtered/unfiltered status specified for each parameter.

If access to industry data cannot be obtained, a completely new groundwater monitoring program must be initiated that would also require industry participation to identify appropriate sampling sites based on plans for future shale gas development.

- **Establish a coordinating entity to develop and implement surface-water and groundwater monitoring plans in the Susquehanna River Basin, with representation from water monitoring organizations, shale gas industry, domestic well owners, and public citizens.**

To answer the case-study policy question, the right water-quality data must be collected in the right locations with the right supporting information. Water monitoring organizations and academia are tasked with collecting the right water-quality data, but industry involvement is necessary to identify the right monitoring locations and provide the right supporting information. Industry cooperation is needed to provide access to existing data and to identify locations of new HVHF wells so appropriate sampling sites can be identified. Ongoing coordination between water monitoring agencies and the shale gas industry will provide the necessary updates on locations of active production and technological advances in shale gas extraction practices that may affect water monitoring strategies.

Because domestic wells are the most accessible locations for monitoring groundwater quality related to shale gas development, it is critical to include domestic well owners in the process to obtain access to domestic wells for sampling before and after new shale gas development. To gain public trust, water monitoring programs should engage the people living in affected areas in addition to independent experts. According to the Council of Canadian Academies (2014), citizens will have greater faith in water monitoring results if they can influence the design, access the results, and comment throughout the process.

Engagement of all the critical stakeholders (water monitoring organizations, shale gas industry, citizens, and domestic well owners) will improve water data coordination, sharing, and analysis to better understand the water-quality impacts of shale gas development in the Susquehanna River Basin. Stakeholder participation will promote confidence in the analytical results and lead to a better understanding of the risks to water resources in the Susquehanna River Basin. With this information, real policies and regulations can be implemented to protect water resources in the Susquehanna River Basin.

A coordinating entity is needed to facilitate coordination of sampling plans among water monitoring organizations so data collection, analysis, and interpretation will be compatible and comparable across monitoring sites. Improved data documentation and data sharing will facilitate the use of water data for answering the case-study policy question. Tools such as the Consortium of Universities for the Advancement of Hydrologic Sciences, Inc. HydroDesktop (Consortium of Universities for the Advancement of Hydrologic Sciences, Inc., 2014), Water Quality Exchange (WQX) (U.S. Environmental Protection Agency, 2014) and the Water Quality Portal provide the infrastructure for organizations to format and share their data, but greater participation is needed. Consistent, thorough data documentation and wider availability of data sources through these services will increase the value of water-quality data from all monitoring agencies and reduce the amount of time needed to access and prepare data for new applications. A continued commitment to water-quality data-sharing systems is essential for maximizing use of existing water-quality data.

1.4 Conclusion

Targeted, robust monitoring networks for both surface water and groundwater are critical for identifying whether the influx of shale gas development activity in the Susquehanna River Basin is generating adverse changes in water quality. The existing water-quality data in the Susquehanna River Basin are inadequate to serve this purpose. This report presents strategies for collecting the water data needed to detect whether shale gas development activities are contaminating surface water or groundwater in the Susquehanna River Basin in a policy-relevant time frame. Key steps to generating the needed information include increased monitoring in strategic locations, design and implementation of a systematic groundwater monitoring program, and a long-term commitment to water-quality monitoring in the Susquehanna River Basin supported by a coordinating entity. Water-quality data collection and analysis, with participation from the key stakeholders, can answer this urgent policy question of critical importance to the Northeast-Midwest region and prepare for the questions that will emerge with further growth of the shale gas industry. The sooner the region gets started, the better.

Summary of information needs to answer "Do Shale Gas Development Activities Contaminate Surface Water or Groundwater in the Susquehanna River Basin?"

- Increase monitoring at a subset of targeted surface water monitoring sites.
- Maintain data collection and analysis at enhanced surface water monitoring sites for a minimum of 10 years and as long as shale gas development activities continue in the Susquehanna River Basin.
- Design and implement a systematic, long-term groundwater monitoring program, building on data collected by shale gas industry, if appropriate.
- Establish a coordinating entity to develop and implement surface water and groundwater monitoring plans in the Susquehanna River Basin, with representation from water monitoring organizations, the shale gas industry, domestic well owners, and public citizens.

2. Case Study Introduction and Background

2.1 Purpose of the Case Study

Throughout its history, the United States has made major investments in assessing natural resources, such as soils, timber, oil and gas, and water. These investments allow policy makers, the private sector and the American public to make informed decisions about cultivating, harvesting or conserving these resources to maximize their value for public welfare, environmental conservation and the economy. As policy issues evolve, new priorities and challenges arise for natural resource assessment, and new approaches to monitoring are needed. For example, new technologies for oil and gas development or alternative energy sources may present new risks for water resources both above and below ground. There is a need to evaluate whether today's water monitoring programs are generating the information needed to answer questions surrounding these new policy priorities.

The Northeast-Midwest Institute (NEMWI), in cooperation with the U.S. Geological Survey (USGS) National Water-Quality Assessment (NAWQA) Program, initiated this project to explore the types and amounts of water data needed to address water-quality related policy questions of critical concern to today's policy makers and whether those data are currently available. The collaborating entities identified two urgent water policy questions and conducted case studies in the Northeast-Midwest region to determine the water data needed, water data available, and the best ways to fill the data gaps relative to those questions. This report details the output from one case study and focuses on the Susquehanna River Basin, a data-rich area expected to be a best-case scenario in terms of water data availability.

The policy question that framed this case study is: ***Do shale gas development activities contaminate surface water or groundwater in the Susquehanna River Basin?*** This case study investigates the water data needed to answer this specific, timely policy question and whether water-quality data are available to answer this question.

The consideration of whether water-quality data can be applied to issues beyond the purpose for which they were originally collected was considered for both technical and practical purposes. Water-quality monitoring programs are usually designed to meet a stated objective or follow a historical precedent. Data collected for one monitoring objective may not be directly applicable for another objective, due to the location of monitoring sites, frequency of monitoring, parameters measured, and methods used to meet that objective. The technical requirements of a new policy question must be considered when determining if water-quality data can be used for a purpose other than its original objective. In addition, this case study examines the practical considerations of whether water-quality data collected by different monitoring agencies in different jurisdictions can be taken together to support regional decision-making.

2.2 Case Study Approach and Organization of the Report

Shale gas development was selected by the project team as the case study subject due to the importance of the issue to policy makers and the NEMW region. The NEMWI interviewed Congressional and Senate staff and decision-makers participating in the Blue Ribbon Project Steering Committee to identify their most urgent water-quality questions related to shale gas development to help design relevant case studies. Twelve Congressional offices and 18 Steering Committee members, including representatives from the Federal Government, States, cities, industry, and environmental groups, participated in these discussions and their responses were distilled into priority policy questions.

While the technical potential for shale gas development has been apparent for some time, it is only recently that horizontal-drilling technology and multi-stage hydraulic fracturing has allowed for the exploration and development of shale formations for natural gas and oil in the United States. Because hydraulic fracturing involves injection of high volumes of water and chemicals deep underground at high (above lithostatic) pressures, it has raised questions about potential contamination of surface water and groundwater and human and environmental health and safety implications. Decision-maker priorities regarding shale gas development reflected concerns about human and ecological health as they relate to possible chemical contamination from shale gas development. Respondents cited the chemicals used in hydraulic fracturing and the chemicals produced during well operation as possible risks to human or ecosystem health, but they were unsure of whether water-quality data show these chemicals ending up in surface water or groundwater. Thus, the Blue Ribbon Project Steering Committee selected the following question around which to structure this case study:

"Do shale gas development activities contaminate surface water or groundwater?"

This question makes the connection between a congressional-level policy issue and a need for water data to inform that issue; it is of urgent importance to decision-makers in the Marcellus Shale region. The Susquehanna River Basin was selected as the geographic location for applying the case-study policy question to determine the availability of water-quality data. The number of unconventional wells in Pennsylvania grew from less than 200 in 2007 to greater than 9,300 as of August 2015 (Pennsylvania Department of Environmental Protection, 2015a). The rapid growth of high-volume hydraulic fracturing (HVHF) in the Susquehanna River Basin has raised concerns about the potential for degraded water quality and potential impacts on drinking water aquifers throughout the basin. A divisive debate has emerged between citizens who support shale gas development and the economic development it brings, and others who oppose the practice based on water-quality and other environmental concerns. This conflict remains unresolved among the public due in part to a lack of information regarding water-quality and other potential environmental impacts.

Understanding the water-quality impacts of shale gas development is a priority for the Northeast-Midwest region, but it is clear that understanding the potential impact of unconventional oil and gas development and hydraulic fracturing on water quality is an emerging national concern. The major conclusions of a recent draft report by the U.S. Environmental Protection Agency (USEPA) studying the impacts of hydraulic fracturing on drinking water resources nationwide generated significant interest and concern from both sides of the debate. The draft report did not find evidence of widespread systemic impacts on drinking water in the United States; however, the limited amount of data collected before and during hydraulic fracturing activities reduced the investigators' ability to determine whether hydraulic fracturing affected water quality in cases of alleged contamination (U.S. Environmental Protection Agency, 2015a).

Despite this high-profile interest, the question of whether shale gas development contaminates surface water or groundwater remains largely unresolved. Government agencies, academia, and volunteer groups have been collecting water-quality data for different parameters in different locations since shale gas development started in the Susquehanna River Basin in 2007, but it is not clear whether the systematic data collection effort necessary to generate answers to questions surrounding shale gas development and water quality has been made. In summary, this case study topic represents an urgent environmental and public health priority for the Northeast-Midwest region, a national energy policy priority, and an issue for which a comprehensive and analytical look at water-quality data needs and data availability has the potential to create a consensus regarding the data needed to inform future management actions and policy solutions.

The case study approach consisted of three main tasks:

1. Identify and describe the types and amounts of water-quality data needed to answer the case-study policy question;
2. Assess the extent to which those data are available and usable; and
3. Identify the additional data that would be needed to more effectively answer the case-study policy question and estimate the level of effort to collect those data.

As described throughout this report, there is a wide range of data types needed to answer this policy question. However, the scope of this study is focused only on the water-quality monitoring data that are needed. Additional data types are discussed but are not the focus of this study.

Shale gas development activities are connected to a range of potential contamination pathways that might allow a contaminant from a development activity to come in contact with surface water or groundwater; for example, methane gas can mobilize during well drilling and migrate along the well bore, or land clearing and construction can result in sediment contamination from runoff. The Technical Advisory Committee (TAC) identified the most likely potential water contamination pathways associated with shale gas development and considered the water-quality data that would be most important for identifying any impacts on surface water or groundwater for each pathway. Sections 2.3 and 2.4 provide background information regarding shale gas development activities and contamination pathways in the Susquehanna River Basin.

To identify the water data available to answer the case-study policy question, the USGS assembled a broad, multi-agency compilation of water data called the National Data Aggregation, with special focus on compiling a Susquehanna data set for the Susquehanna River Basin (Chapter 5). This Susquehanna data set served as the primary basis for assessing existing water-quality monitoring in the Susquehanna River Basin.

Conceptual models were developed to identify the types of data needed to answer the case-study policy question. A hypothesis was proposed for surface water to guide development of the monitoring design for answering the case-study policy question, and a statistical analysis was completed to illustrate the quantity of data needed to detect statistically significant changes in concentration for the most available water-quality parameters (Chapter 6). Surface-water monitoring sites from the Susquehanna data set were mapped and analyzed to identify the number of sites that meet the sampling criteria (Chapter 7). Data gaps were identified, monitoring recommendations were made for filling the gaps, and a range of costs for new monitoring were estimated (Chapter 8).

This process was repeated using a hypothesis for groundwater (Chapters 9 through 11). Chapter 12 summarizes recent developments in surface-water and groundwater-quality data related to shale gas development in the Susquehanna River Basin. Chapter 13 provides findings and recommendations regarding both surface water and groundwater monitoring in the Susquehanna River Basin.

2.3 Background

The Susquehanna River Basin was selected for this case study because it is undergoing significant shale gas development in the Marcellus Formation (hereafter referred to as the Marcellus Shale) (Figure 1), and it is rich in water-quality data collected by volunteer groups, State, interstate and Federal agencies. These factors make the basin a best case scenario for having water-quality data to answer the case-study policy question. Much of the Susquehanna River Basin also is underlain by the Utica Formation (hereafter referred to as the Utica Shale), which has been largely undeveloped, with only a few wells listing the target formation as the Utica Shale as compared to thousands of wells targeting the Marcellus Shale (Pennsylvania Department of Environmental Protection, 2015a; New York State Department of Environmental Conservation, 2015). The Marcellus Shale extends from southern New York across Pennsylvania, and into western Maryland, West Virginia, and eastern Ohio (Soeder and Kappel, 2009).

2.3.1 Description of the Susquehanna River Basin

The Susquehanna River Basin drains 27,510 square miles covering half the land area of Pennsylvania as well as portions of New York and Maryland (Susquehanna River Basin Commission, 2014a). The basin has 63 percent forest cover, a population of more than 4 million people, and more than 49,000 miles of rivers, streams, creeks, and brooks (Susquehanna River Basin Commission, 2014a). Figure 2 shows the land cover in the basin, with locations for wells in the Marcellus and Utica Shales in Pennsylvania that most likely require high-volume hydraulic fracturing (HVHF). In December of 2014, the State of New York decided that HVHF development should not proceed because the current scientific information is insufficient to determine the level of risk to public health and whether the risks can be adequately managed (New York State Department of Health, 2014).

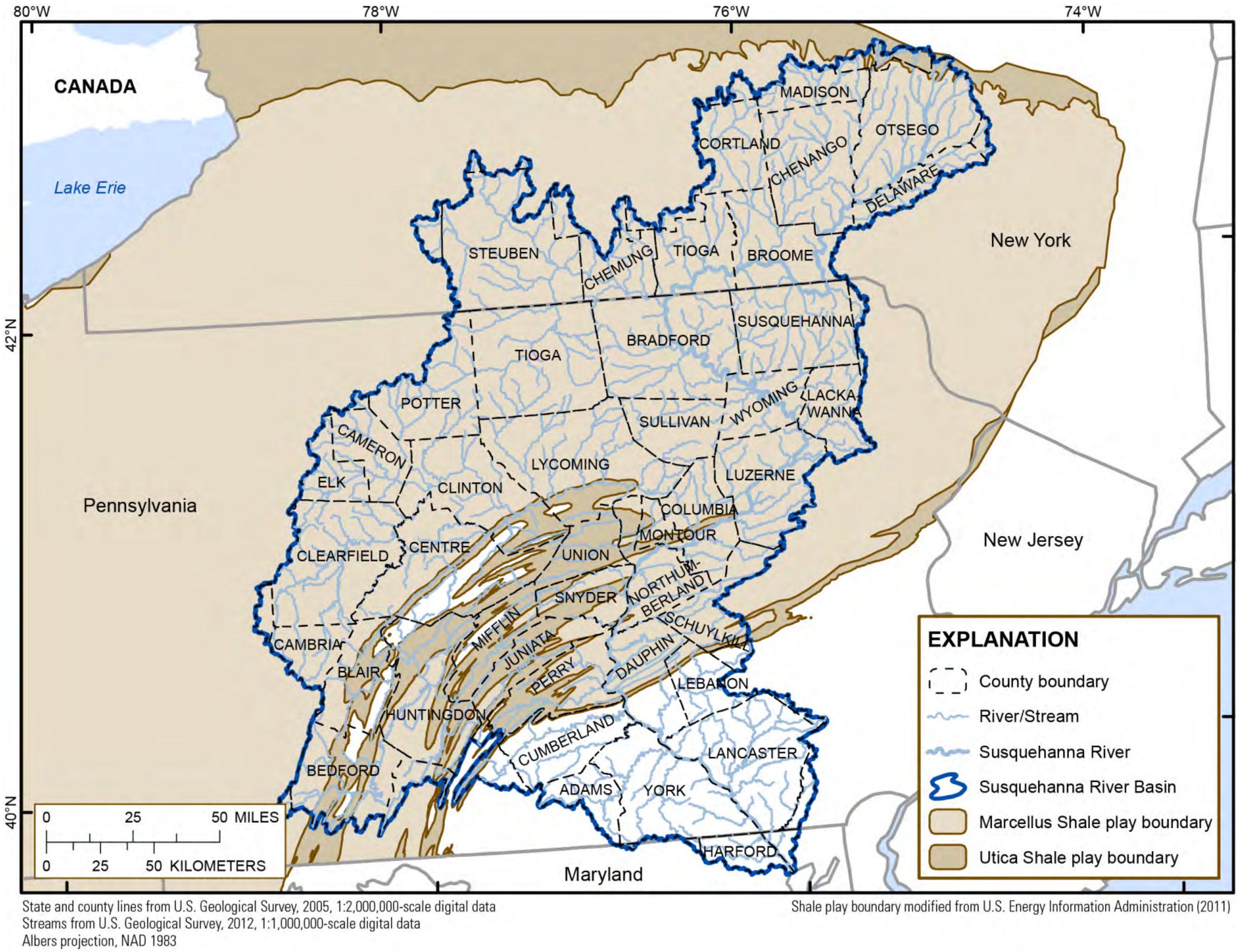
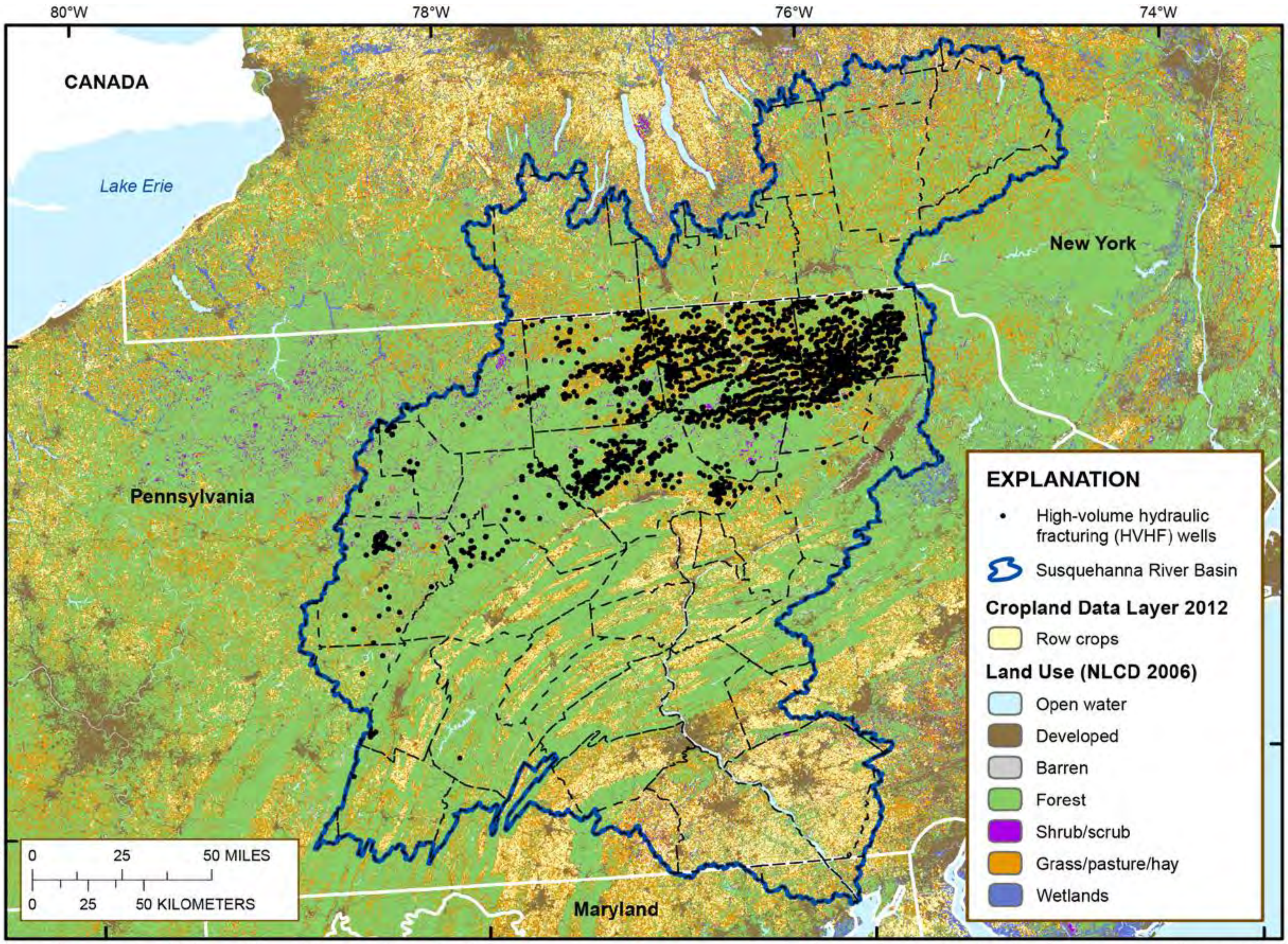


Figure 1. Susquehanna River Basin in relation to the extent of the Marcellus and Utica Shales.



State and county lines from U.S. Geological Survey, 2005, 1:2,000,000-scale digital data
Albers projection, NAD 1983

Wells from Pennsylvania Department of Environmental Protection (2015a)
Cropland Data layer from U.S. Department of Agriculture (2012)
NLCD 2006 from Fry and others (2011)

Figure 2. Land cover in the Susquehanna River Basin and the distribution of high-volume hydraulic fracturing wells.

2.3.2 Historic and recent trends in oil and gas development in the Susquehanna River Basin

Pennsylvania has a legacy of oil and gas production going back as far as the Drake Well in 1859 near Titusville; in the mid-1870s, the western regions of Pennsylvania were the center of crude oil production in the United States (Williamson et al., 1966). Hydraulic fracturing has been used in Pennsylvania since the 1950s, and prior to that, explosives were used to create artificial fractures for the purpose of increasing production, with the first commercial operation in 1865 (Madrigal, 2013). Since the 1980s, nearly all wells drilled in Pennsylvania have been hydraulically fractured (Pennsylvania Department of Environmental Protection, 2010).

Definitions

Shale Gas Development: The complete life cycle of gas development activities for hydrocarbon extraction from shale formations, beginning with clearing and construction and including drilling, hydraulic fracturing, disposal of waste fluids, well production and depletion, well-closure and site remediation, and long-term post-closure monitoring.

Conventional Well: Conventional oil and gas wells are typically vertical and target relatively permeable formations like sandstone or carbonate rocks. Conventional wells typically use low volumes of hydraulic fracturing fluids (50,000-100,000 gallons) and target discrete pay zones (the rock in which oil or gas are recoverable) over relatively short intervals (measured in tens of feet).

Unconventional Well: These wells target low permeability source-rock formations such as shale, coal seams, tight sands, and others. They can be vertical or horizontal wells. Horizontal unconventional wells typically require high volumes of hydraulic fracturing fluids.

Conventional oil and gas wells are typically vertical and target relatively permeable formations like sandstone or carbonate rocks, and unconventional wells target low permeability shale formations. Conventional wells typically use low volumes of hydraulic fracturing fluids (50,000-100,000 gallons) and extend relatively short distances in the target formation (measured in tens of feet). This report focuses on the potential adverse outcomes of hydraulic fracturing done at multiple stages along horizontal sections of unconventional wells that can extend several thousand feet into the target formation, and that use high volumes (2-10 million gallons) of hydraulic fracturing fluids. In this report, horizontal unconventional wells are presumed to require high-volume hydraulic fracturing and are referred to as “HVHF wells.” To simplify the discussion of HVHF wells throughout this report, the term ‘HVHF well’ refers to wells that have been drilled (see section 5.5). The term ‘shale gas development’ in this document refers to the complete life cycle of gas development activities for hydrocarbon extraction from shale formations, including high-volume hydraulic fracturing.

All of the HVHF development in the Susquehanna River Basin is currently occurring in Pennsylvania. In December of 2014, the State of New York declared a moratorium on HVHF development (New York State Department of Health, 2014), although conventional vertical wells have been and continue to be hydraulically fractured in New York (New York State Department of Environmental Conservation, 2011).

The number of unconventional wells in Pennsylvania has grown from 181 in 2007 to greater than 9,300 as of August, 2015 (Pennsylvania Department of Environmental Protection, 2015a). Not all wells targeting the Marcellus or Utica Shale are HVHF wells, as some test wells may be vertical and do not involve high volumes of hydraulic fracturing fluids. Unconventional wells drilled through 2007 were mostly vertical wells (87.8 percent), but since then only 9.2 percent have been vertical wells through August 2015.

Figure 3 shows the current extent of oil and gas wells in the Susquehanna River Basin classified by whether the target or producing formation is in the Marcellus or Utica Shale, and additional detail about well types is provided in Figure 4. Several well types in Figure 4 may need explanation: a dry hole is a well that did not encounter enough oil or gas to complete the well; disposal wells are used to inject wastewater or chemicals into deep underground formations; and injection wells are used to re-pressurize depleted oil reservoirs to enhance the recovery of the remaining oil at surrounding wells.

Figure 4 shows that there are 10,445 wells in the Susquehanna River Basin that are not HVHF gas wells. Many of these wells correspond to vertical wells found in the southwest corner of the Susquehanna River Basin. Approximately 99 percent of the 4,810 HVHF gas wells are in the Marcellus Shale. There are approximately 636 storage or monitoring wells related to the storage of natural gas in depleted reservoirs. Figure 3 and Figure 4 represent information obtained from online State well records that were queried 8/27/2015 (New York) and 8/25/2015 (Pennsylvania). The most recent spud date (the day when the process of drilling begins at a well) in these well records was 8/21/2015.

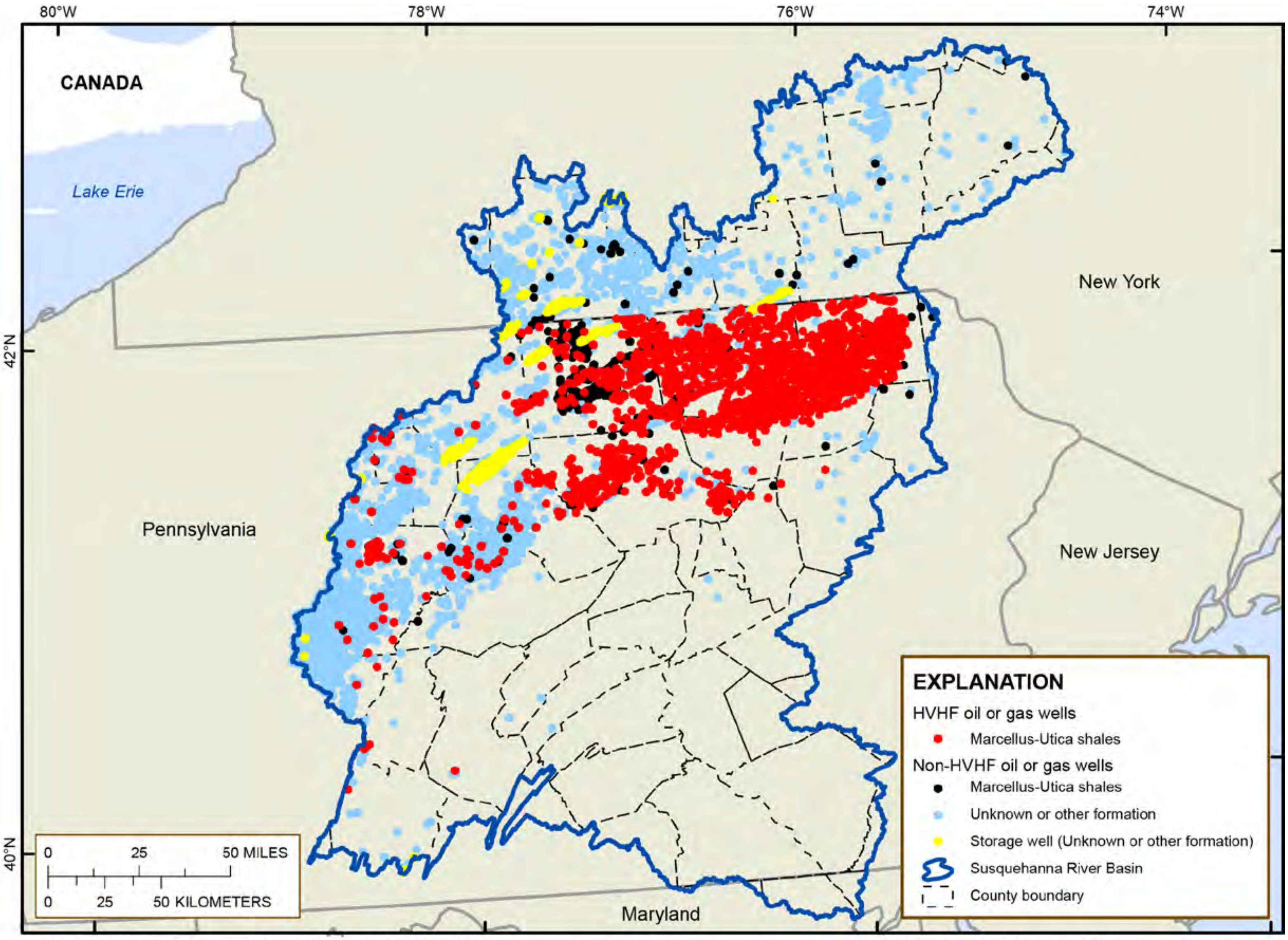
The introduction of shale gas development in Pennsylvania has contributed to its growth in natural gas production. In 2009, Pennsylvania ranked tenth in natural gas production, but in 2014, it was second to only Texas as the largest producers in the nation (U.S. Energy Information Administration, 2015). Of the 16 states that produce shale gas, Pennsylvania ranked second in shale gas production for 2014 (Figure 5). Annual shale gas production in Pennsylvania has steadily grown from 89 billion cubic feet in 2009 to approximately 4,000 billion cubic feet in 2014. Although production totals are much lower than in Pennsylvania, an upward trend in production from the Marcellus and Utica Shales can be seen in West Virginia and Ohio (Figure 5).

Definitions

Vertical Well: This is a well-drilling method used over the past 100 years. Single-stage hydraulic fracturing has been used to increase production in vertical conventional wells since the 1950's.

Hydraulic Fracturing Well: Any well that uses hydraulic fracturing technology to increase well production, including conventional vertical and unconventional horizontal wells.

High-Volume Hydraulic Fracturing Well: In this report, horizontal unconventional wells are presumed to require high-volume hydraulic fracturing and are referred to as "HVHF" wells. HVHF wells employ hydraulic fracturing at multiple stages along horizontal sections that can extend several thousand feet in the target formation. In the Susquehanna River Basin, HVHF wells typically use high volumes (2-10 millions of gallons) of hydraulic fracturing fluids in the Marcellus Shale.



State and county lines from U.S. Geological Survey, 2005, 1:2,000,000-scale digital data
Albers projection, NAD 1983

Wells from Pennsylvania Department of Environmental Protection (2015a)
and New York State Department of Environmental Conservation (2015)

Figure 3. Current extent (8/21/2015) of oil and gas wells in the Susquehanna River Basin classified by whether the target or producing formation is in the Marcellus or Utica Shale, is unknown, or is a storage well in another formation.
[Abbreviations: HVHF, high-volume hydraulic fracturing]

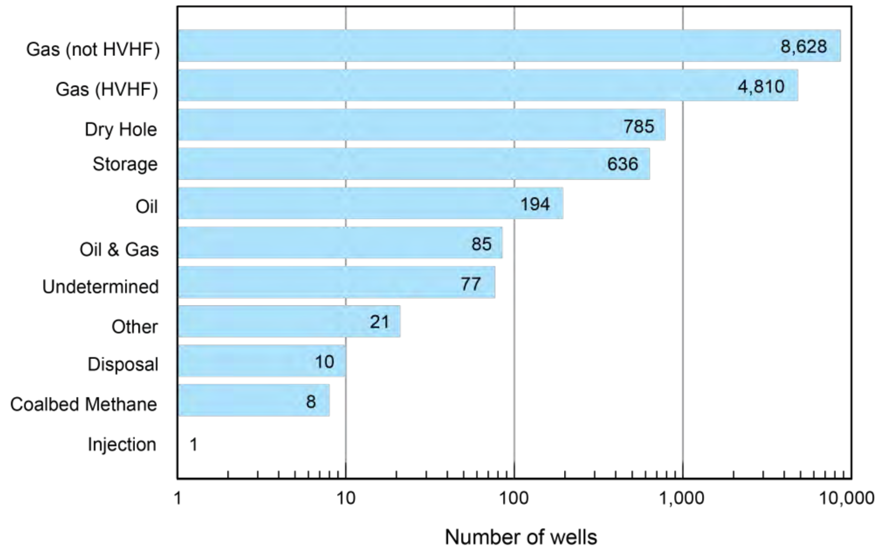


Figure 4. Types of wells related to oil and gas development in the Susquehanna River Basin (note that the x-axis uses a log scale; data from Pennsylvania Department of Environmental Protection (2015a) and New York State Department of Environmental Conservation (2015)).
[**Abbreviations**: HVHF, high-volume hydraulic fracturing]

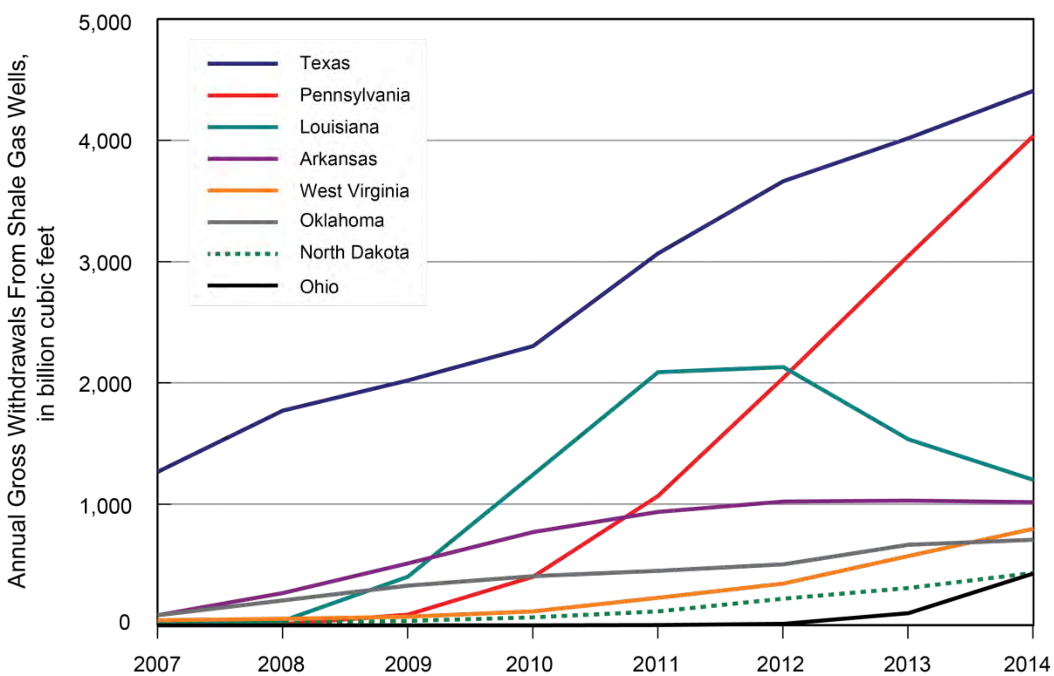


Figure 5. Annual gross withdrawals of natural gas from shale gas wells for the top 8 states that produce shale gas (data from U.S. Energy Information Administration, 2015).

2.3.3 Summary of previous recommendations regarding data and monitoring needs

Many previous and ongoing studies in the U.S. and in the Susquehanna River Basin have identified a need for data collection and analysis to determine whether shale gas development is affecting surface-water or groundwater resources.

The U.S. Congress requested the U.S. Environmental Protection Agency (USEPA) to conduct a study of the impacts of hydraulic fracturing on drinking water resources. Subsequently, the USEPA is undertaking a multi-year and national-scale effort to explore this issue. A progress report was released in December 2012 and an external review draft was released in June 2015 (U.S. Environmental Protection Agency, 2015a). The draft report did not find evidence of widespread systemic impacts on drinking water in the United States; however, the limited amount of data collected before and during hydraulic fracturing activities reduced the agency's ability to determine whether hydraulic fracturing affected water quality in cases of alleged contamination. The USEPA study includes a case study in the Susquehanna River Basin in Bradford and Susquehanna Counties. The draft report found that there is no systematic collection, reporting, or publishing of pre-drilling and post-fracturing groundwater monitoring data that could indicate whether hydraulic fracturing-related fluids have or have not migrated to shallow aquifers.

Battelle (2013) quantified the historical data available to characterize surface-water and groundwater quality for Bradford and Susquehanna Counties prior to 2007 to identify whether these data could be used to determine if shale gas production affects water quality. This time period is (mostly) prior to shale gas development in these counties. Battelle (2013) examined background levels for an expansive suite of parameters that might be indicative of impairments from shale gas activity and found that there are many instances of water-quality parameters exceeding screening criteria from such causes as legacy mining, road runoff, and agricultural land use, showing that impairments can be due to activities other than shale gas development. Battelle notes that overall, there is a lack of historical water-quality data in these counties. This lack of data, in combination with other causes of impairment, makes it difficult to determine whether recent hydraulic fracturing has impacted water quality (Battelle, 2013).

The Scientific and Technical Advisory Committee (STAC) for the Chesapeake Bay Program found that existing monitoring programs may not be adequate to assess shale gas development effects on surface waters. The STAC also found that few high-quality baseline water assessment studies have been undertaken prior to shale gas development. Participants in this workshop agreed upon the need for more baseline assessments and for long-term monitoring of shale gas development effects (Scientific and Technical Advisory Committee, 2013).

Bowen et al. (2015) analyzed nationally available surface-water chloride and specific conductivity data to identify whether national trends related to unconventional oil and gas development are apparent. They identified the current limitations of existing national water-quality monitoring programs for answering questions about shale gas development. Sufficient data for trend analysis were available for chloride and specific conductance in surface water for only 155 monitoring sites nationwide. Groundwater data were insufficient for conducting trend analyses. The authors recommend increased sample size and duration of sample collection and analysis for water-quality indicators related to shale gas development for regional assessments of potential effects of hydraulic fracturing.

The Health Effects Institute developed a strategic research agenda on the potential impacts of unconventional oil and gas development in the Appalachian region. This effort found limited availability of systematically collected surface-water and groundwater quality data before and after oil and gas development is initiated, or during the lifespan of wells (Health Effects Institute Special Scientific Committee on Unconventional Oil and Gas Development in the Appalachian Basin, 2015). Their strategic research agenda recommends research on surface-water and groundwater quality conditions before, during, and after oil and gas development; and identifying the sources of water contamination that are potentially related to oil and gas development.

2.4 Shale Gas Development Activities and Potential Contamination Pathways.

Hydraulic fracturing is just one step of shale gas development that could result in water contamination. The TAC considered the full life cycle of shale gas development activities, from land clearing and construction through well closure, and the potential contamination pathways that might allow a contaminant from the development activity to come in contact with surface water or groundwater. The TAC prioritized shale gas activities they determined to have the highest potential for water-quality impacts, and those in which the water-quality data are most relevant for identifying contamination events that are not otherwise easily observed, such as groundwater contamination or long-term migration of contaminants. The priority shale gas activities identified by the TAC include:

- land clearing and construction activities,
- well drilling,
- hydraulic fracturing and the well completion process,
- well production, and
- cumulative impacts.

Water-quality concerns associated with shale gas development include:

- the potential for methane contamination in groundwater,
- the effects of land clearing and construction on surface water,
- the potential contamination of surface water or groundwater from surface spills,
- subsurface migration of injected chemicals and underground contaminants such as brine and naturally occurring radioactive material, and
- the handling and disposal of waste (solids and fluids) created during these processes.

Detailed background on shale gas development is readily available through websites hosted by the U.S. Energy Information Administration, the Penn State Marcellus Center for Outreach and Research, FracFocus, and many others. The National Geographic has an interactive website showing a typical HVHF well in the Marcellus Shale and the hydraulic fracturing process (National Geographic, 2014), and Penn State Public Broadcasting has created a similar website (Penn State Public Broadcasting, 2011). The following sections provide an overview of the priority shale gas development activities, associated potential contamination pathways, and potential contaminants that may be found in surface water and groundwater. This section is provided as background; readers may skip to Chapter 3 to focus on the details of this case study.

2.4.1 Land clearing and construction activities

Clearing and construction involves clearing and grading of land, road building, pipeline and pad building and any other forms of land disturbance in preparing for shale gas development. The typical well pad in the Marcellus Shale comprises an area about 5 acres in size (2.5 football fields). As technology has improved, laterals (horizontal sections of the well) in the Marcellus Shale can now reach up to 10,000 feet. Longer laterals extract gas from a larger drainage area, and the drainage area per well pad in the Marcellus Shale has grown to about 640 acres (1 square mile) or greater with some companies drilling 6 or more wells per pad. The horizontal drilling of multiple wells from a well pad results in a smaller environmental “footprint” (area disturbed) per well as compared to conventional gas development (U.S. Department of Energy, 2013), although HVHF development in the Marcellus Shale involves extensive earth disturbance including roads, well pads and pipelines (Pennsylvania Department of Environmental Protection, 2012a). Slonecker et al. (2014) document the landscape consequences of shale gas development in several counties in Pennsylvania.

Land clearing and road, pipeline and well-pad development can cause sediment to enter waterways through short-term disturbances related to construction, and can create longer term problems due to an increase in impervious surfaces. Impervious surfaces change the hydrodynamic energy balance in a watershed, creating a mechanism for increased runoff and sediment transport to streams. This sediment can harm the natural environment, particularly in pristine areas prevalent in the Susquehanna River Basin such as cold-water fisheries. Studies have analyzed the relationship between watershed imperviousness and biotic integrity in streams and identified negative effects starting at 10 to 15 percent imperviousness (Schueler, 1994; Brabec et al., 2002), although more recent research suggests that effects begin to occur at the earliest stages of urban development (Coles et al., 2012).

In Pennsylvania, approximately 60 percent of existing and permitted pads occur on slopes at risk to some excess surface-water movement and local erosion (Drohan and Brittingham, 2012). Development in pristine watersheds is a primary concern; in Pennsylvania, the ‘sweet spot’ of shale gas development in Susquehanna and Bradford Counties overlies some of the most “exceptional quality” watersheds in the State (Brantley, 2013). Clearing and construction is common to all types of development. However, the scale of clearing and construction for shale gas development in these pristine, unique watersheds may result in degraded water quality; nearly 25 percent of Pennsylvania HVHF well pads are being built in core forest areas (Fisher, 2012).

Oil and gas operators must develop and implement a written Erosion and Sediment Control Plan when:

- planned earth-disturbance activities will result in a total earth disturbance of 5,000 square feet or more,

- the earth disturbance activity has the potential to discharge to water classified as High Quality or Exceptional Value water, or
- if 5 acres or more of earth will be disturbed over the life of the project (Pennsylvania Department of Environmental Protection, 2012a).

In December 2012, Pennsylvania adopted a policy whereby agency staff interpret “project,” for purposes of whether a permit is required, to mean all substantially connected well sites, access roads, pipelines, other service lines, support facilities, and/or other oil and gas activities (Pennsylvania Department of Environmental Protection, 2012a). Multiple well pads in a concentrated area raise concerns regarding the cumulative impact of clearing and construction. Many well pads are less than 5 acres, making many well pads drilled before December 2012 exempt from the State's erosion and sediment control permit requirements (Johnson, 2010).

Changes in macroinvertebrate communities in streams and turbidity can help characterize the effect of well pad clearing and associated construction (roads, gathering pipelines, compressor sites) due to shale gas development.

2.4.2 Well drilling

After construction of the well pad, a small drilling rig will move in and drill the vertical section of the well known as the “tophole,” and set and grout the well casings. The well “drilling” activity is defined by the presence of a drilling rig. After each vertical well is drilled and grouted, a larger rig is used to drill the transition (curved part of the well from vertical to horizontal) and the horizontal part of the well (Pennsylvania State University, 2014), though sometimes the large rig may do all the vertical and horizontal drilling. The laterals extend horizontally for a distance from 3,000 to 10,000 feet into the shale gas formation (U.S. Department of Energy, 2013). The “total time to drill each well is about three to six weeks depending on the depth and length of the horizontal well, so if there are four wells on a well pad, you could expect the big rig to be there for about three to six months”(Pennsylvania State University, 2014).

HVHF wells are lined with telescoping layers of steel casing and cement grout to prevent transport of the fluids and hydrocarbons inside the production casing to the surrounding soil, bedrock, and groundwater aquifers, and to prevent transport between different layers of the surrounding soil, bedrock, and groundwater aquifers. Typically, thousands of feet separate the shallow fresh water aquifers and the deep area where hydraulic fracturing occurs. Improper well construction and cementing of the well annulus could allow gas, brine, or hydraulic fracturing fluids or their residual fluids to migrate up through gaps in the cement grout or along the improperly sealed well annulus into overlying freshwater aquifers. Methane occurs naturally in Devonian formations which lie above the Marcellus Shale, underscoring the importance of proper well construction.

2.4.2.1 Solid waste

A 12-inch diameter horizontal borehole through 5,000 feet of Marcellus Shale would result in approximately 3,900 cubic feet of waste mud and rock, weighing approximately 5.3 tons (Soeder, 2011).

Black shales, such as the Marcellus Shale, contain naturally-occurring heavy metals and other elements that can be detrimental to the environment if mobilized and concentrated. Most of this solid waste is disposed in licensed landfills as long as radioactivity limits at a given landfill are not exceeded.

2.4.2.2 Spills

Spills can occur during any stage of the shale gas development process. When spills occur on land, surface water is affected to different degrees depending on what is spilled (fuel, cuttings, fluids used in hydraulic fracturing, wastewater, etc.), how much is spilled, and how much reaches the waterway. A review of shale gas-related spills on land in Pennsylvania shows that most spills were small and did not have environmental impacts as they were contained within the boundaries of the pad site (Considine et al., 2012). For those spills directly into surface waters, *"… in most cases, these events are minor. Our analysis tracks all types of spills from a gallon of diesel fuel to hundreds of barrels spilled into the many small creeks and ponds in rural Pennsylvania…The impacts of these events varied by the amount of fluids spilled. Our analysis indicates that on average 105 gallons of fluid were spilled for minor water contamination events"* (Considine et al., 2012). Spills can infiltrate into shallow groundwater. When spills occur, groundwater is affected to different degrees depending on what is spilled, how large the spill is, and how much is cleaned up before it drains to the aquifer (Drollette et al., 2015). The effects of spills are currently monitored, assessed, and documented by the Pennsylvania Department of Environmental Protection (PADEP), and information on documented spills is publicly available through the offices of PADEP. Of the 3,533 wells drilled between 2008 and August 2011, 149 had minor land spill violations (defined as less than 400 gallons) and 9 had major land spill violations (defined as greater than 400 gallons) (Considine et al., 2012); there were 8 major surface-water contamination violations and 258 minor surface-water contamination spills as determined by the authors' evaluation of environmental impact (Considine et al., 2012).

A national review of spills related to hydraulic fracturing by the USEPA (U.S. Environmental Protection Agency, 2015b) found that spills related to hydraulic fracturing were frequently low volume events of up to 1,000 gallons, with relatively few high volume events over 20,000 gallons. Of the spills that the USEPA authors determined were associated with hydraulic fracturing, half reached surface-water or groundwater resources.

2.4.2.3 Mechanical Disturbances, Turbidity, and Metals

Other forms of possible contamination from well drilling may be indirect. Mechanical disturbances, such as vibrations from drilling activity, could loosen iron oxide (rust) particles from the casings of domestic wells, leading to increased turbidity of well water. Arsenic and selenium that occur naturally in the drinking water could be mobilized into groundwater if iron oxide coatings with adsorbed arsenic or selenium are mobilized. Mechanical disturbance could also mobilize strontium and barium from sulfate or carbonate scales on the interior casings of water wells (Fontenot et al., 2013). Lowered water levels can cause changes in pH, redox, or other chemical properties that in turn could cause desorption of metals such as arsenic and selenium from soil and unconsolidated material (Fontenot et al., 2013), which could enter groundwater when water levels return to previous levels.

Conventional oil and gas wells have been drilled for over 100 years; the most commonly validated impact from conventional oil and gas drilling activity on private water supplies is a short-term turbidity problem, which generally resolves quickly (New York State Department of Environmental Conservation, 2011) and can occur whenever the aquifer is penetrated by a drilling operation.

2.4.3 Hydraulic fracturing and well completion

Activities after the drill rig leaves but prior to production are called "well completion," and include a number of closely coordinated steps including hydraulic fracturing (also known as 'fracking' or 'fracing'). The peak in on-site fluid storage and transport occurs during well completion when fracturing fluids are transported by truck or pipeline and pumped into a well, increasing the possibility of spills. Four to five million gallons of freshwater are used to hydraulically fracture each HVHF well in Pennsylvania (Susquehanna River Basin Commission, 2013).

After the drilling rig is removed, a crane is used to lower a casing perforation tool into the well bore to perforate the production casing and grout seal around the casing in preparation for hydraulic fracturing. Hydraulic fracturing is the injection of a highly pressurized mixture of water, chemicals and sand (or other proppant) into sections of the well (each section or stage is fractured in 300- to 500-foot sections using multiple stages spaced along the lateral section reaching a total length of up to 10,000 feet). In each stage, hydraulic fracturing fluid is injected under pressure through the perforations to fracture the reservoir rock, which allows natural gas and other hydrocarbons to flow along the fractures to the well more easily. Sand is typically used to 'prop' open the fracture when the fluid pressure is released to prevent the newly-opened fractures from closing. Hydraulic fracturing uses heavy-duty pump trucks, material blenders, and wellhead equipment to mix and inject the fracturing fluid into the well, through the perforations, and into the shale formation. The hydraulic fracturing step typically occurs within a few weeks or months of the well drilling, dependent on the project schedule. The staged process of hydraulic fracturing requires several days for each well that is hydraulically fractured (Pennsylvania State University, 2014). After hydraulic fracturing, the well is allowed to 'unload' spent hydraulic fracturing fluids, proppant, and drill cuttings, and the well may be tested to measure the 'shut-in pressure build-up' for estimating future production. The gas produced during unloading may either be processed and fed into a gathering pipeline or discharged through a controlled flaring process prior to well production, although the USEPA has banned flaring subsequent to January 2015. On average, 10 percent of the injected fluids return to the surface via the well over several days or weeks during this period (Vidic et al., 2013).

2.4.3.1 Chemicals Used in Hydraulic Fracturing

A Congressional report identifies approximately 750 chemicals and components that have been used in hydraulic fracturing fluids, including 29 components that may be hazardous if introduced into the water supply (U.S. House of Representatives, 2011). Some of the chemicals used in hydraulic fracturing are carcinogenic (U.S. House of Representatives, 2011) and others are known or suspected endocrine disrupting chemicals (Kassotis et al., 2013).

Industry practice is evolving and several companies are beginning to disclose the chemicals they use in hydraulic fracturing. The website *fracfocus.org* provides a chemical disclosure registry for industry: "The changes in state public disclosure laws are occurring so fast that posting a comprehensive list of all states contemplating or preparing laws in this area is not possible as it would change on a frequent basis." (FracFocus, 2013).

The fracturing fluids typically consist of more than 99 percent water and sand, with additional additives used to address a wide range of issues that are site specific, so there is no single formula for the volumes or types of each additive (FracFocus, 2013). Water use for hydraulic fracturing in the Marcellus and Utica Shale area ranges from 10,001 to 36,620 cubic meters per well (Gallegos et al., 2015a). Typically, 3 to 12 chemicals are added (FracFocus, 2013); the chemicals act as friction reducers, corrosion inhibitors, scale inhibitors, biocides, acids, and surfactants. Friction reducers enable smooth (laminar) flow of the fluid into the rock formation at a higher rate and reduced surface pressure than if water alone were used. Corrosion inhibitors prevent corrosion of metal pipes. Scale inhibitors are added to the fracturing fluid to limit the precipitation of salts and metals in the shale formation and inside the well; scaling can limit the productivity of a well. Biocides prevent microorganism growth, and reduce fouling of the fractures. Acids are also used to remove drilling mud near the wellbore area after perforation and to dissolve soluble minerals in the surrounding formation. Surfactants (alcohols such as methanol or isopropanol) may also be added to reduce the fluid surface tension to aid fluid recovery (Vidic et al., 2013). Chemical additives may represent a small fraction of the total volumes involved, but given the large volume of fluid used for hydraulic fracturing these small fractions can add up to significant volumes. For example, one 4-million gallon fracturing job in the Marcellus Shale used 937 gallons of hydrochloric acid and 29 gallons of methanol, despite both chemicals representing less than 0.01 percent of the total fluid by weight (Cooley and Donnelly, 2012). Gallegos et al. (2015b) published a report summarizing publicly available data regarding hydraulic fracturing treatments, water volumes, proppants, treatment fluids, and additives used in hydraulic fracturing in the United States from 1947 through 2010.

Due to the large number of chemicals used in hydraulic fracturing, not all can be monitored. Benzene, toluene, ethylbenzene, and xylene (BTEX compounds) were identified as chemicals of concern in a Congressional report on hydraulic fracturing fluid chemicals, and are commonly used in hydraulic fracturing (U.S. House of Representatives, 2011.) The BTEX chemicals have been used to assess water quality in domestic wells near shale gas development (Fontenot et al., 2013) and are among the best options for analyzing groundwater for the presence of chemicals used in hydraulic fracturing fluids.

2.4.3.2 Produced Water

In this report, "produced water" is defined as any water that returns to the surface during shale gas development. This broad definition of produced water includes water that returns to the surface after hydraulic fracturing is completed and before well production commences—which some have defined as "flowback" water. In the Marcellus Shale, an average of 10 percent of the water injected for hydraulic fracturing will come out of the wells (Vidic et al., 2013). This produced water is a mixture of injected fracturing fluids and water from the target or surrounding formations that may contain dissolved salts and metals, naturally occurring radionuclides, dissolved organic carbon, and natural gases including methane,

ethane, and propane. These parameters can be measured to determine whether produced water is contaminating surface water or groundwater. Each well in the Marcellus Shale generates as much as 200 tons of salt in the first few months of production; these salts are found in highly saline produced waters (Vidic et al., 2013). The produced water increases in salinity and decreases in volume over time, and at the end of the 90-day period the fluid has a salinity of about 6 to 10 times that of seawater (Hayes, 2009). The produced water from the Marcellus Shale is the second saltiest and most radiogenic of all waters produced from sedimentary basins in the United States where shale gas development is occurring (Vidic et al., 2013).

The radioactive materials and dissolved salts and metals in produced water can have potential ecological and human health impacts if released to surface water or groundwater, and all of these contaminants have potential environmental impacts. Some researchers have postulated that brine originating from the Marcellus Shale can move through natural features in areas yet to be developed for shale gas, further complicating efforts to tie brine contamination to recent shale gas development (Warner et al., 2012).

The safe storage and disposal of produced water associated with shale gas development is a major challenge. The average quantity of shale gas-related produced water generated in Pennsylvania during 2008 to 2011 was 26 million barrels per year, representing a fourfold increase compared with the period preceding shale gas development (Vidic et al., 2013). The handling of produced water from HVHF wells is evolving. In the early years of the shale boom, produced water was sent to municipal wastewater treatment plants in Pennsylvania. The municipal wastewater treatment plants could not treat the radionuclides, total dissolved solids (TDS), or dissolved metals in the wastewater, which led to public health issues such as bromate in drinking water supplies (Brantley, 2013). Practices have evolved, and produced water is no longer sent to municipal wastewater treatment plants due to environmental concerns about TDS entering streams (Soeder et al., 2014). Now, about 90 percent of produced water from HVHF wells in the Marcellus Shale is treated and re-used (Groom, 2013). When it can no longer be recycled, produced water is injected into a deep underground disposal well or treated at an industrial wastewater treatment plant (Soeder et al., 2014).

Produced water can be sent to private industrial wastewater treatment plants and then discharged to streams and rivers if discharge standards are met. These licensed facilities are primarily located in the Ohio River Basin, but several are located in the Susquehanna River Basin. Warner et al. (2013a) describe in detail the treatment process of one such treatment facility in western Pennsylvania that exclusively treats the produced brine from oil and gas wastewater. Several contaminants in the halogen family such as bromide, sodium and chloride are conservative, that is, treatment plants do not remove these contaminants from the waste stream. Others contaminants, such as radium and barium, are substantially reduced but not completely removed. The researchers found high levels of radioactivity in sediments near the treatment plant outflow and above levels that would prohibit their disposal at a licensed treatment facility (Spotts, 2013); such levels would pose environmental risks of radium bioaccumulation in localized areas (Warner et al. 2013b). However, ascertaining that shale gas development is the source of the contamination is complicated by the fact that discharge into that stream had long been permitted for conventional oil and gas brines, and that all of these brines are known to contain radioactive radium, which can sorb to bottom sediments (Brantley et al., 2013).

New PADEP standards dictate that any oil and gas produced water must be treated to have a TDS concentration of no greater than 500 parts per million prior to discharge, which is consistent with Pennsylvania drinking water standards (Abdalla and Drohan, 2009). There are new treatment plants being built to meet these standards and at least one is now in operation in Williamsport, Pennsylvania (Pennsylvania State University, 2014), and construction has begun on a facility in Standing Stone, Pennsylvania.

2.4.3.3 "Frack Hit"

A "frack hit" or "downhole communication" occurs when fractures from a HVHF well intersect with an abandoned well, or an existing well that is producing gas. An existing well that is already producing gas is not equipped to handle high pressures associated with active hydraulic fracturing; the pressures of approximately 10,000 pounds per square inch can push oil, gas, and hydraulic fracturing fluids and water up the borehole of an abandoned or a producing well, causing a blowout or overflowing the holding tank that collects brine. Several such incidents have occurred in the West (Vaidyanathan, 2013). When a blowout or uncontrolled venting occurs, the potential exists for large amounts of fluids and gases to be released from the wells, despite mitigation efforts by operators. Once a frack hit occurs, negative environmental impacts are almost impossible to avoid (Considine et al., 2012). While this pathway is more germane to contamination of surface water, it is possible that spills from well blowouts could infiltrate to groundwater. Individual states are responsible for documenting spills and frack hits; there is no comprehensive federal database describing known well blowouts.

2.4.3.4 Pressure changes and Methane Mobilization

Another potential source of water-quality degradation near HVHF wells can be methane gas and sediment moving through fractured aquifer systems. Cloudy water, discoloration, bad taste, and methane gas may all be due to materials pre-existing in the fractures that become entrained and transported by a surge in the fractures from the high-pressure air used during drilling (Geng et al., 2014). The methane is naturally occurring, but mobilized by shale development activities (Vidic et al., 2013). This can be an issue particularly in northeast Pennsylvania, where approximately 80 percent of domestic wells have detectable levels of methane due to natural conditions (Warner et al, 2013a; Molofsky et al, 2013). These pressure gradients may also allow brine to move through faults, fractures, or abandoned wells; and cementing and casing deficiencies may allow brine to reach freshwater aquifers.

Methane will tend to move toward areas of low pressure, which can occur when water is withdrawn from a domestic well as the water level in the well is lowered. The lower water level (hydraulic head) may allow for methane to exsolve from the groundwater and accumulate as gas in the well casing, or dissolved methane can remain in the well water to be pumped into the domestic water supply system. In areas of natural methane occurrence such as northeastern Pennsylvania, when a well is pumped (the head is decreased), a spike in methane gas levels in the well annulus will be observed due to the pressure change caused by pumping (B. Smith, EnviroClean Products And Services, oral commun., July 2014).

Methane is not toxic but represents a flammability and explosion risk, and can be associated with lowered property values and impairment of domestic wells for drinking water use (Osborn et al., 2011). Methane that enters groundwater as a solute is not considered a health hazard with respect to ingestion and is therefore not regulated in the United States (Vidic et al., 2013). While methane is tasteless and odorless, it can cause taste and odor problems: methane bubbling into domestic wells can suspend fine-grained sediments that accumulate at the bottom of wells, causing normally clear water to suddenly become "colored, turbid, slimy, and smelly" (Gorody, 2012). Low oxygen levels can cause increased solubility of such elements as iron and arsenic and reduce sulfate to sulfide, causing taste and odor problems (Vidic et al., 2013).

Methane occurs naturally in many parts of the Susquehanna River Basin in groundwater, and consists of two types. Microbial methane is formed in relatively shallow aquifers from the decomposition of organic material, as can occur in swamps or in landfills. Thermogenic methane is formed also from organic sources but at high temperatures and pressures, deep in the earth. Methane can migrate from coal seams, shale or other source rock formations, or glacial sediments as drift gas (Vidic et al., 2013). Methane can also derive from anthropogenic sources such as gas storage fields, coal mines, landfills, gas pipelines, and abandoned gas wells.

2.4.3.5 Induced Fractures

A common question is whether the hydraulic fracturing process can create pathways by which gas, brine, or fluids used in fracturing could rapidly migrate from the Marcellus Shale into the shallow subsurface. While risk assessments indicate that the probability of this happening is low (Soeder et al., 2014), there are broad public concerns about the safety of this practice.

Several scientists have investigated the hydrogeology of the Marcellus Shale and illustrate many obstacles to contaminants moving vertically in the aquifer network (Carter et al., 2013, Engelder, 2012; Flewelling et al, 2013; Flewelling and Sharma, 2014, Saiers and Barth, 2012). These investigations are detailed below.

Obstacles include the depth of the Marcellus Shale formation with many layers of different rock types between the Marcellus Shale and shallow aquifers. These different rock types have heterogenous fracture patterns, none of which are demonstrated to penetrate the Upper Devonian stratigraphic units (Carter et al., 2013).

Unidirectional flow up from the Marcellus Shale towards the shallow aquifers is not expected. Horizontal flow typically occurs in high conductivity aquifers and vertical flow in layers of low conductivity. The bedrock above the Marcellus Shale is stratified, consisting largely of interbedded shale, mudstone, and siltstone units with low conductivity, and lesser proportions of limestone and thin sandstone units of higher conductivity (Harper, 1999 and U.S. Geological Survey, 2006 as cited in Saiers and Barth, 2012).

Fluid movement in the Marcellus Shale is further limited by physics. The nature of fluid transport in non-saturated conditions, as encountered in the Marcellus Shale and some overlying rock formations, would severely constrain fluid transport (Saiers and Barth, 2012). The capillary tension in the Marcellus Shale is likely to prevent migration of liquids (Engelder, 2012; Soeder, 1988). Recent research suggests that capillary and osmotic forces move fracturing fluid into the shale rather than out of it (Engelder et al., 2014). This may, in part, explain the relatively small volume of produced water when compared to the quantity injected during hydraulic fracturing.

Fractures from hydraulic fracturing are relatively small. After examining over 12,000 hydraulic fracturing operations across North America, Flewelling et al. (2013) concluded that direct hydraulic connection between tight formations like shale and shallow groundwater via induced fractures and existing bedrock structures (joints, fractures, or faults) is not a realistic expectation based on the limitations on fracture height growth. The results of this analysis showed that the maximum height of microseismic events was less than 600 meters above well perforations and that induced fractures were comparatively small (radii on the order of 10 meters or less).

Flewelling and Sharma (2014) have modeled the theoretical travel time of a contaminant in the Marcellus Shale to shallow aquifers. Their model shows that even in conservative scenarios "permeability is low, upward flow rates are low, and mean travel times are long (often greater than 10^6 years). Consequently, the recently proposed rapid upward migration of brine and [hydraulic fracturing] fluid, predicted to occur as a result of increased [hydraulic fracturing] activity, does not appear to be physically plausible."

In the short term, hydraulic connections between the Marcellus Shale and shallow aquifers is not likely given the long distances and confining layers that separate the shale from shallow groundwater aquifers in the Marcellus Shale; however, the shorter the distance between fracturing activity and overlying aquifer, the more likely that contamination could occur (Vidic et al, 2013).

As one drills deeper into the bedrock, groundwater typically becomes more salinized (more salty), and eventually becomes a brine, which can be many times more salty than seawater. There is a research need to understand whether natural brine transport pathways could present a risk for contamination of aquifers in hydraulic fracturing areas (Vidic et al., 2013; Warner et al., 2013a).

2.4.3.6 Cement Grout Integrity

The high pressures introduced into a well during the hydraulic fracturing process place a great deal of stress on the casing and cement grout. The weak points tend to be the boundary surfaces—interfaces between the cement grout and casing, or the cement grout and the formation. When the shear strength placed on one of these interfaces during hydraulic fracturing is greater than the shear strength at the interface itself, the cement grout may shrink, swell, or crack and detach from the formation or steel casing. Fractures created this way can extend significant distances along the interface, grow and link up, or enter the formation itself, resulting in the development of a long crack known as a microannulus.

The most common cause for methane contamination in domestic wells (termed 'stray gas') is a buildup of gas pressure in and around casing cements, which causes methane gas to migrate along the well annulus to local aquifers (Gorody, 2012). In Pennsylvania, methane gas is extracted from the Marcellus Shale, but it is also found in the overlying Middle and Upper Devonian formations. If a HVHF well is not properly grouted, it is possible that methane could migrate from any of the gas-bearing formations that the borehole intersects, along the well annulus or through the grout prior to it hardening, and reach overlying drinking-water aquifers. Problems with well grouting have been cited by researchers as occurring in approximately 1 to 2 percent of wells (Considine et al., 2012). Other researchers investigating well construction problems in Pennsylvania found problems in 3.4 percent of wells (Vidic et al., 2013). Vidic et al. (2013) provide a detailed discussion of potential geological issues that can cause well grouting problems, and discuss best practices for mitigating these problems.

2.4.4 Well production

Well production is defined as the step when a HVHF well begins producing fluids and gas. This happens after the fracturing fluid has been allowed to flow back, and the surface infrastructure is complete. The surface infrastructure includes valves, piping, separators, storage tanks, a gas compressor and other equipment.

Migration of brines, fracturing fluids and gas could occur along microannuli in improperly sealed wells during well production. Abandoned wells could also provide a migration corridor (Harrison, 1985).

2.4.5 Cumulative impact

The cumulative effects of shale gas development include all the individual shale gas development activities taken together, including the cumulative impact of long-term and low-level upstream spills, clearing of land, and other ongoing contamination events that might not be detectable by themselves but would manifest as changes in water quality that might increase over time. The individual activities of shale gas development taken together, even if they do not have significant water-quality impacts on their own, may result in cumulative impacts on surface-water or groundwater quality. In addition to the cumulative impact of multiple processes, there is the cumulative impact of multiple well pads within a given area that also must be considered.

Cumulative effects will likely be manifested through changes in surface water or groundwater for most water-quality parameters, except sediments due to erosion. Contaminated aquifers that discharge to streams can result in long-term contamination of surface water (Winter et al., 1998). Long-term monitoring of surface water can indicate problems with groundwater contamination, particularly during periods of low-flow since groundwater contributes to streams in most physiographic and climatic settings (Winter et al., 1998). Given the potential for episodic releases of contaminants during the individual processes of shale gas development (including accidental release of hydraulic fracturing chemicals and produced waters or brines), regular discrete and continuous water-quality monitoring of streams and rivers in watersheds undergoing shale gas development activities is .

Table 1 summarizes these shale gas development activities and the contamination pathways.

Table 1. Summary of priority shale gas development activities and potential contamination pathways identified by the Technical Advisory Committee.

Shale gas development activity	Surface or groundwater impact?	Contamination pathways with the highest potential for water-quality impacts
Land clearing and construction	Surface Water	• Sediment loadings to streams may increase during clearing and construction activities. • Erosion may increase due to long-term changes in flow patterns.
Well drilling	Groundwater	• Drilling can open surficial pathways that allow preexisting native gas (methane) to leak into water wells.
Hydraulic fracturing and well completion	Groundwater	• Spills of chemicals onsite can infiltrate groundwater. • Fracturing may push fracturing fluids, proppants, or brines through faults or fractures to the fresh water aquifer. • Fracturing may mobilize methane into the freshwater aquifer. • Unidentified abandoned wells or “frack hits” may allow contaminants to reach freshwater aquifers.
Well production	Groundwater	• Pressure gradients may allow contaminants to move through faults or fractures. • Cementing and casing deficiencies may allow contaminants to reach freshwater aquifers. • Unidentified abandoned wells may allow contaminants to reach freshwater aquifers. • Long-term migration of contaminants could occur.
Cumulative impacts	Both	• The accumulated effects of shale gas development activities over space and time for all groundwater pathways, including waste disposal and well-closure.

3. Most Critical Data Types Needed to Answer the Case-Study Policy Question

The first case study task was to describe, and quantify where possible, the types and amounts of water-quality data that are needed to answer **"Do shale gas development activities contaminate surface water or groundwater?"** in the Susquehanna River Basin**.**

Conceptual models were developed to delineate the multiple types of data necessary for answering the case-study policy question, both water-quality and other types of data. These conceptual models for surface water (Figure 6 and Figure 7) and groundwater (Figure 8) connect shale gas development activities with potential contamination pathways and corresponding water-quality parameters, which in turn are connected to endpoints of concern to decision makers. The conceptual models make clear that many data types are needed to measure contamination of surface water or groundwater related to shale gas development. The TAC further refined the scope of this case study to more clearly identify and prioritize the types of data needed to answer the policy question. This section describes the most important data types selected by the TAC for measuring whether shale gas development contaminates surface water or groundwater. The specific data types that were evaluated in this case study directly or were examined as supporting data are shown in Table 2.

3.1 Study Design Needed to Answer the Case-Study Policy Question

The conceptual models (Figure 6, Figure 7, and Figure 8) can be distilled into three critical components that must be available to answer the case-study policy question, as shown in Figure 9. The right water data must be available in the right locations with the right supporting information. The selection of appropriate monitoring sites is critical for answering the case-study policy question. For surface water, monitoring sites must be located in watersheds with HVHF wells and in reference watersheds in each ecoregion. Monitoring sites in these types of watersheds allow for the detection of water-quality changes in watersheds with HVHF well development and comparison with undeveloped watersheds to identify whether water-quality changes are resulting from HVHF development. The study design for groundwater again requires monitoring sites in the right places: networks of groundwater sampling sites are needed with each sampling site located within 1 mile of an HVHF well.

Table 2. Water-quality and ancillary data types necessary to answer the case-study policy question, identifying the data types evaluated in this report.

Data type	Study status
Water-quality data	
Groundwater quality data	Quantified
Surface-water quality data	Quantified
Ancillary data	
Oil and gas well locations and abandoned wells	Supporting analysis
Plugged, abandoned and orphaned wells	Supporting analysis
Landfills and wastewater treatment plants that accept shale gas waste	Supporting analysis
Land Use	Supporting analysis
Aquifers	Supporting analysis
Ecoregions	Supporting analysis
Other contamination sources	Not evaluated
Streamflow	Evaluated
Weather variables	Not evaluated
Endpoints of concern	Not evaluated

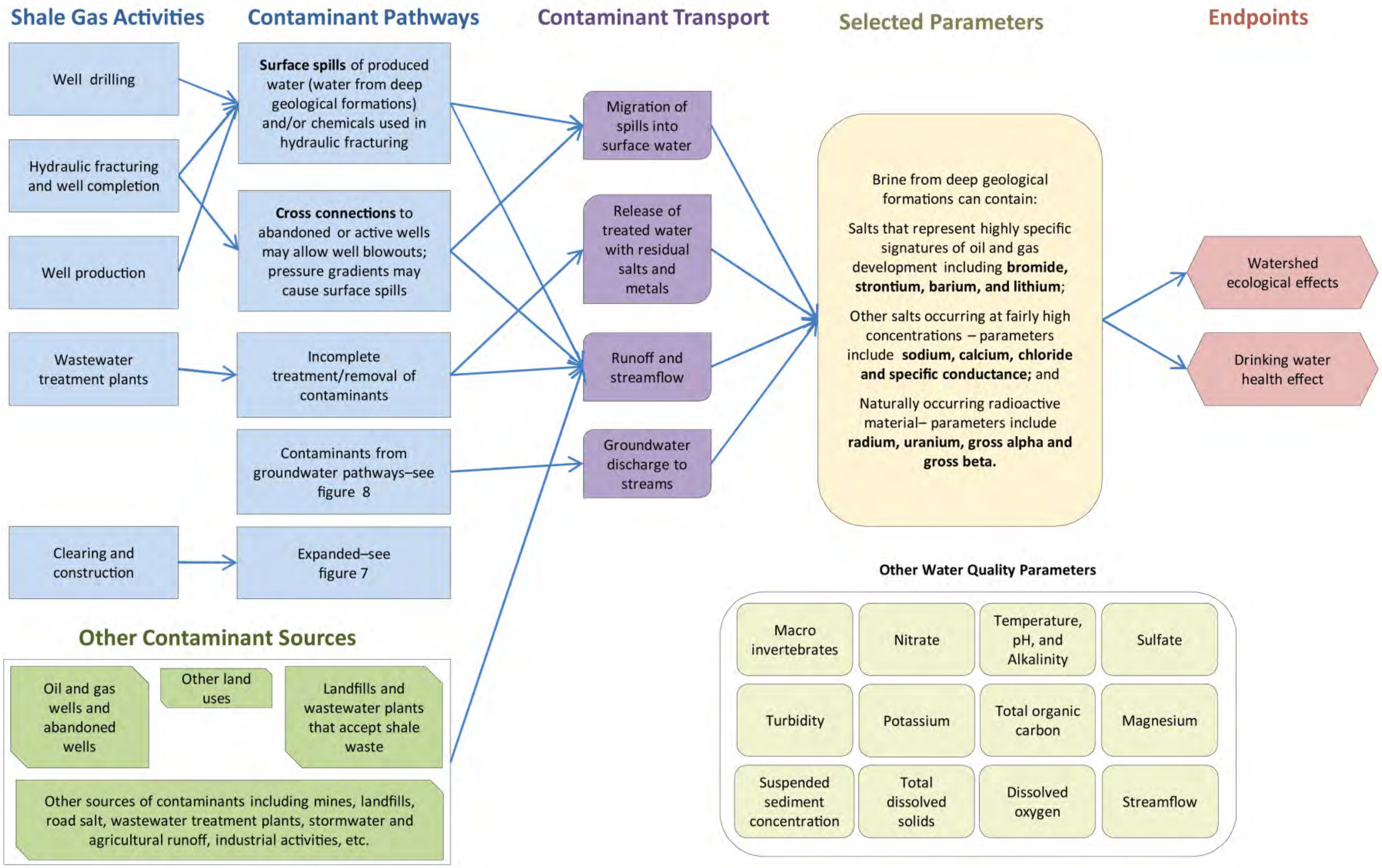


Figure 6. Conceptual model that identifies water-quality and ancillary data needed to understand the impacts of shale gas development on surface water.

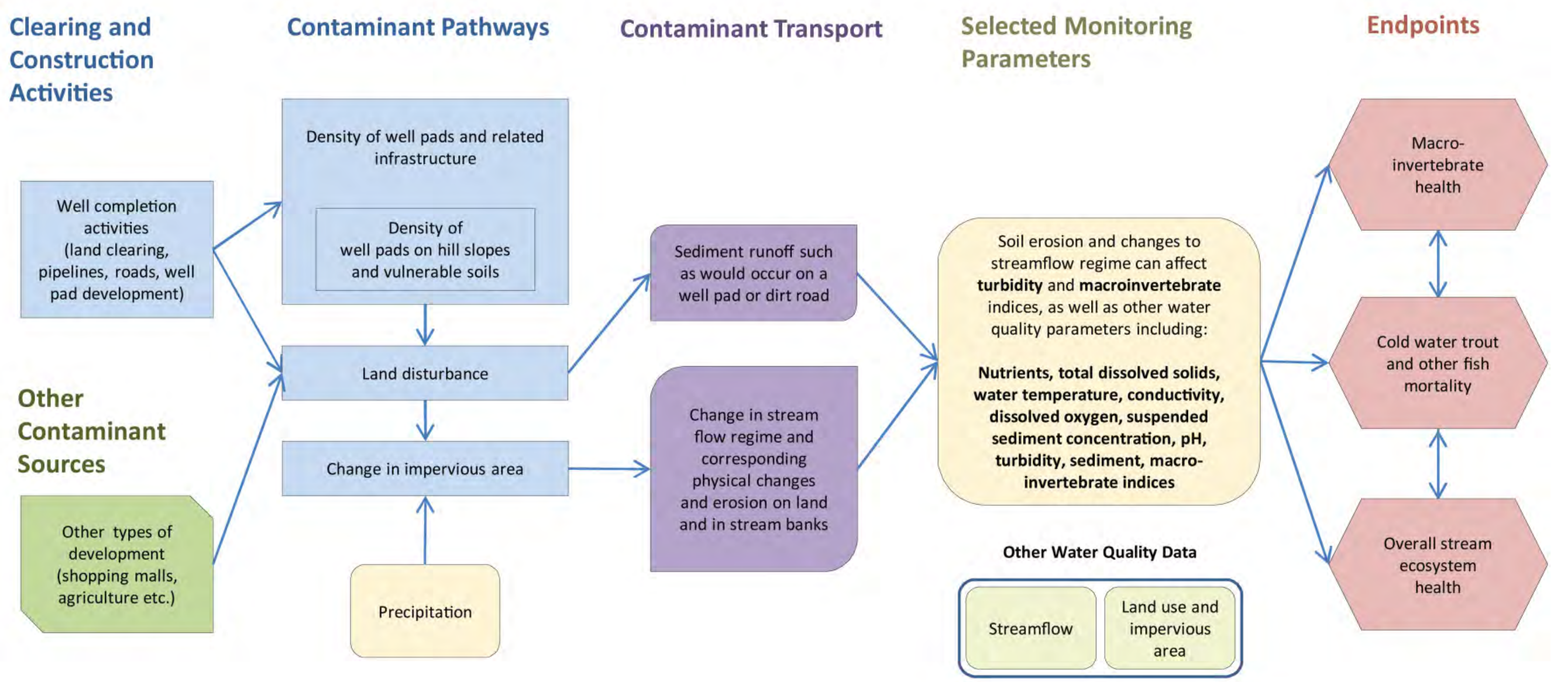


Figure 7. Conceptual model that identifies water-quality and ancillary data needed to understand the impacts of land clearing and construction due to shale gas development on surface water.

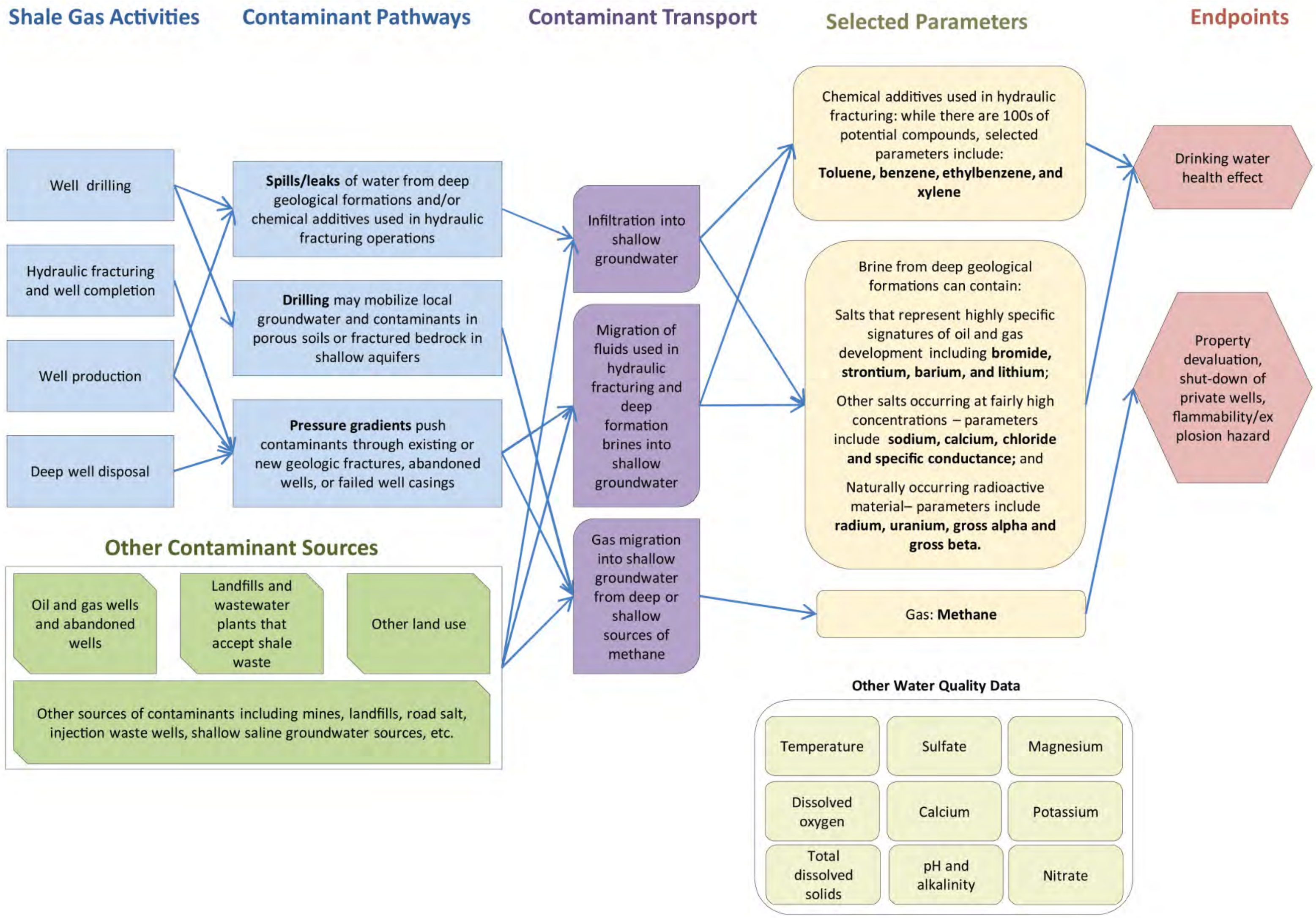


Figure 8. Conceptual model that identifies water-quality and ancillary data needed to understand the impacts of shale gas development on groundwater.

Next, water-quality and streamflow data at the appropriate monitoring sites must be available with sufficient sampling frequency and duration to evaluate trends in concentration over time. Finally, data on shale gas development, geology, climate, and other changes in land use in the monitored area must be available to correlate water-quality change with shale gas development activity. Without this information, the relationship between shale gas development and water quality cannot be evaluated, even if shale gas development is causing water-quality change.

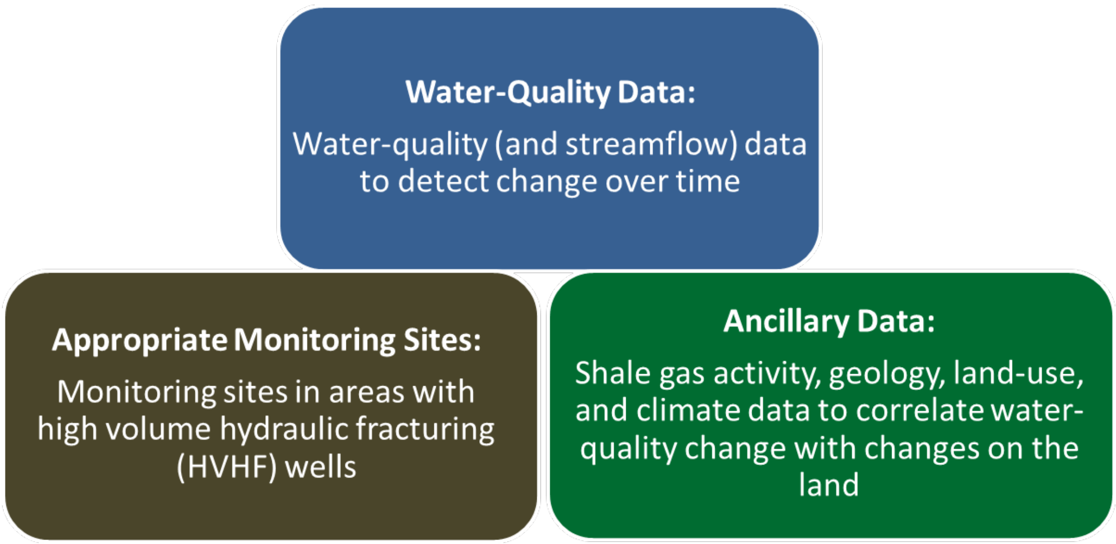


Figure 9. Study design needed to answer "Do shale gas development activities contaminate surface water or groundwater?"

3.2 Shale Gas Activities and Contaminant Pathways

The shale gas activities and contaminant pathways shown in the conceptual models (Figure 6, Figure 7, and Figure 8) were discussed in detail in section 2.4.

3.3 Contaminant Transport

In surface water, contaminant transport from shale gas development activities includes migration of spills into surface water, and release of treated water that contains residual salts and metals. Streamflow transports contaminants downstream and affects parameter concentrations in surface water. Streamflow information is needed to identify changes in a parameter over time due to differences in streamflow as opposed to other causes. Flow and stage data are regularly collected by the USGS and State monitoring agencies. USGS streamgages are often used to measure flow at sites where another agency collects water-quality data. These data collected by USGS are usually stored in the National Water Information System (NWIS) database. Streamflow data for locations within the Susquehanna River Basin were quantified and compiled for this project.

Groundwater discharge from contaminated aquifers to streams is another potential source of contaminant transport. Gradual changes and cumulative effects in groundwater will likely manifest in surface water over longer periods of time. If indeed there is a long-term change in surface-water quality it is likely related to a change in groundwater quality, a cumulative, chronic change versus an acute, spike effect.

For clearing and construction activities, sediment runoff during precipitation events and erosion during high flow are additional forms of contaminant transport that can result from shale gas development.

For groundwater, infiltration from surface spills and gas migration from geologic sources of methane to shallow groundwater are the primary forms of contaminant transport. Migration of hydraulic fracturing fluids and migration of produced water to shallow ground water are other possible means of transport to freshwater aquifers, as discussed in section 2.4.3.

3.4 Water-Quality Parameters

The TAC identified a set of priority parameters that is most useful for determining whether shale gas development is changing surface-water or groundwater quality. Table 3 identifies these priority parameters and the role of each parameter in identifying whether shale gas development might have been the source of contamination. To develop this suite of parameters, the TAC chose parameters that were: 1) possible indicators of shale gas development, 2) important for characterizing and interpreting water-quality data, and 3) most readily available in the existing record. The parameters summarized in Table 3 are referred to throughout this document as the “suite of priority surface-water parameters” and “suite of priority groundwater parameters.”

The priority parameters include:

- possible indicators of oil or gas field brines,
- parameters present in brine from deep geologic formations,
- indicators of landscape disturbance,
- parameters that characterize source water type,
- parameters present in chemical additives for hydraulic fracturing,
- parameters that are mobilized by pressure gradients and can serve as indicators of faulty well casings,
- redox indicators, and
- non-specific indicators of contamination.

Table 3. Suite of priority monitoring parameters selected by the Technical Advisory Committee to answer the question "Do shale gas development activities contaminate surface water or groundwater?" in the Susquehanna River Basin.

Parameter	Surface water laboratory measurement	Groundwater laboratory measurement	Surface water field measurement	Groundwater field measurement	Reason for Selection[1]
Barium, dissolved	x	x			A
Barium, total	x				A
Bromide	x	x			A
Calcium	x	x			B, C
Chloride	x	x			B, C
Gross alpha	x	x			A
Gross beta	x	x			A
Lithium	x	x			A
Magnesium	x	x			C
Nitrate	x	x			B, F
Phosphorus, total	x				B
Potassium	x	x			C
Radium-226, -228	x[2]	x			A
Sodium	x	x			C, D
Strontium	x	x			A
Sulfate	x	x			C, F
Suspended sediment concentration	x				B
Total dissolved solids	x	x			A, B
Total organic carbon	x				B, G
Uranium	x	x			A
Benzene		x			D
Toluene		x			D
Ethylbenzene		x			D
Xylene		x			D
Methane		x			E, F
Alkalinity			x	x	B, C, G
Dissolved oxygen			x	x	C,G,F
pH			x	x	C, G
Specific conductance			x	x	A, B, C, G
Turbidity			x	x	B, G
Water temperature			x	x	B, G
Streamflow			x		H

[1]A: Possible indicator of oil or gas field brine based on known or reported presence in oil or gas field brines, produced waters, or flow-back waters from hydraulically fractured wells; B: Indicator of landscape disturbance; C: Characterizes water type; D: Potentially present in chemical additives used for hydraulic fracturing; E: Mobilized by pressure gradients and indicator of faulty well casing; F: Redox indicator; G: Low cost non-specific indicator of potential contamination; H: Needed to determine influence of flow on parameter variability.

[2]Testing for radium-226, and -228 in surface water is recommended only if there is a change in gross alpha and gross beta; if measurements of gross alpha and gross beta indicate possible radiological contamination, the TAC recommends enhanced monitoring for radium-226 and -228 as a diagnostic tool.

The TAC selected metals and salts that are concentrated in brines and produced waters as possible indicators of oil or gas field brines for both surface water and groundwater. Calcium, chloride, and sodium occur at fairly high concentrations in produced water, and laboratory analysis for these parameters is inexpensive. Dissolved and total barium, bromide, strontium, and lithium are parameters that can represent highly specific signatures of oil and gas development to determine whether contamination has resulted from release of brines to the environment from produced water. While bromide, strontium and barium may be considered highly specific signatures of produced waters (Vidic et al., 2013), these elements have been found in groundwater samples in northeastern Pennsylvania that pre-date shale gas development (Molofsky et al., 2011). Barium is also commonly found in groundwater samples from wells that penetrate restricted groundwater zones in northeastern Pennsylvania (Williams et al., 1998; restricted groundwater zones are defined as being present in bedrock and in some cases in till and in confined aquifers that directly overlie bedrock.) The TAC also chose a general water-quality parameter, specific conductance, as a non-specific indicator of dissolved metals and salts.

Naturally occurring radioactive material (NORM) also can be present in brines. Radioactivity is the release of energy in the form of gamma rays and energetic particles (alpha and beta particles); uranium, gross alpha and gross beta are included in the suite of parameters. Decay of uranium can be a considerable source of gross alpha in shallow oxic aquifers, but uranium is unlikely to be found in high concentrations in Marcellus Shale brine because the brine is anoxic. For shallow aquifers, it will be important to document the presence of uranium and its contribution to ambient radioactivity in groundwater.

If measurements of gross alpha and gross beta indicate possible radioactive contamination, the TAC recommends additional analysis for radium-226 (Ra-226) and radium-228 (Ra-228) as a diagnostic tool. Analyses of Ra-226 and Ra-228 and gross alpha/gross beta are relatively expensive compared to other parameters, so analysis for gross alpha/gross beta is recommended as an initial screen using standardized analytical methods approved by the USEPA.

Total organic carbon was selected as an indicator of fluids used in hydraulic fracturing for surface water, and changes in turbidity can be used to indicate impacts of landscape disturbance on streams. Biological data, such as assessments of macroinvertebrate communities, are general indicators of stream health and can be used to indicate potential acute or chronic upstream disturbances**.** In Pennsylvania, macroinvertebrate monitoring is used to determine if streams are meeting designated uses under the Clean Water Act. (A. Gavin, Susquehanna River Basin Commission, oral commun., July 2013).

Additional organics are appropriate for sampling in groundwater; however, due to their volatility they are not as informative in surface water. Methane was selected by the TAC as an important parameter for monitoring in groundwater because it can be mobilized by various shale gas development activities. The TAC determined that BTEX compounds are the best parameters to determine if fluids used in hydraulic fracturing have migrated into drinking water aquifers, although the TAC noted that deep-water chemistry can modify BTEX compounds due to mixing with deep brines at high temperature and pressure. Some drawbacks to monitoring for BTEX compounds are that they are not unique to shale gas development, they are expensive to analyze, and can be prone to contamination during sampling. With energy companies potentially switching to alternative fracturing-fluid formulations, BTEX compounds may be less common in the future. Another drawback is that data on BTEX compounds are not readily available in the existing record, based on the Susquehanna data set and the data compilation efforts of Battelle (2013), but BTEX data were collected by Chesapeake Energy prior to shale gas development (Siegel et al., 2015b).

Bromide and chloride are useful indicators because once dissolved, they tend to be conservative in groundwater (Davis et al., 1998) and can be used to characterize water type. The ratio of chloride to bromide, plotted against concentrations of either bromide or chloride, can be used to identify the sources of chloride in groundwater and in reconstructing the origin and movement of groundwater (Whittemore, 1988; Davis et al., 1998; Mullaney et al., 2009; Chaplin, 2012). Chloride, barium, and TDS can be useful indicators in groundwater because of the high concentration of these parameters in produced water relative to typical background concentrations in Pennsylvania groundwater (Boyer et al. 2012).

Other water-quality data are indicators of stream health and can be used to indicate potential acute or chronic upstream disturbances. While not directly related to shale gas development, some water-quality data and field measurements are critical for understanding and interpreting water-quality data. For example, some contaminants are associated with certain water types as characterized by their major-ion chemistry. Major ions include calcium, chloride, potassium, magnesium, sodium, sulfate, and bicarbonate as alkalinity. Field parameters include dissolved oxygen, pH, specific conductance, turbidity, water temperature, and streamflow.

As surface-water and groundwater quality are researched and better understood, and shale gas development technology advances, the priority parameters may change over time. For example, iodine may be a promising groundwater tracer (Lu et al., 2015), and distinctive lead isotope ratios may be useful for identifying contamination sources (Johnson and Graney, 2015).

3.5 Other Contaminant Sources

Attributing a contaminant to a specific activity like shale gas development is complicated by the fact that many of the contaminants associated with shale gas development can occur naturally, or they are associated with other human activities such as conventional oil and gas development, various industrial activities, septic or animal wastes, road deicing, or other processes. Identifying a source of pollution therefore relies on the availability of additional supporting data including the condition of water quality prior to shale gas development, locations of conventional and HVHF wells and other sources of contaminants, and understanding the local geology and geochemistry. Ancillary data, as presented in Figure 9, are critical for identifying potential sources of contamination and to differentiate them from shale gas development sources. While this study focuses on the availability of water-quality data, the following section describes the other data types that are needed to help answer the policy question.

Potential contaminants from shale gas development that are similar to those associated with conventional oil and gas development include: methane gas and other hydrocarbons, brines (salts), metals (barium, strontium, etc.), radioactive parameters (primarily radium-226, radium-228), synthetic chemicals, fuels, and drilling mud and cuttings. Fluids used in hydraulic fracturing of conventional wells are an additional potential contaminant. Bromide, strontium and barium may be present in oil-field brines associated with historic or conventional oil and gas development, and in some areas may leak into surface water and groundwater from improperly sealed and/or abandoned oil and gas wells. Water produced from HVHF wells resembles brines produced from conventional gas wells that tap into other Paleozoic formations in the region that range in age from Ordovician to Devonian (Haluszczak et al., 2013). When evaluating temporal changes in these parameters near HVHF wells an inventory of existing or abandoned conventional and HVHF wells or former well fields is needed to determine if any other sources of deep brine exist in the area being evaluated.

Figure 10 shows locations of non-HVHF wells, including known plugged, abandoned and orphaned oil and gas wells. Non-HVHF oil and gas wells include storage wells, which are also shown in Figure 10. Not shown in any of the figures are 996 wells (all types) without coordinates that are estimated to be within the Susquehanna River Basin based on the location information from the well records (county, township or town). The locations of older wells installed decades ago are not well documented and the location of many plugged, abandoned and orphaned oil and gas wells is unknown. These are the wells that are most likely to have not been sealed properly. According to Pennsylvania Department of Environmental Protection (2013):

> *"Since the first commercial oil well was drilled in Pennsylvania in 1859, it is estimated that 300,000 oil and gas wells have been drilled in the state. Only since 1956 has Pennsylvania been permitting new drilling operations, and not until 1985 were oil and gas operators required to register old wells. In the prior years of the oil and gas industry in Pennsylvania, many wells were not properly plugged when abandoned. An unplugged abandoned well can be a hazard to the health and safety of people and cause pollution to the environment. Air, water and soil contamination can be attributed to leaking wells."*

Wastewater treatment plants that accept fluid waste from HVHF wells may themselves be a source of priority parameters in surface water or groundwater. Figure 11 shows waste management facilities in the Susquehanna River Basin that accepted waste from unconventional wells within the basin during the period 2011 to June, 2015 (Pennsylvania Department of Environmental Protection, 2015b).

Surface-water and shallow groundwater quality are often affected by prior and existing land uses; each land use is associated with its own suite of potential contaminants. Many of these contaminants include the priority parameters associated with shale gas development. In order to understand the likely origin of a particular parameter, the land uses in the watershed must be well understood. The main causes of documented water-quality impairments in Bradford and Susquehanna Counties are associated with agriculture and road runoff (Battelle, 2013), some of which have been well characterized and others not. For example, Johnson et al. (2015) identified and quantified regional and road salt sources along the New York/Pennsylvania border. Other sources of similar contaminants discussed by Battelle (2013) are summarized below:

"Studies have shown elevated levels of chloride, dissolved organic carbon, and sulfate concentrations in monitoring and domestic wells in proximity to septic tanks (Katz et al., 2011). Non-point sources of stormwater runoff and industrial activities are associated with metals, salts, pH, siltation, and suspended solids. Habitat modification can cause stream bank instability and significant soil erosion and sediment pollution in nearby streams. Non-coal mineral mining is associated with total suspended solids, turbidity, temperature changes, pH, and oil and grease from runoff and washing operations in the vicinity of mechanical equipment and vehicles. Coal mining is associated with acid mine drainage and contamination parameters including: metals, sulfate, and general water quality (for example, total dissolved solids, pH)."

Information was obtained for the Susquehanna River Basin to characterize monitoring sites, discussed in sections 5.4 and 5.5, but detailed study of potential sources of contaminants associated with various land uses upstream (or upgradient) of monitoring locations was beyond the scope of this study.

3.6 Endpoints of Concern to Policy Makers

The potential impact of shale gas development activities on surface water and groundwater is of interest to policy makers due to the potential for ecological and human health effects. Depending on the concentrations of contaminants from shale gas development, contaminants in drinking water sources may exceed maximum contaminant limits and affect human health. There are examples of elevated bromide in sources of drinking water resulting in increased disinfection byproducts in finished drinking water (Brantley, 2013). These contaminants can also result in ecological effects, and sensitive macroinvertebrate communities can be affected even at low concentrations. Increased erosion and turbidity can impact cold water trout populations that are popular tourism attractions in the region.

The potential for contaminants in groundwater can also result in adverse drinking water health effects. Private homeowner wells are not subject to the same sampling requirements as surface-water utilities under the Safe Drinking Water Act. As a result, unknown drinking water quality at private homes and contaminant concentrations in domestic wells may exceed maximum contaminant levels that would have been monitored in a community water system. Methane presents another endpoint of concern because methane in groundwater at high concentrations presents an explosion hazard.

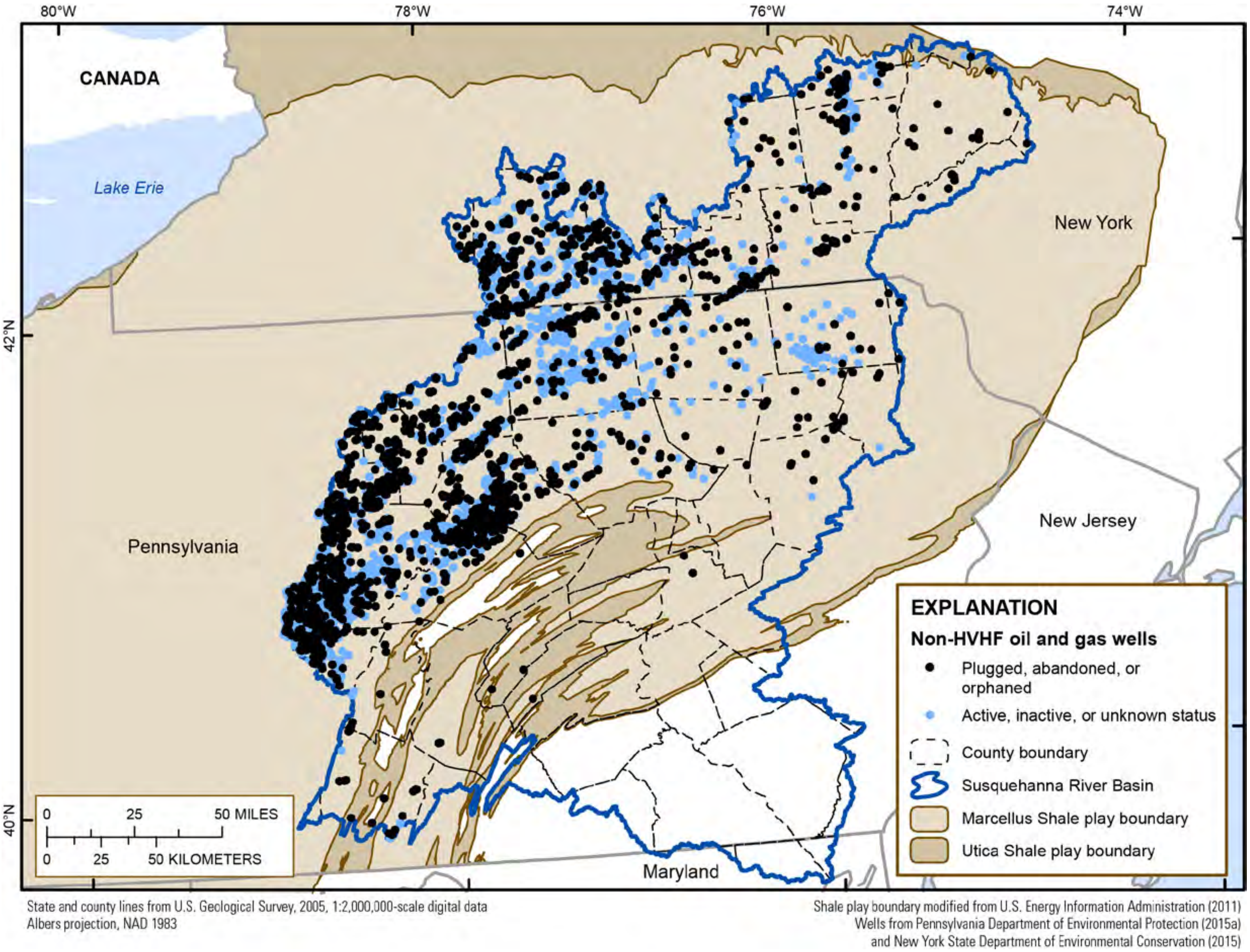
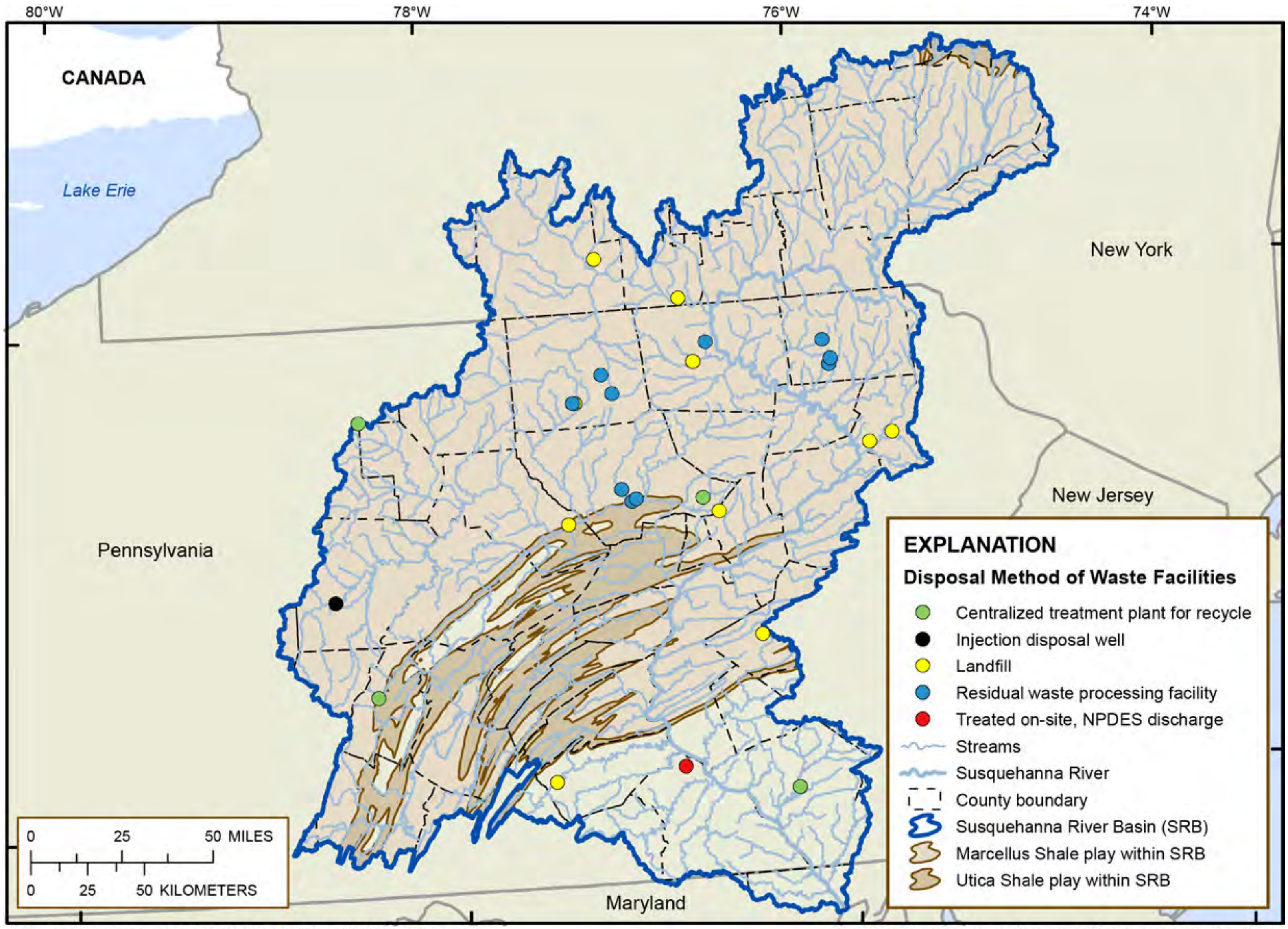


Figure 10. Non-HVHF oil and gas wells and known plugged, abandoned, and orphaned wells in the Susquehanna River Basin.
[Abbreviations: HVHF, high-volume hydraulic fracturing]



State and county lines from U.S. Geological Survey, 2005, 1:2,000,000-scale digital data
Streams from U.S. Geological Survey, 2012, 1:1,000,000-scale digital data
Albers projection, NAD 1983

Facilities data from Pennsylvania Department of Environmental Protection (2015b)
Shale play boundary modified from U.S. Energy Information Administration (2011)

Figure 11. Facilities accepting waste generated from Pennsylvania shale gas development in the Susquehanna River Basin (New York and Pennsylvania) between 2011 and 2015.
[Abbreviations: NPDES, National Pollutant Discharge Elimination System]

4. Criteria Used to Define the Scope of the Case Study Evaluation

The conceptual models make clear that many data types are needed to identify whether shale gas development is changing surface-water or groundwater quality. The TAC selected additional criteria to further define the scope of the case study to more clearly identify and prioritize the types and quantities of data needed to answer the policy question.

4.1 Focus on Cumulative Effects of Shale Gas Development.

The TAC narrowed the scope of the case study to focus on the cumulative effects of shale gas development, which were described in section 2.4.5, and the water data needed to know whether cumulative effects of shale gas development are contaminating surface water or groundwater. Cumulative effects of shale gas development, as defined in this document, include the long-term impacts of the individual activities of shale gas development and the associated potential contamination pathways taken together, and the cumulative impacts of multiple well pads within a geographic area. The TAC decided to focus on the ability to detect the cumulative effects of shale gas development because water-quality monitoring is critical to identifying this type of contamination. Short-term spills of concentrated brines are likely to be ephemeral, with spikes in concentrations. Acute incidents are best detected through on-site monitoring, accident reporting with targeted response monitoring, and through continuous water-quality monitors in receiving streams. Cumulative effects of shale gas development were selected as the study focus because these are more subtle to detect, a more targeted monitoring approach is required to detect changes, and water-quality monitoring is the only path to identifying low level and long-term changes.

The conceptual models (Figure 6, Figure 7, and Figure 8) illustrate the shale gas development activities that comprise cumulative effects and their potential for changing surface-water and groundwater quality in the Susquehanna River Basin. In the absence of water-quality data, the long-term cumulative effects of shale gas development on water quality will be unknown. If contamination occurs undetected in the headwaters of the Susquehanna River Basin, degradation of downstream water bodies could occur over the long term throughout the basin and in the Chesapeake Bay.

4.2 Suite of Priority Parameters and Streamflow

The TAC agreed that no one parameter can identify whether shale gas development is the source of contamination if a change in that parameter concentration is detected. By monitoring a suite of parameters, more information is available to identify the source of the contamination. The collective judgment of the TAC was that there is no ‘magic bullet’ for monitoring the effects of shale gas development on water quality. Each shale gas development activity, from initial well pad development to production of gas from a completed well is associated with different pathways that contribute different potential contaminants, so monitoring for just one of the parameters would miss multiple types of potential contamination.

Consequently, the TAC recommended a suite of priority water-quality parameters to determine if the cumulative impact of shale gas development activities has resulted in contamination in the Susquehanna River Basin. The suite of priority parameters is based on the specific hydrology, geology, past and current land use, and other environmental concerns identified in the Susquehanna River Basin as described in section 3.4; other parameters may be appropriate for other regions. Table 3 summarizes the suite of priority parameters chosen by the TAC to answer the policy question. The entire suite of priority surface-water parameters are needed to determine whether shale gas development is changing surface-water quality. In addition, a continuous record of streamflow is recommended at each monitoring site so the effects of variable streamflow on constituent concentrations can be evaluated. The entire suite of priority groundwater parameters are needed to determine whether shale gas development is changing groundwater quality.

The TAC noted that as industry practices change, so do the risks and potential monitoring priorities. For example, a modified set of monitoring parameters might be needed if industry changes to a new set of chemicals to use in fracturing fluids. As practices evolve, a mix of Marcellus Shale specific and non-specific indicators, spatial coverage across the Susquehanna River Basin, and adaptive management will allow for flexibility in water-quality interpretation over time.

4.3 Data Before and After Shale Gas Development

To be able to detect whether shale gas development is changing surface-water or groundwater quality in the Susquehanna River Basin, monitoring data should be available for a given monitoring site both before and after shale gas development began. The year 2007 was selected for this case study to represent the year that shale gas development began in the Susquehanna River Basin and was used to screen data basin-wide. For specific locations or monitoring sites, the first year of shale gas development within a watershed or aquifer may be later than 2007.

4.4 Surface-Water and Groundwater Data

The case-study policy question addresses both surface-water and groundwater quality. Surface water and groundwater are part of the same water cycle; long-term contamination of surface water usually occurs as a result of contaminated groundwater discharged to streams. Surface water and groundwater were separated in this report due to the different strategies typically used to monitor water quality in each hydrologic setting. Further criteria specific to surface-water and groundwater monitoring strategies are provided in Chapters 6 and 9, respectively.

5. Compilation and Screening of Water Data

In the Susquehanna River Basin, water-quality data are collected by at least 40 different organizations, and there is no single data repository used by all these organizations. In 2011, the USGS began assembling a multi-agency compilation of water-quality data into a centralized data set where water-quality data were merged into a consistent format. One of the objectives in assembling this data set, the National Data Aggregation, was to assist the USGS and other organizations in determining the availability and suitability of historic and current monitoring data for addressing regional, multi-state, and national-scale water-resources issues (Argue et al. 2014). A subset of the National Data Aggregation covering the Susquehanna River Basin ("Susquehanna data set" in this report) served as the primary basis for assessing existing water-quality monitoring in this case study. The Susquehanna data set also includes subsequent data additions resulting from suggestions from members of the TAC familiar with local monitoring programs. A complete multi-agency water monitoring data set was a critical component for completing this case study, and obtaining and processing this data set required expertise, time, and resources.

5.1 Water Data Types and Sources

The Susquehanna data set consists of water-quality and associated hydrologic data collected by Federal, State, and regional governmental agencies and non-governmental organizations. The compilation consists of a "snapshot" of historic and current monitoring data available at the time of the project. The data sources in the Susquehanna data set included online data repositories as well as data obtained directly from monitoring organizations. The largest online data sources included the NWIS database (U.S. Geological Survey, 1998a and U.S. Geological Survey, 2002a) and the USEPA Storage and Retrieval (STORET) Data Warehouse (U.S. Environmental Protection Agency, 2014). In addition, a significant effort was made to obtain monitoring records that were in electronic format but were not in NWIS or STORET by contacting individual agencies and organizations as well as locating data available in published reports.

The Susquehanna data set focused on water-quality data that were readily available and in electronic format. As a consequence, some industry data and studies subject to confidentiality agreements with homeowners were not included in the Susquehanna data set, although some researchers have obtained access to these data through special agreements. Industry data and other data subject to confidentiality agreements are discussed in detail in section 10.2.

The types of data targeted in the Susquehanna data set included water-quality records for surface-water and groundwater, and aquatic macroinvertebrate bioassessments. Water-quality records for surface water and groundwater consisted of inorganic, organic, and physical parameters, while the bioassessments included information on presence and diversity of macroinvertebrates at the stream site. Information was also compiled showing the available period of record for continuous surface-water monitoring sites.

The time lags between submission of a data request and receipt of the data varied substantially, ranging from immediate data transfer to 6 months or more. As a result, most data sets in the Susquehanna data set are complete through 2010, but some data sets include data collected through 2012 and into 2013.

Overall, the Susquehanna data set is considered complete through 2010. The earliest data records in the Susquehanna data set are from 1923 for surface water and 1925 for groundwater.

After the data were received, they were processed, recoded, and formatted to generate a data set in a uniform format such that the data could be queried, compared, and summarized. The data processing tasks included ensuring consistent naming of the parameters and chemical species measured; uniform coding of data qualifiers and units of measurement; removal of duplicate sites and data to the extent possible; and documentation of the organization's field and laboratory comment codes.

Table 4 lists organizations that collected surface-water-quality data in the Susquehanna River Basin and Table 5 lists organizations that collected groundwater-quality data. Both tables list the data providers from which those organizations' data were retrieved and are ranked by the number of sampling sites. For example, data collected by the Susquehanna River Basin Commission (SRBC) were retrieved from the STORET and Consortium of Universities for the Advancement of Hydrologic Sciences, Inc. (CUAHSI) databases, as well as directly from SRBC itself. Figure 12 shows that there are many more organizations collecting surface-water data than collecting groundwater data.

5.2 Water-Quality Monitoring Programs in the Susquehanna River Basin

Water-quality monitoring programs in the Susquehanna River Basin serve many different purposes including assessing trends in water quality, setting protection levels, identifying emerging problem areas, obtaining information in support of Clean Water Act requirements, and regulatory actions. This section provides a description of several monitoring programs in the basin that were included in the Susquehanna data set. There are new monitoring programs that have begun recently in the Susquehanna River Basin to look more closely at potential impacts of shale gas development in the region. Data for some of these programs were not available in time to be included in the Susquehanna data set, and they are described in sections 12.2 and 12.3.

5.2.1 Surface water

The SRBC has deployed an extensive network of continuous, real-time monitoring sondes throughout the Marcellus Shale area of the Susquehanna River Basin, with 58 monitoring sites as of October 2013. The monitoring network was established in January 2010 in response to unconventional shale gas development activities (Susquehanna River Basin Commission, 2012a; Susquehanna River Basin Commission, 2015). In addition, the SRBC collects discrete water samples four times a year at the monitoring sites. The PADEP, in coordination with the USGS, maintains a fixed-station water-quality network (WQN) to support assessment of surface-water quality in the State (Pennsylvania Department of Environmental Protection, 2005). The USGS measures streamflow and collects field chemistry samples and biological samples from designated stream and lake sites across the State.

Table 4. Sources and data providers of surface-water-quality data in the Susquehanna data set, ranked by number of sampling sites.
[Abbreviations: Same, data provider is the same as the data source; ALLARM, Alliance of Aquatic Resource Monitoring; CHRESI, Catskill Headwaters Research Institute; CUAHSI, Consortium of Universities for the Advancement of Hydrologic Sciences, Inc.; NGO, non-governmental organization; NURE, National Uranium Resource Evaluation; NWIS, National Water Information System; STORET, U.S. Environmental Protection Agency Modern STOrage and RETrieval data warehouse; USGS, United States Geological Survey]

Source name	Source type	Data provider	Number of sites	Number of records	Number of parameters
Pennsylvania Fish and Boat Commission	State	Same	7,756	36,768	27
U.S. Department of Energy – NURE Program	Federal	USGS	3,249	38,090	13
USGS National Water Information System (NWIS)	Federal	Same	1,749	614,060	62
Susquehanna River Basin Commission	Regional	STORET	859	84,617	37
Maryland Department of Natural Resources	State	Same	148	11,511	16
Pennsylvania Department of Environmental Protection	State	CUAHSI	116	4,355	28
Dickinson College – ALLARM	Academia	CUAHSI	90	180	2
Community Science Institute	NGO	Same	89	2,315	6
Dauphin County Conservation District	County	Same	89	7,828	17
Otsego County Conservation District	County	CHRESI	64	3,850	6
Susquehanna River Basin Commission	Regional	Same	57	12,226	33
New York Department of Environmental Conservation	State	Same	47	30,794	35
Maryland Department of the Environment	State	STORET	46	6,848	16
Susquehanna River Basin Commission	Regional	CUAHSI	45	4,590	23
New York Department of Environmental Conservation	State	STORET	40	19,837	34
Cortland County Soil and Water Conservation District	County	Same	39	5,655	24
Centre County Senior Environmental Corps	Volunteer	Same	37	11,887	12
U.S. Army Corps of Engineers, Baltimore District	Federal	Same	35	11,996	11
Shermans Creek Conservation Association	Volunteer	CUAHSI	34	2,799	5
Conodoguinet Creek Watershed Association	Volunteer	CUAHSI	32	2,135	4
Shermans Creek Conservation Association	Volunteer	Same	21	2,170	5
Pennsylvania State University	Academia	CUAHSI	18	898	27
Pennsylvania Senior Environmental Corps and Nature Abounds	Volunteer	Same	17	3,328	15

Source name	Source type	Data provider	Number of sites	Number of records	Number of parameters
U.S. Environmental Protection Agency	Federal	STORET	17	240	15
U.S. National Park Service	Federal	STORET	16	1,370	31
Clearwater Conservancy	Volunteer	Same	12	23,104	37
State University of New York at Oneonta	Academia	CHRESI	11	877	37
Northumberland County Conservation District	County	Same	11	4,513	8
Countryside Conservancy	Volunteer	Same	10	1,543	7
Clinton County Conservation District	County	Same	9	1,080	12
God's Country Trout Unlimited Chapter	Volunteer	CUAHSI	8	909	4
The E.L. Rose Conservancy of Susquehanna County	Volunteer	CUAHSI	7	192	2
Lancaster County Conservation District	County	Same	3	309	1
Pennsylvania Department of Environmental Protection	State	STORET	2	2,771	35
Maryland Department of Natural Resources	State	STORET	1	4	4
Appalachia Hydrogeologic and Environmental Consulting, Inc.	Private	CUAHSI	1	32	32
Renfrew Institute, Antietem Watershed Association	Volunteer	Same	1	198	7
Keystone Watershed Monitoring Network	Volunteer	STORET	1	136	8
Potomac Appalachian Trail Club	Volunteer	STORET	1	12	7
U.S. Department of Agriculture, Agricultural Research Service	Federal	Same	1	7,595	3
Virginia Department of Environmental Quality	State	Same	1	48	8

Table 5. Sources and data providers of groundwater-quality data in the Susquehanna data set, ranked by number of sampling sites.
[Abbreviations: CHRESI, Catskill Headwaters Research Institute; CUAHSI, Consortium of Universities for the Advancement of Hydrologic Sciences, Inc.; NGO, non-governmental organization; NURE, National Uranium Resource Evaluation; same, data provider is the same as the data source; USGS, United States Geological Survey]

Source name	Source type	Data provider	Number of sites	Number of records	Number of parameters
USGS National Water Information System	Federal	same	4,235	129,898	69
Pennsylvania Department of Environmental Protection	State	USGS	2,692	67,465	37
U.S. Department of Energy – NURE Program	Federal	USGS	2,265	25,375	12
Pennsylvania Topographic and Geologic Survey	State	same	279	1,317	48
Community Science Institute	NGO	same	96	2,172	25
Pennsylvania State University	Academia	CUAHSI	66	518	25
Chester County Health Department	County	same	56	112	2
State University of New York at Oneonta	Academia	CHRESI	48	1,310	30
ClearWater Conservancy	Volunteer	same	7	5,686	21
Appalachia Hydrogeologic and Environmental Consulting, Inc.	Private	CUAHSI	1	32	31

At the county level, conservation districts in Pennsylvania have deployed several continuous monitors to monitor changes due to agriculture, development, or land use changes including shale gas development.

New York State Department of Environmental Conservation’s Division of Water has a surface-water monitoring program, called the Rotating Integrated Basin Studies program. The objectives of this program are to asses and characterize water quality of all New York State surface waters on a rotating basis. In addition, routine trend monitoring is done at fixed sites across the State.

Volunteer stream monitoring groups that are actively sampling stream water quality to identify potential shale gas development impacts in the Susquehanna River Basin include the Centre County Pennsylvania Senior Environmental Corps, the Alliance for Aquatic Resource Monitoring (ALLARM), and Trout Unlimited (TU).

Centre County Pennsylvania Senior Environmental Corps focuses on the streams of Centre County, Pennsylvania. The organization works cooperatively with many organizations including the ClearWater

Conservancy, the Centre County Conservation District, the Pennsylvania Chapter of Trout Unlimited, and the PADEP. The organization provides its monitoring data online, as well as its documentation of training methods and quality control procedures. Volunteers monitor in areas of shale gas development in Beech Creek Watershed, and submit monthly samples for analysis to Lock Haven University. The ClearWater Conservancy includes the Spring Creek Watershed Association, which maintains several stream monitoring sites.

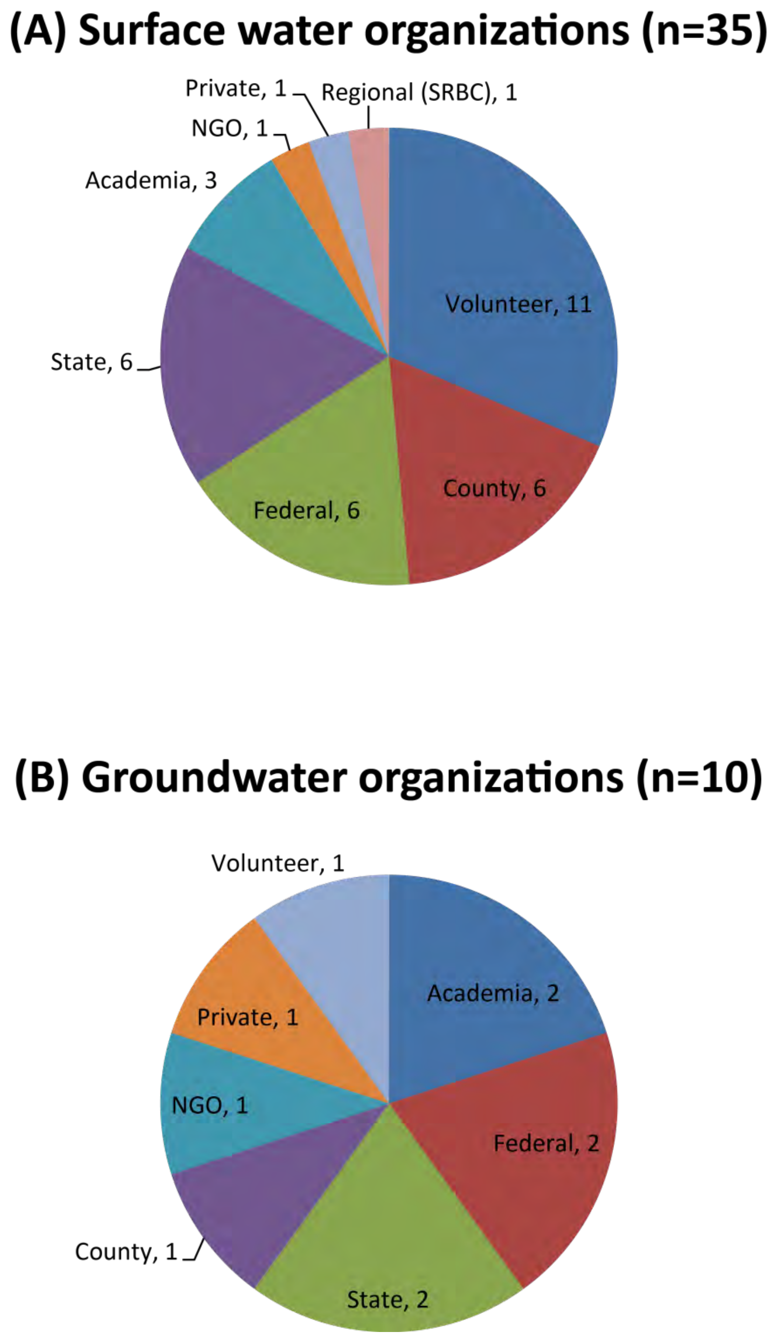


Figure 12. Number of organizations collecting water-quality data in the Susquehanna data set by type of organization for (A) Surface water, and (B) Groundwater.

ALLARM is based out of Dickinson College and has provided technical support to communities that wish to monitor water quality for the impacts of shale gas development, providing training manuals, protocols, and workshops. ALLARM-trained groups have sampled water quality at 272 sites in Pennsylvania and New York, and ALLARM has analyzed 116 of those sites, although only 23 of these sites had nearby shale gas development at the time of sampling (C. Wilderman, ALLARM, written commun., May, 2014). The water-quality data collected include specific conductance, total dissolved solids, stream stage, barium, and strontium. ALLARM is developing a central database to share the water data (C. Wilderman, ALLARM, written commun., May, 2014).

Trout Unlimited (TU) has developed the Pennsylvania Coldwater Conservation Corps, an extensive volunteer stream monitoring program designed to gather baseline data in watersheds where shale gas development is not yet occurring and to identify acute pollution events in watersheds where shale resources are actively being developed. Data collected by TU include conductivity, dissolved solids, temperature, turbidity, pH, barium, and strontium. Water-quality data have been collected for 237 sites in Pennsylvania and in Broome County, New York; 108 of these sites are in the Susquehanna River Basin. TU uses its Conservation Success Index (CSI)—a science-based landscape-scale geospatial analysis and planning tool—to focus monitoring efforts in watersheds most likely to experience shale gas development. Volunteers upload data to TU's online data portal hosted by CitSci.org, an online database developed by Colorado State University for citizen science projects. Currently, TU staff is conducting quality assurance/quality control (QA/QC) on the 3,350 records collected by TU volunteers over the past 4 years.

The Pennsylvania Fish and Boat Commission's (PFBC) Unassessed Waters Initiative is a partnership between the PFBC and qualified universities, research entities, and conservation organization partners to conduct stream surveys. These groups are directed by the PFBC to sample previously unassessed potential wild trout streams that are determined to be most at risk due to development and Marcellus Shale drilling operations and to use sampling protocols established by the PFBC (Weisberg, 2011). Although sites are typically sampled only once through this program, the number of stream sites has increased from 152 sampled in 2008 to 587 in 2012.

5.2.2 Groundwater

The National Uranium Resource Evaluation (NURE) program was initiated by the Atomic Energy Commission in 1973 with a primary goal of identifying uranium resources in the United States. As a result of this program a large number of surface-water and groundwater samples were collected in the 1970s. The NURE data set represents the largest set of available groundwater records, but those data are limited to one sample per site and usually lack important site information, such as well type, well depth, aquifer, and aquifer material.

The non-profit Community Science Institute has implemented a program in the southern counties of New York where shale gas development is most likely to occur if New York lifts its ban on high-volume hydraulic fracturing. The goal is to create a baseline of groundwater-quality data by pooling the results of tests from domestic wells. Homeowners can choose which parameters to have tested, and an extensive list of parameters is offered that includes about half of the suite of parameters identified in Table 3. None of the

Community Science Institute groundwater sampling sites in the Susquehanna data set have been sampled for all of their offered parameters. The pooled results of domestic well tests are presented in an online database using anonymous formats that safeguard homeowners' privacy. The Community Science Institute has tested over 300 domestic wells for baseline analysis since 2009 and has asked these homeowners for their permission to enter their results anonymously into their database, with over 80 percent giving permission to do so (Community Science Institute, 2015). In order to protect homeowners' privacy, drinking water wells are mapped to 1-mile grid squares.

Individual studies have assessed groundwater quality in selected parts of the Susquehanna River Basin. These studies have described background levels of methane in several counties in the Susquehanna River Basin that may experience shale gas development in the future, for example, Heisig and Scott (2013), and Sloto (2013).

The New York State Department of Environmental Conservation's (NYSDEC) Division of Water has a groundwater monitoring program in collaboration with the USGS. As a part of the groundwater program, samples were collected from 16 public and 14 domestic wells in the upper Susquehanna River Basin in order to characterize water quality in sand and gravel aquifers, as well as bedrock aquifers (Reddy and Risen, 2012). In another study in the upper Susquehanna River Basin conducted in 2004 to 2005, samples were taken from 20 public-supply wells and 13 domestic wells in order to describe the chemical quality of groundwater, representing areas of high vulnerability to contamination (Hetcher-Aguila and Eckhardt, 2006).

5.3 Screening Available Monitoring Data

The TAC prepared a screening list of water-quality parameters relevant to shale gas development in the Susquehanna River Basin. The presence of these parameters might indicate impacts from shale gas development on surface water (Table 6) or groundwater (Table 7). This screening list was used in an initial step to identify the quantities of monitoring data available and assess the completeness of the metadata in the Susquehanna data set. The TAC used best professional judgment and a variety of data sources to identify the parameters for both surface water and groundwater. The lists include parameters that the TAC knew to be associated with each contamination pathway, for example changes in suspended sediment concentration in surface water can be associated with clearing and construction (Streets, 2012). Many of the groundwater parameters were selected from a USEPA case study in Bradford County, Pennsylvania (U.S. Environmental Protection Agency, 2015c). The TAC included parameters required for sampling by the PADEP for producers of shale gas waste (Pennsylvania Department of Environmental Protection, 2011), as well as the most frequently used and most toxic chemicals referenced in a Congressional report on hydraulic fracturing fluid components (U.S. House of Representatives, 2011). The TAC also included more common parameters that are useful in measuring changes in water quality but are not necessarily contaminants specific to shale gas development. The parameters for surface water and groundwater are presented in Table 6 and Table 7, respectively, and the locations of the monitoring sites with water-quality data for those parameters are shown in Figure 13 and Figure 14, respectively.

Table 6. Comprehensive list of surface-water parameters selected by the Technical Advisory Committee for querying the Susquehanna data set, ranked by number of sampling sites.
[Abbreviations: NA, not applicable]

Parameter	Fraction type: Filtered/Unfiltered/ Combined/Unknown[1]	Number of sites	Number of data records	Number of data sources
Water temperature	NA	14,044	56,340	31
pH	NA	13,611	58,639	30
Specific conductance	NA	13,548	59,070	26
Alkalinity	Combined	11,672	28,053	24
Hardness	NA	7,325	24,142	11
Streamflow	NA	5,589	43,317	13
Chloride	Combined	5,425	26,553	15
Sodium	Combined	5,059	16,898	12
Dissolved oxygen	NA	4,435	37,066	24
Fluoride	Combined	3,417	9,906	6
Aluminum	Unknown	3,342	4,515	5
Uranium	Combined	3,316	3,557	4
Manganese	Unknown	3,093	4,085	4
Bromide	Combined	2,817	4,216	7
Sulfate	Combined	2,539	29,693	18
Nitrate	Combined	2,133	29,942	29
Phosphorus, total	Unfiltered	1,834	26,091	14
Ammonia	Combined	1,753	25,990	13
Iron, total	Unfiltered	1,742	21,331	9
Total dissolved solids	Combined	1,662	19,149	13
Manganese, total	Unfiltered	1,599	16,958	8
Aluminum, total	Unfiltered	1,378	14,915	9
Turbidity	NA	1,300	16,756	16
Calcium, total	Unfiltered	1,261	13,236	9
Magnesium, total	Unfiltered	1,255	13,435	9
Orthophosphate	Combined	1,099	23,293	19
Nitrite + Nitrate	Combined	959	16,549	8
Total suspended solids	NA	956	14,301	10
Iron, dissolved	Filtered	943	10,178	5
Zinc	Combined	940	13,796	13
Total Kjeldahl Nitrogen	Combined	918	13,309	7
Manganese, dissolved	Filtered	895	10,066	6

Parameter	Fraction type: Filtered/Unfiltered/ Combined/Unknown[1]	Number of sites	Number of data records	Number of data sources
Magnesium	Unknown	885	1,706	3
Magnesium, dissolved	Filtered	858	10,285	6
Calcium, dissolved	Filtered	851	10,310	5
Acidity	NA	849	11,802	10
Potassium	Combined	836	8,249	10
Aluminum, dissolved	Filtered	704	7,896	9
Biochemical oxygen demand	NA	702	4,673	4
Silica	Combined	678	5,899	7
Lead	Combined	579	11,969	11
Copper	Combined	565	12,071	11
Nickel	Combined	521	10,813	10
Barium, total	Unfiltered	491	3,785	6
Arsenic	Combined	436	3,958	8
Cadmium	Combined	436	5,301	10
Strontium, total	Unfiltered	435	1,818	5
Selenium	Combined	325	2,488	7
Chromium	Combined	323	2,852	8
Phosphorus, total dissolved	Filtered	264	4,270	7
Boron	Combined	257	1,245	3
Stage	NA	247	7,595	4
Strontium, dissolved	Filtered	236	1,189	3
Mercury	Combined	227	2,367	6
Lithium, total	Unfiltered	214	722	3
Suspended sediment concentration	NA	196	1,533	3
Silver	Combined	194	1,646	5
Cobalt	Combined	187	1,552	5
Barium, dissolved	Filtered	154	1,071	4
Lithium, dissolved	Filtered	150	786	1
Beryllium	Combined	144	865	4
Molybdenum	Combined	139	641	2
Iron	Unknown	118	1,232	6
Phenols	Unfiltered	96	1,710	2
Gross beta, total	Unfiltered	88	1,020	2
Gross alpha, total	Unfiltered	87	1,013	2

Parameter	Fraction type: Filtered/Unfiltered/ Combined/Unknown[1]	Number of sites	Number of data records	Number of data sources
Methylene blue active substances	Unfiltered	72	586	2
Strontium	Unknown	57	528	3
Gross alpha	Unknown	51	89	2
Oil and grease	Unfiltered	51	651	2
Gross beta	Unknown	50	88	2
Chemical oxygen demand	NA	49	976	1
Calcium	Unknown	48	869	2
Naphthalene	Combined	32	94	2
Gross alpha, dissolved	Filtered	27	226	1
Gross beta, dissolved	Filtered	27	226	1
Radium-226	Filtered	12	114	2
Radium-228	Filtered	10	11	2
Benzene	Unfiltered	8	39	1
Toluene	Unfiltered	8	39	1
Barium	Unknown	1	23	1
Lithium	Unknown	1	23	1
Methane	Filtered	1	1	1
Sulfur	Combined	0	0	0
Ethylene glycol	Combined	0	0	0
Thorium	Combined	0	0	0
Surfactants	Combined	0	0	0

[1]For some parameters, the TAC chose to report filtered, unfiltered, and unknown fractions together.

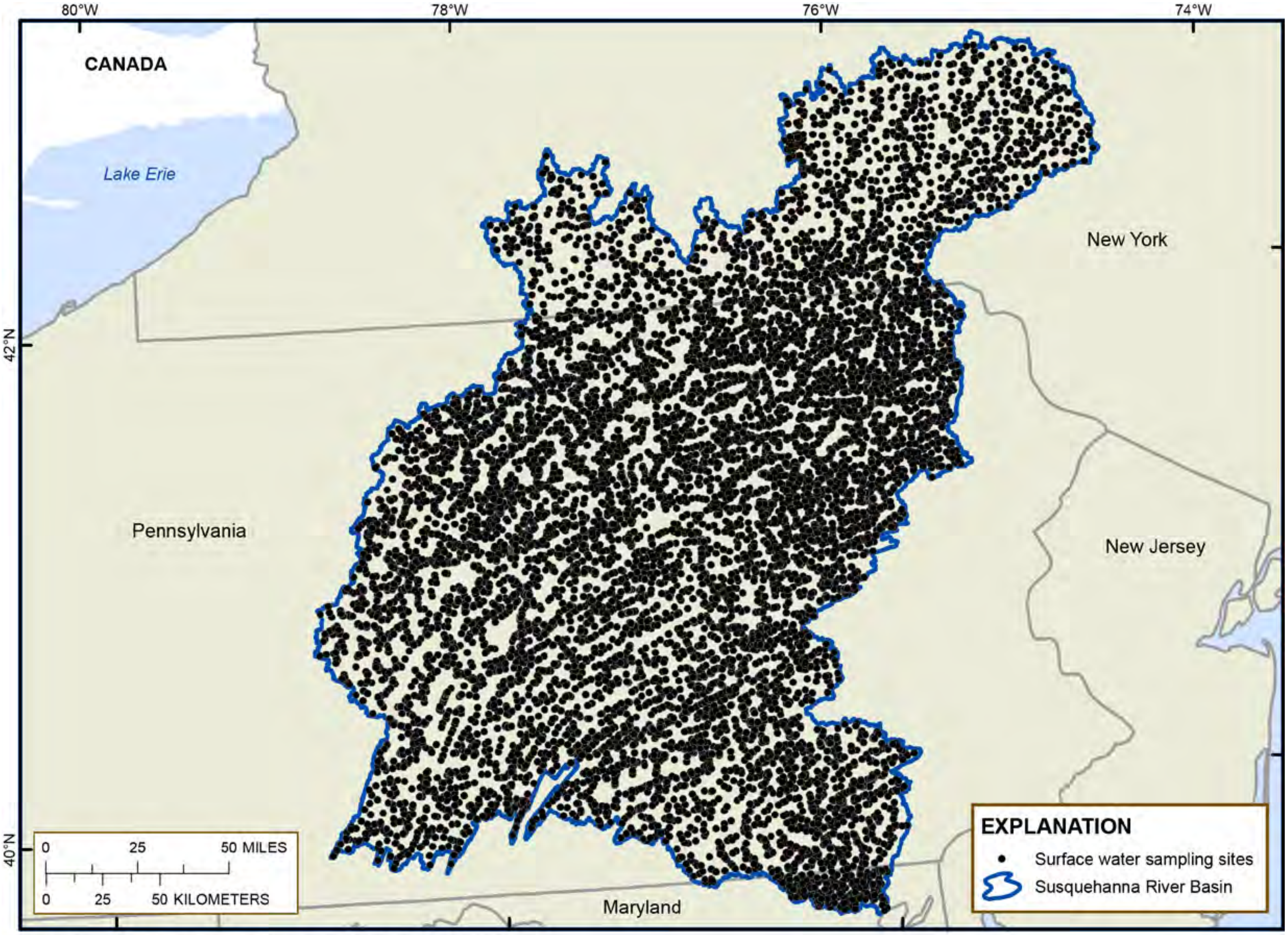
Table 7. Comprehensive list of groundwater parameters selected by the Technical Advisory Committee for querying the Susquehanna data set, ranked by number of sampling sites.

Parameter	Number of sites	Number of data records	Number of data sources
pH	6,082	12,302	10
Specific conductance	5,997	10,878	8
Chloride, dissolved	5,648	7,099	5
Nitrate	5,402	9,997	8
Alkalinity	5,303	7,111	6
Water temperature	5,115	8,203	7
Sodium, dissolved	5,010	6,491	7
Manganese, dissolved	4,596	5,972	7
Fluoride, dissolved	4,554	5,162	4
Magnesium, dissolved	4,378	5,714	6
Aluminum, dissolved	3,584	4,442	7
Sulfate, dissolved	3,250	4,514	4
Total dissolved solids	3,214	6,912	5
Calcium, dissolved	2,919	4,173	4
Hardness	2,908	6,459	5
Iron, dissolved	2,774	4,223	6
Uranium, dissolved	2,714	2,781	5
Potassium, dissolved	2,634	3,759	4
Nitrite	2,428	7,302	3
Ammonia	2,291	7,389	3
Bromide, dissolved	2,194	2,357	5
Arsenic	1,745	4,232	6
Zinc, dissolved	1,484	2,406	6
Nitrite + Nitrate	1,392	4,333	2
Dissolved oxygen	1,222	3,156	5
Chromium, dissolved	1,211	2,117	6
Nickel, dissolved	1,191	2,024	5
Lead, dissolved	1,153	2,072	6
Cadmium, dissolved	1,118	1,982	6
Iron, total	1,036	4,424	5
Manganese, total	787	2,609	5
Barium, total	668	2,494	5
Copper, dissolved	650	1,566	6
Barium, dissolved	633	1,348	5

Parameter	Number of sites	Number of data records	Number of data sources
Zinc, total	580	2,662	4
Lead, total	578	2,658	4
Chromium, total	569	2,651	3
Strontium, dissolved	546	1,190	4
Cadmium, total	537	2,614	4
Mercury, total	534	2,422	3
Calcium, total	531	3,677	5
Total Kjeldahl Nitrogen	513	1,473	1
Boron, dissolved	494	677	2
Sodium, total	468	3,838	4
Potassium, total	460	3,444	3
Copper, total	453	2,507	4
Cobalt, dissolved	439	665	4
Strontium, total	424	439	4
Magnesium, total	423	3,553	4
Methylene blue active substances	422	433	2
Chloride, total	410	3,786	5
Toluene	385	695	4
Benzene	372	680	4
Ethylbenzene	368	675	4
Silver, dissolved	368	606	2
Xylene	362	665	3
Aluminum, total	356	630	4
Nickel, total	340	585	4
Dissolved organic carbon	335	791	2
Lithium, dissolved	335	520	3
Methane	302	344	2
Selenium, dissolved	299	472	3
Mercury, dissolved	278	407	6
Sulfate, total	263	3,492	3
Molybdenum, dissolved	257	415	3
Fluoride, total	229	821	2
Beryllium, dissolved	199	309	2
Lithium, total	164	177	1
Acidity	161	699	2

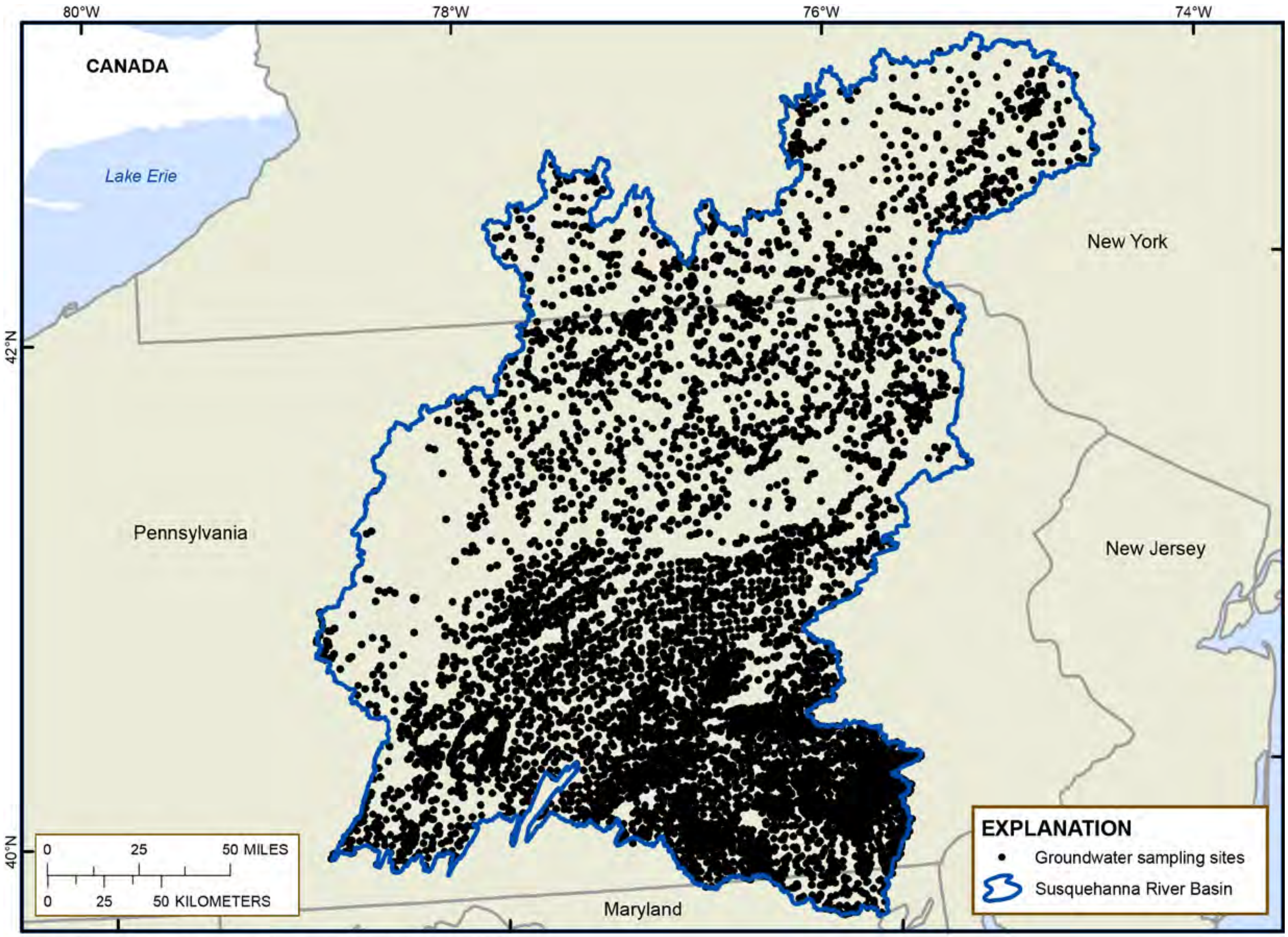
Parameter	Number of sites	Number of data records	Number of data sources
Cobalt, total	161	217	3
Silver, total	161	242	3
Beryllium, total	159	215	3
Molybdenum, total	150	163	1
Phenols	139	438	3
Naphthalene	134	153	2
Oxygen-18:Oxygen-16 ratio in water	132	134	2
Selenium, total	131	217	3
Gross alpha, total	129	129	2
Gross beta, total	129	129	2
Hydrogen-2:Hydrogen-1 ratio in water	129	131	2
Gross alpha, dissolved	116	122	2
Gross beta, dissolved	116	120	2
Radium-226	111	116	2
Radium-228	111	115	2
Total suspended solids	102	282	3
Uranium, total	102	112	1
Turbidity	101	103	2
Chemical oxygen demand	100	100	2
Biochemical oxygen demand	83	83	1
Ethane	63	63	2
Boron, total	51	56	2
Carbon-13:Carbon-12 ratio in inorganic carbon	42	44	1
Silica	39	51	1
Sulfur, dissolved	38	72	1
Carbon-13:Carbon-12 ratio in methane	36	36	1
Hydrogen-2:Hydrogen-1 ratio in methane	36	36	1
Oil and grease	20	23	1
Iodide, dissolved	9	27	1
Sulfate, unknown fraction	7	8	1
Dissolved inorganic carbon	6	6	1
Manganese	2	2	1
Bromide, total	1	1	1
Propane	1	1	1
2-butoxythenol	0	0	0

Parameter	Number of sites	Number of data records	Number of data sources
Acetic acid	0	0	0
Acrylamide copolymer	0	0	0
Aliphatic alcohol polyglycol ether	0	0	0
Ammonium acetate	0	0	0
Beryllium	0	0	0
Butane	0	0	0
Diesel	0	0	0
Ethoxylated alcohol	0	0	0
Ethoxylated nonyl phenol	0	0	0
Ethylene glycol (1,2-ethanediol)	0	0	0
Formaldehyde	0	0	0
Guar gum	0	0	0
Heavy aromatic petroleum naphtha	0	0	0
Heavy catalytic reformed petroleum naphtha	0	0	0
Hydrotreated light petroleum distillates	0	0	0
Isopropanol	0	0	0
Light catalytic cracked petroleum distillates	0	0	0
Medium aliphatic solvent petroleum naphtha	0	0	0
Methanol	0	0	0
Mineral spirits (Stoddard solvent)	0	0	0
Petroleum distillates	0	0	0
Petroleum gas oils	0	0	0
Polyethylene glycol	0	0	0
Polyethylene-polypropylene glycol	0	0	0
Polypropylene glycol	0	0	0
Propargyl alcohol (2-propyn-1-ol)	0	0	0
Propylene glycol (1,2-propanediol)	0	0	0
Silver	0	0	0
Strontium-87:Strontium-86 ratio	0	0	0
Thorium-228	0	0	0
Thorium-230	0	0	0
Thorium-232	0	0	0



State and county lines from U.S. Geological Survey, 2005, 1:2,000,000-scale digital data
Albers projection, NAD 1983

Figure 13. Surface-water monitoring sites in the Susquehanna River Basin with water-quality records for at least one of the comprehensive list of surface-water parameters selected by the Technical Advisory Committee (n=14,730)(Table 6).



State and county lines from U.S. Geological Survey, 2005, 1:2,000,000-scale digital data
Albers projection, NAD 1983

Figure 14. Groundwater sampling sites in the Susquehanna River Basin with water-quality records for at least one of the comprehensive list of groundwater parameters selected by the Technical Advisory Committee (n=9,761) (Table 7).

This screening process was a first attempt at a data summary; the process provided insight into several issues with the available water data including the existence of duplicate records and missing metadata for some sites. In order to perform spatial analysis on groundwater data, the aquifer should be known. Many sampling sites did not include metadata identifying the aquifer, so an attempt to identify the aquifer at each groundwater monitoring site was completed. For sites identified as springs, the aquifer was estimated from a geologic map using a Geographic Information System (GIS). Some wells did not have the aquifer identified but the well depth was identified. In some instances, the well depth could be used to determine if the well was finished in bedrock or unconsolidated (alluvial) deposits by comparing the well depth to a thickness map of unconsolidated deposits (Soller et al., 2012). For many wells, the aquifer could not be determined beyond “unknown bedrock” or “unknown.” Some aquifers were grouped for the analyses presented in this case study. For example, the aquifer "Middle Ordovician Carbonates and Shales" is a combination of several limestone and shale formations between the Reedsville Formation and the Beekmantown Group.

Table 3 identifies the suite of priority parameters that are a subset of the comprehensive list that the TAC identified as most useful for determining whether shale gas development is changing surface-water or groundwater quality.

5.4 Land Cover and Streamgage Matching

Two primary types of ancillary spatial information were used in this case study—spatial data that delineated hydrologic boundaries and were used to compute the drainage areas for surface-water monitoring sites, and descriptive spatial data such as land cover and geologic maps. The contributing drainage area of each surface-water monitoring site was computed using the database Geospatial Attributes of Gages for Evaluating Streamflow (GAGES) (Falcone et al., 2010; James A. Falcone, U.S. Geological Survey, written commun., December 2013), NWIS, GIS, or a combination of these sources. Land cover was determined using ArcGIS V10 software (Esri, Inc., 2012), the 2006 National Land Cover Database (Fry et al., 2011), and the National Agricultural Statistics Service Cropland Data Layer (U.S. Department of Agriculture, 2012).

5.5 Oil and Gas Well and Disposal Datasets

Online databases maintained by the NYSDEC and PADEP were used to compile an inventory of all types of oil and gas wells within the Susquehanna River Basin. New York’s oil and gas well database was downloaded on August 27, 2015 (New York State Department of Environmental Conservation, 2015). Pennsylvania’s oil and gas well database was accessed on August 28, 2015, and the “Spud Data Report” option was used to download well records for the entire State (Pennsylvania Department of Environmental Protection, 2015a). The spud date for a well is the date that drilling began. In both states, many wells are too old for the actual spud date to be known but a generic spud date of 1/1/1900 (NY) or 1/1/1800 (PA) is used for the well record. Well records without a latitude-longitude were deleted because their locations could not be displayed in maps. Well records in New York without a spud date were retained if they had a completion date or a plugged date; however, if a well record did not have one of these dates, the record was deleted under the assumption that the well was never drilled (for example, the permit expired).

NYSDEC's well records include formation names, but the spud data reports from the PADEP do not. The PADEP's "Oil and Gas Well Formations Report" option was used to download well records with formation names, which were merged with the well records from the spud data report. Newer, unconventional wells were more likely to have formations identified than older, conventional wells.

The intention of the oil and gas well inventory is to provide a measure of well locations in the Susquehanna River Basin, especially HVHF wells. The resulting inventory of oil and gas wells is composed of active wells, inactive wells, and plugged or abandoned conventional and unconventional wells (Figure 3 and Figure 10). The most recent spud date in the oil and gas well inventory is 10/20/2014 in New York and 8/21/2015 in Pennsylvania.

The PADEP identifies unconventional wells in their well records database. For the purposes of this report, a subset of these wells was classified as HVHF wells. Since the state databases are not explicit in designating HVHF wells, the identification of HVHF wells includes some interpretation. HVHF wells were assumed if the well was classified as unconventional and the well orientation was horizontal or directional. Vertical wells in unconventional formations were not considered to be HVHF wells. There are a total of 4,810 wells in the oil and gas well inventory that were designated as HVHF wells using these criteria (Figure 2, Figure 4). Over 99 percent of the HVHF wells produce from the Marcellus Shale and 0.3 percent (15) produce from the Utica Shale. Most of the wells have a horizontal orientation; only two have a directional orientation. The NYSDEC does not identify unconventional wells in their well records database, but there are wells that target or produce from unconventional formations. Since there is a ban on HVHF in New York, all of the HVHF wells in the Susquehanna River Basin are in Pennsylvania.

Figure 15 identifies the density of HVHF wells by the USGS's 12-digit hydrologic unit code (HUC12 watersheds). The highest density of HVHF wells is in the area close to the northern Pennsylvania border with New York and in the eastern portion of the basin.

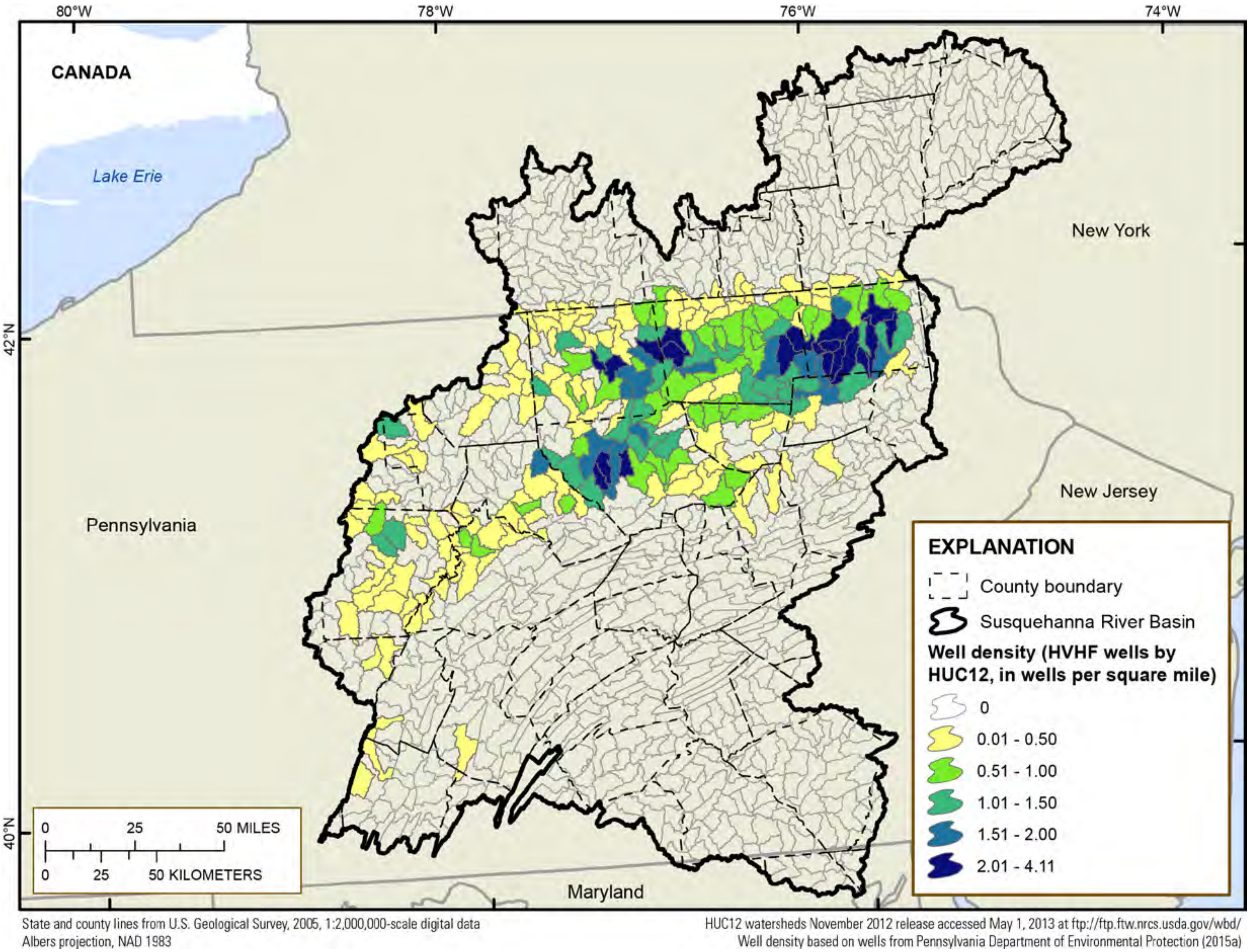


Figure 15. Density of HVHF wells in the Susquehanna River Basin by HUC12 watershed.
[Abbreviations: HVHF, high-volume hydraulic fracturing, HUC12, 12-digit hydrologic unit code boundaries]

6. Quantities of Surface-Water Data Needed to Answer the Case-Study Policy Question

This chapter defines the quantifies of surface-water data needed to answer “Do shale gas development activities contaminate surface water or groundwater?” in the Susquehanna River Basin, building on the criteria discussed in Chapter 4. The overarching objectives for answering the case-study policy question are to (1) be able to determine whether concentrations of parameters associated with shale gas development have changed over time (identify temporal trends), and (2) determine whether the detected change is the result of shale gas development (identify cause of trend). These objectives, plus a testable hypothesis, were used to develop a sampling plan that defines the types of surface-water data, frequency and duration of monitoring, location criteria, and number of monitoring sites needed to answer the case-study policy question.

6.1 Focus Parameters – Surface Water

From the suite of priority surface-water parameters, the TAC identified three surface-water focus parameters to be used as examples to quantify the data needed to answer the policy question. These focus parameters are specific conductance and dissolved and total barium. Specific conductance was the parameter with the most records and the longest periods of record for individual monitoring sites in the Susquehanna data set, making it the parameter most likely to have enough data to detect whether a statistically significant change has occurred. Barium is a highly specific signature of produced waters (Vidic et al., 2013) and can serve as a possible indicator of shale gas activity, whereas specific conductance can show potential but non-specific impacts. Total barium had among the longest and most plentiful records in the Susquehanna data set. Total and dissolved barium concentrations can be quite different in surface water, and may represent different aspects of shale gas development. For example, dissolved barium may be most associated with produced waters, while total barium is more closely associated with drilling mud or cuttings due to its presence in drilling mud.

6.2 Hypothesis – Surface-Water Quality

Clear identification of monitoring objectives and the hypotheses to be tested are critical for determining a successful water-quality monitoring design (Sanders and others, 1983; Gilbert, 1987; U.S. Environmental Protection Agency, 2002 and 2006). Identifying an appropriate monitoring design, to which the existing monitoring data could be compared, was important for determining the availability of data and gaps in current monitoring. The following hypothesis reflects a possible scenario of shale gas development resulting in water-quality change, and is used to identify the quantity of data needed to answer the policy question.

- Cumulative effects of shale gas development using high-volume hydraulic fracturing will cause impacts to surface-water quality at the small watershed scale over the long-term, resulting in statistically significant increases of specific conductance and/or barium concentrations at sampling sites that monitor watersheds smaller than approximately 70 square miles.

This hypothesis is referred to as the “surface-water hypothesis” throughout this document.

The following core set of water data is needed to achieve the first overarching objective for answering the case-study policy question, identify temporal trends in water quality:

- Specific conductance or barium data collected with sufficient frequency and duration to detect trends in flow-adjusted concentrations, and
- Streamflow data near the water-quality monitoring site, to calculate flow-adjusted concentrations.

Constraints on the types of monitored watersheds are expected to maximize the ability to statistically discern such trends:

- Locate monitoring sites in watersheds with a high density of HVHF wells and complimentary reference monitoring watersheds with no planned HVHF development, and
- Maintain consistency in sampling design between watersheds to allow comparison between high density and reference watersheds.

To determine whether observed trends in specific conductance or barium concentrations result from shale gas development activities (the second objective), additional data and analysis are needed for monitored watersheds to determine which factors are playing a role in observed trends:

- Data documenting shale gas development activities within monitored watersheds,
- Data on changes in other potential sources of specific conductance and barium within monitored watersheds, and
- Data on climate conditions including precipitation and streamflow.

This chapter identifies the quantities of specific conductance, total barium, and dissolved barium data needed to detect whether concentrations are changing. While these focus parameters provide an illustration of data needed, the full suite of priority surface-water parameters is needed to determine if contamination observed is related to the cumulative effects of shale gas development activities in the Susquehanna River Basin or is related to a different source of contaminants.

6.3 Statistical Design to Test the Hypothesis – Surface-Water Monitoring

Trend detection is ideal for determining if long-term and/or low-magnitude changes are occurring in surface water. To test the surface-water hypothesis, sufficient data must be collected using an appropriate monitoring design such that statistically significant trends in water quality can be detected through statistical analysis.

Monotonic trend analysis on flow-adjusted concentrations has been widely used as a statistical tool for assessing water-quality trends, and assumes that the trend is consistently increasing or decreasing over time. Monotonic trend analysis is an appropriate tool for evaluating temporal trends that might result from cumulative shale gas development activities. Log-linear regression was used in this study with the Susquehanna data set to estimate the duration of monitoring required to statistically discern water-quality trends in tributaries to the Susquehanna River. This method requires that there are no large gaps in the data set, no changes in sampling or laboratory methods over time (unless the effect of method changes can be characterized statistically), available data for important covariates that may explain variations in water quality (e.g., other water-quality parameters in Figure 6), streamflow data, and land use activity data for the duration of the study.

Monotonic trend study design also lends itself to bringing data together from multiple organizations. When multiple agencies collect data over the long term at each agency's own monitoring sites, the data record at each site can be analyzed to detect a trend and the trends can be compared across different sites due to long-term consistency at each monitoring site.

6.4 Sampling Frequency

The sampling frequency for testing the hypotheses must support trend analysis of flow-adjusted concentrations. The TAC selected regular monthly water-quality samples for the preferred monitoring frequency needed to capture seasonal variation throughout the year. These samples should not reflect storm conditions. However, samples taken over longer time intervals and including seasonal representation, such as quarterly or bimonthly samples, can also be used.

6.5 Locations and Number of Monitoring Sites

6.5.1 Monitoring sites in each ecoregion with shale gas development

To understand if shale gas development is affecting streams or rivers, it is essential to characterize background water chemistry for each ecoregion. Stream chemistry in each ecoregion is unique, and will respond differently to disturbances or changes in land use. Each ecoregion has its own water chemistry baseline against which change can be assessed.

Given the differences in water chemistry across the ecoregions of the Susquehanna River Basin, monitoring sites are needed in each of four ecoregions with active or planned shale gas development (Figure 16). The Northern Allegheny Plateau spans the northern portion of the Susquehanna River Basin and its open valleys and low mountains support woodlands and agriculture. It is also an ecoregion that has experienced significant shale gas development, particularly in northeastern Pennsylvania. The North Central Appalachians ecoregion is a sedimentary upland that has high hills and low mountains. A substantial portion of the shale gas development in the Susquehanna River Basin is located in this ecoregion. The Ridge and Valley ecoregion contains an area of parallel ridges and valleys, with a mix of forest and agricultural land use. There is not as much shale gas development in this ecoregion overall because much of the shale is thermally overmature and therefore unlikely to contain economically recoverable gas resources. The Central Appalachian ecoregion covers a portion of the Marcellus and Utica Shale area of the Susquehanna River Basin, and is primarily forested with some agricultural land use. This area does not currently have many HVHF wells, but may experience more development in the future.

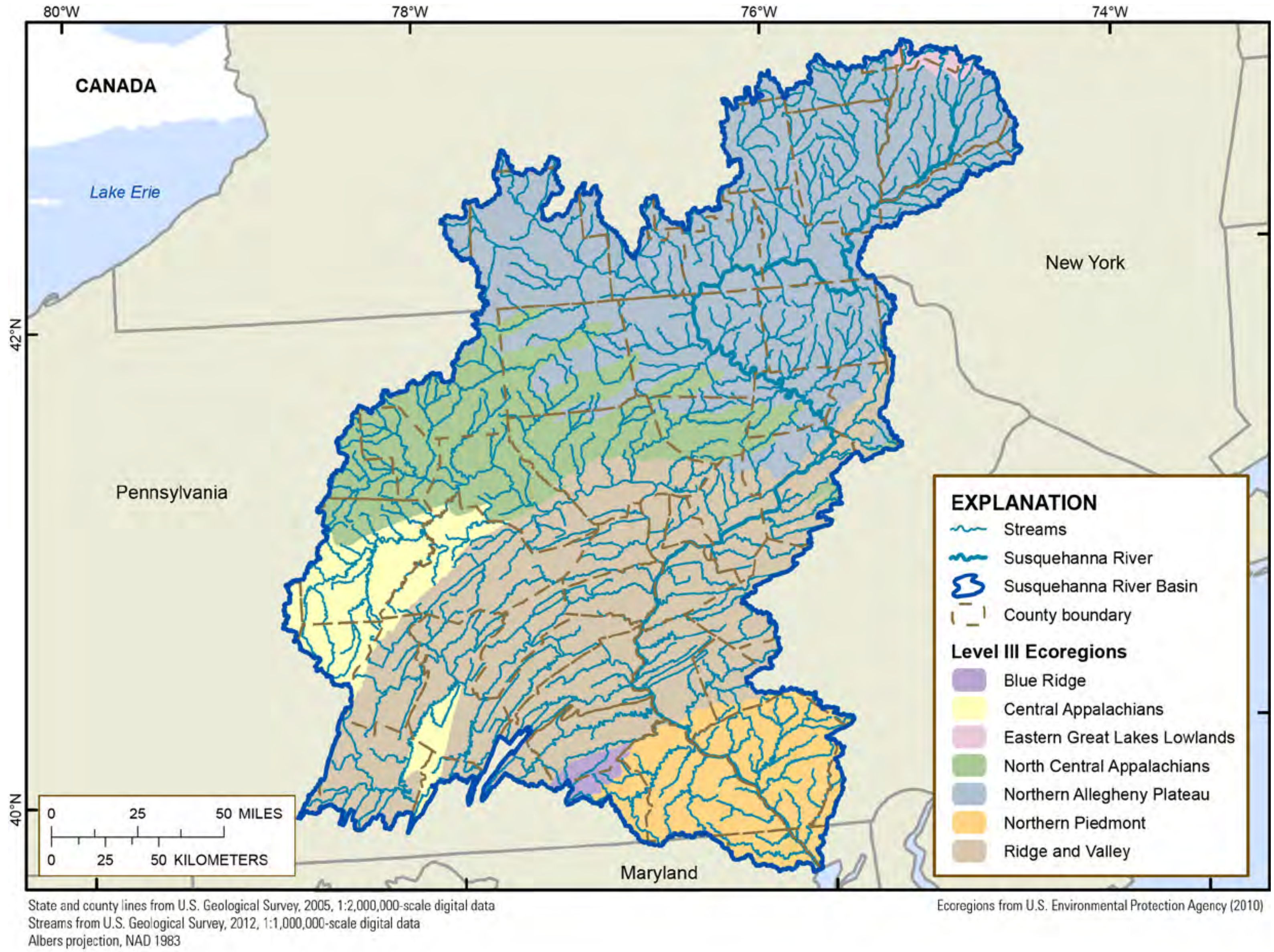


Figure 16. Ecoregions in the Susquehanna River Basin.

Water chemistry in these ecoregions is characterized in Appendix A (Chapter 15) to this report and by the SRBC (2012a, 2015). The differences in baseline concentrations and background variability indicate that trends will express themselves differently in each ecoregion, and monitoring sites in each ecoregion would allow those changes to be quantified.

6.5.2 Monitoring sites in small watersheds

Small watersheds provide the best opportunities to identify pollutants that are primarily derived from a single source; the idea is to isolate factors that might contribute to pollution concentrations, so that the impact from a given source can be better understood (Scientific and Technical Advisory Committee, 2013). The TAC determined that a small watershed size of approximately 70 square miles or less would be ideal for monitoring and took into consideration the decision of the SRBC to use a similar watershed scale of 10 to 70 square miles for its water-quality monitoring network. As a result, watersheds smaller than 70 square miles in the Susquehanna River Basin were selected as the priority monitoring scale for this study. The TAC determined that this size range is small enough to isolate known factors that can have an impact on water quality, yet are potentially large enough to contain upstream shale gas development. While this scale of monitoring may not capture individual spills due to the effects of dilution, this scale is useful for monitoring and detecting the cumulative effects of shale gas development.

The bigger the watershed, the more dilution occurs from water unaffected by a given pollution incident, and the more attenuation of the parameter in the waterway. At the large watershed or regional scale, effects of shale gas development will be more difficult to detect due to dilution. Spills may not be detectable at the large watershed scale because of the larger volume of water associated with large watersheds; upstream shale gas development may not be as dense per unit area as in smaller watersheds. An additional problem is that many of the potential contaminants associated with shale gas development are also associated with other sources such as agriculture, coal mining, road deicing, industrial solvents, fuels, etc. Large watersheds were not selected as a monitoring priority by themselves, but they can provide additional context for interpreting small watershed monitoring results when small watershed monitoring sites are nested within larger monitored watersheds. As a result, opportunities were explored for nested monitoring with larger watersheds in this case study.

6.5.3 Monitoring sites in high density and reference watersheds

It is important to have monitoring sites in both watersheds with a high density of shale gas development, and reference watersheds without shale gas development and minimal expectations for future development within each monitored ecoregion. This experiment/control monitoring design allows for trend detection in watersheds with HVHF wells while also detecting whether similar trends exist in reference watersheds that may result from sources and causes not related to shale gas development. For this study, watersheds with greater than 0.5 HVHF wells per square mile represent medium density watersheds, and watersheds with greater than 2.0 HVHF wells per square mile represent high density watersheds. Watersheds with no HVHF well permits were selected to represent reference watersheds. Figure 17 shows an example of a surface-water monitoring site in the Susquehanna River Basin with several nearby HVHF wells in the Marcellus Shale, and other vertical oil and gas wells.

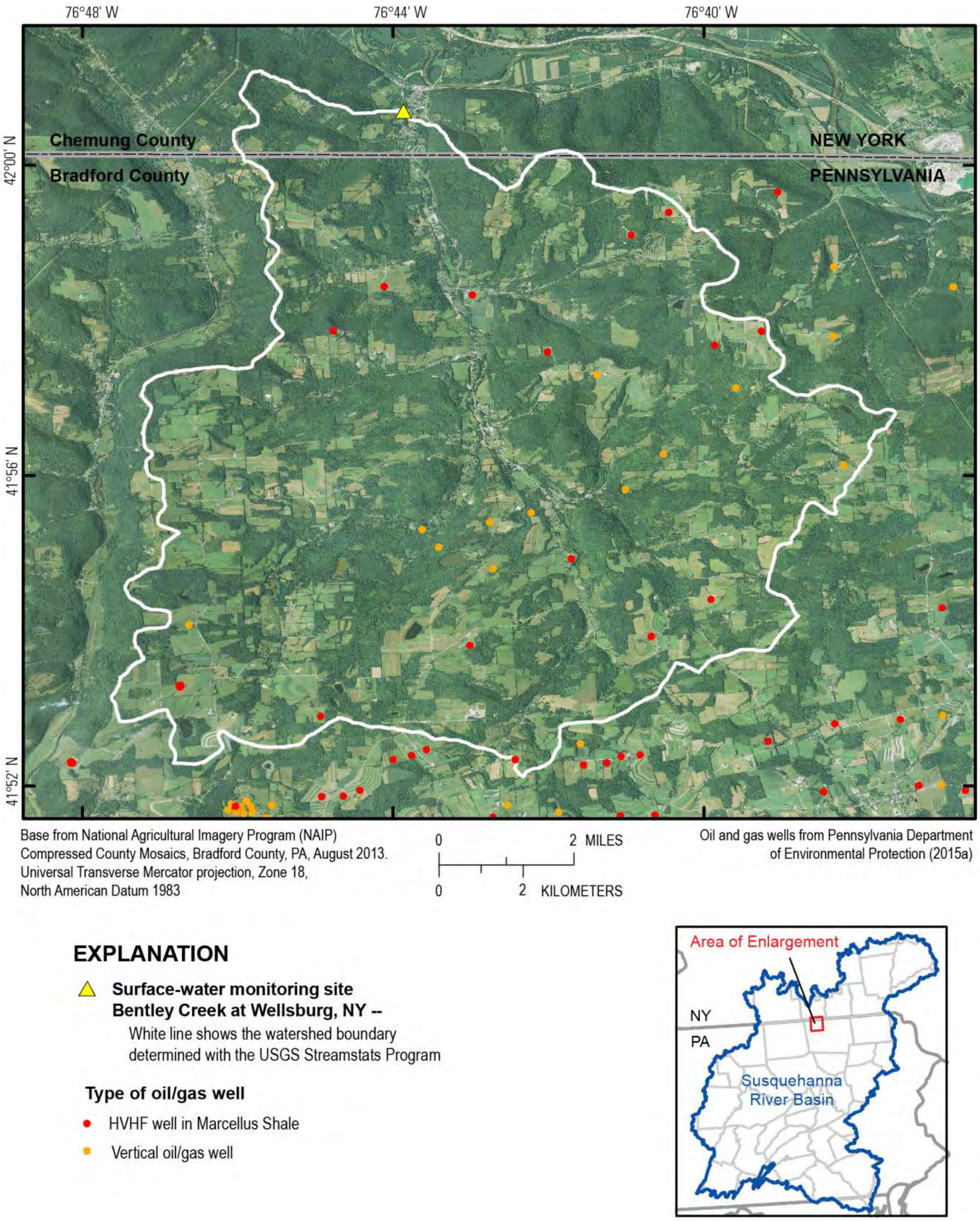


Figure 17. Surface-water monitoring site Bentley Creek at Wellsburg, NY and the locations of oil and gas wells within and near the watershed (as of 8/21/2015). The calculated well density for this watershed is 0.3 HVHF wells per square mile; the watershed is 54.2 square miles.
[Abbreviations: USGS, U.S. Geological Survey; HVHF, high-volume hydraulic fracturing]

At least one watershed with medium or high density of HVHF wells and one reference watershed should be monitored per ecoregion to answer the case-study policy question. As there are four priority ecoregions for monitoring, a minimum of eight monitoring sites are needed to test the hypothesis and answer the case-study policy question. Additional monitoring sites, including nested monitoring sites, will increase the ability to identify potential water-quality changes resulting from shale gas development in the Susquehanna River Basin.

6.6 Duration of Monitoring to Detect Change

The minimum duration of sampling required to statistically discern water-quality trends is influenced by a number of factors, principally the sampling frequency, the inherent variability of the system (i.e. concentrations and streamflow), the ability of the trend method to remove extraneous variability other than land use effects, and the rate and magnitude of water-quality change anticipated in the watershed. The TAC selected an approach for estimating the duration of sampling needed to detect change that included use of focus parameter data from the Susquehanna data set, a 5-parameter log-linear regression flow-adjusted trend analysis model, and a “power analysis”. The power analysis is used to estimate the minimum duration of monitoring needed to identify a statistically significant trend in specific conductance or barium concentrations using a specific sampling frequency, at specified levels of significance and power.

Power, in the statistical context, relates to the probability of not committing a Type II error (Somers, 1997), and in this study it is the ability of a statistical test to detect water-quality trends that may result from shale gas development if that trend actually exists. In watersheds exhibiting a weak water-quality signal (small rate of change) and characterized by highly variable water-quality and/or streamflow systems, longer monitoring records will be required to discern temporal trends.

Type I Error: Detecting an effect that is not present (incorrectly rejecting the null hypothesis).

Type II Error: Failing to detect an effect that is present (incorrectly failing to reject the null hypothesis).

Power analysis uses the characterization of historical natural variability in a parameter of interest, specific conductance and barium in this case, and estimates the number of samples needed to detect a signal over the background variation, or “noise”. The method for the power analysis is presented in detail in Appendix B (Chapter 16). To conduct the power analysis for surface water, a time series with approximately 36 independent samples or more is preferred, and ideally those samples would be collected over a period of roughly 3 to 4 years at approximately monthly intervals.

Figure 18 shows all sites within or near the Marcellus and Utica Shale area in the Susquehanna data set that have barium or specific conductance data in each ecoregion. The Susquehanna data set was queried for monitoring sites with 36 or more monthly records, in small watersheds, with corresponding streamflow data, and located in or near the Marcellus and Utica Shale area to use in the power analysis. Table 8 presents the list of surface-water monitoring sites that were used in the power analysis, sorted by ecoregion.

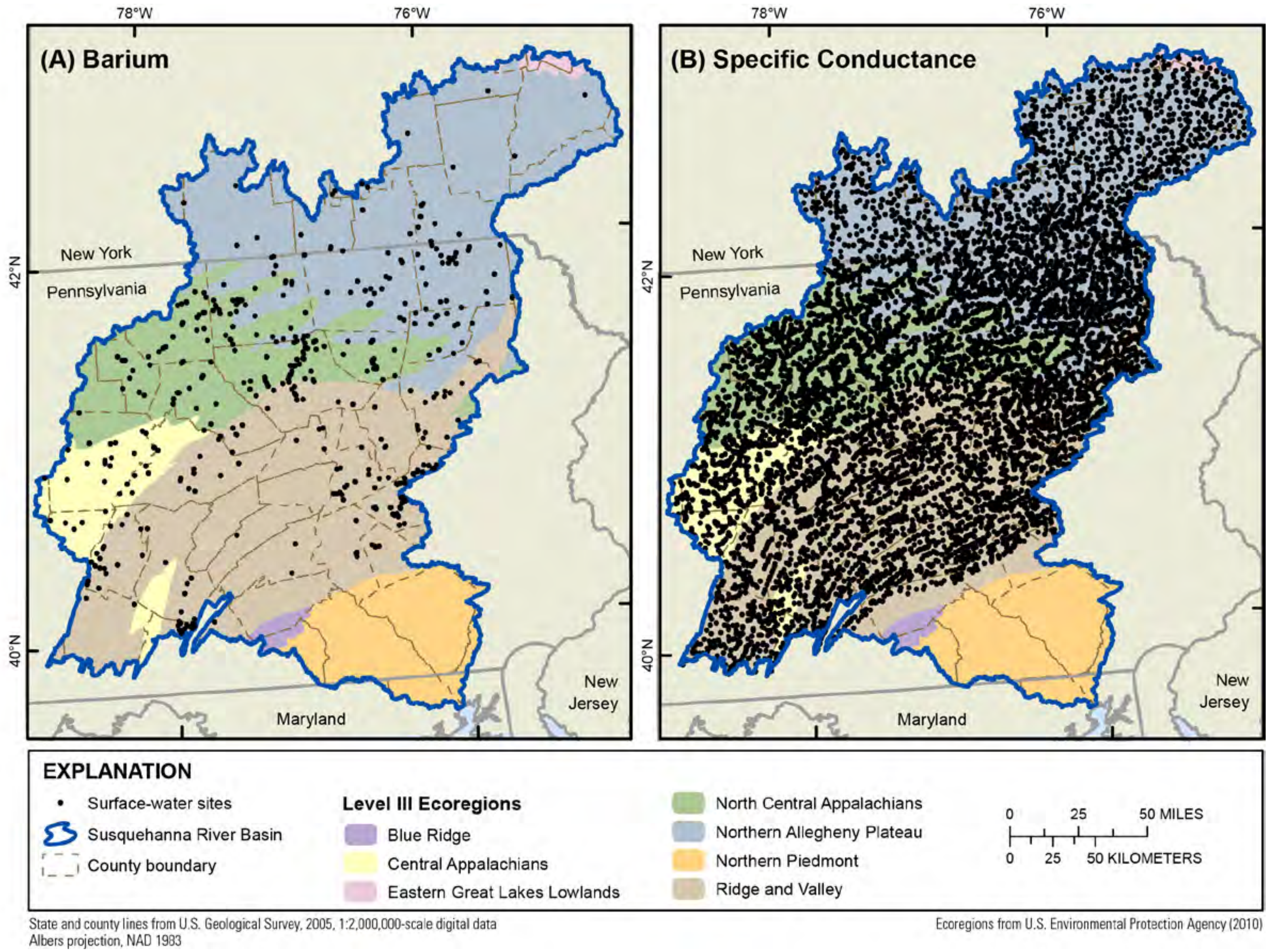


Figure 18. Surface-water monitoring sites in the Susquehanna data set in or near the Marcellus and Utica Shale area by USEPA Level III Ecoregion with data available for (A) Barium (total and dissolved, n=549), and (B) Specific Conductance (n=11,890).
[Abbreviations: USEPA, U.S. Environmental Protection Agency]

Table 8. Data-rich monitoring sites in the Susquehanna data set that were evaluated using power analysis to estimate sample-size requirements for detecting temporal trends in barium and specific conductance, sorted by ecoregion.
[Abbreviations: mi^2, square miles; ecoregion, U.S. Environmental Protection Agency Ecoregion]

Monitoring site identification number	Site name	Drainage area (mi^2)	Water-quality parameter
	Northern Allegheny Plateau ecoregion		
01528000	Fivemile Creek near Kanona, NY	66.8	Specific conductance
BNTY000.9	Bentley Creek at Wellsburg, NY	54.2	Specific conductance
01508800	Factory Brook at Homer, NY	15.8	Specific conductance
01509150	Gridley Creek above East Virgil, NY	10.4	Specific conductance
01533610	Unnamed Tributary to Tunkhannock Creek at Gelatt, PA	9.02	Total barium
	North Central Appalachians ecoregion		
01545600	Young Womans Creek near Renovo, PA	46.2	Dissolved barium, Specific conductance
01548476	Cedar Run above Mine Hole Run near Cedar Run, PA	26.3	Total barium, Specific conductance
01548423	Wilson Creek at Morris, PA	22.8	Specific conductance
01538709	West Branch Fishing Creek near Elk Grove, PA	20.2	Total barium, Specific conductance
	Ridge and Valley ecoregion		
01571820	Swatara Creek at Ravine, PA	43.3	Dissolved barium, Total barium, Specific conductance
01557990	Sinking Run near Spruce Creek, PA	28.3	Total barium, Specific conductance
01569195	Conodoguinet Creek above Reservoir near Roxbury, PA	27.2	Total barium, Specific conductance
0155979602	Bobs Creek below Wallacks Branch at Pavia, PA	22.1	Total barium, Specific conductance
CCPASEC_1216	Lick Run at Howard, PA	11.4	Specific conductance

Figure 19 shows concentration of total barium over time at the West Branch Fishing Creek surface-water monitoring site in the Susquehanna River Basin. Analysis of these data per the power analysis method presented in Appendix B (Chapter 16) yields the graph shown in Figure 20, which shows the number of samples required to detect change at various error levels versus change in median total barium concentration.

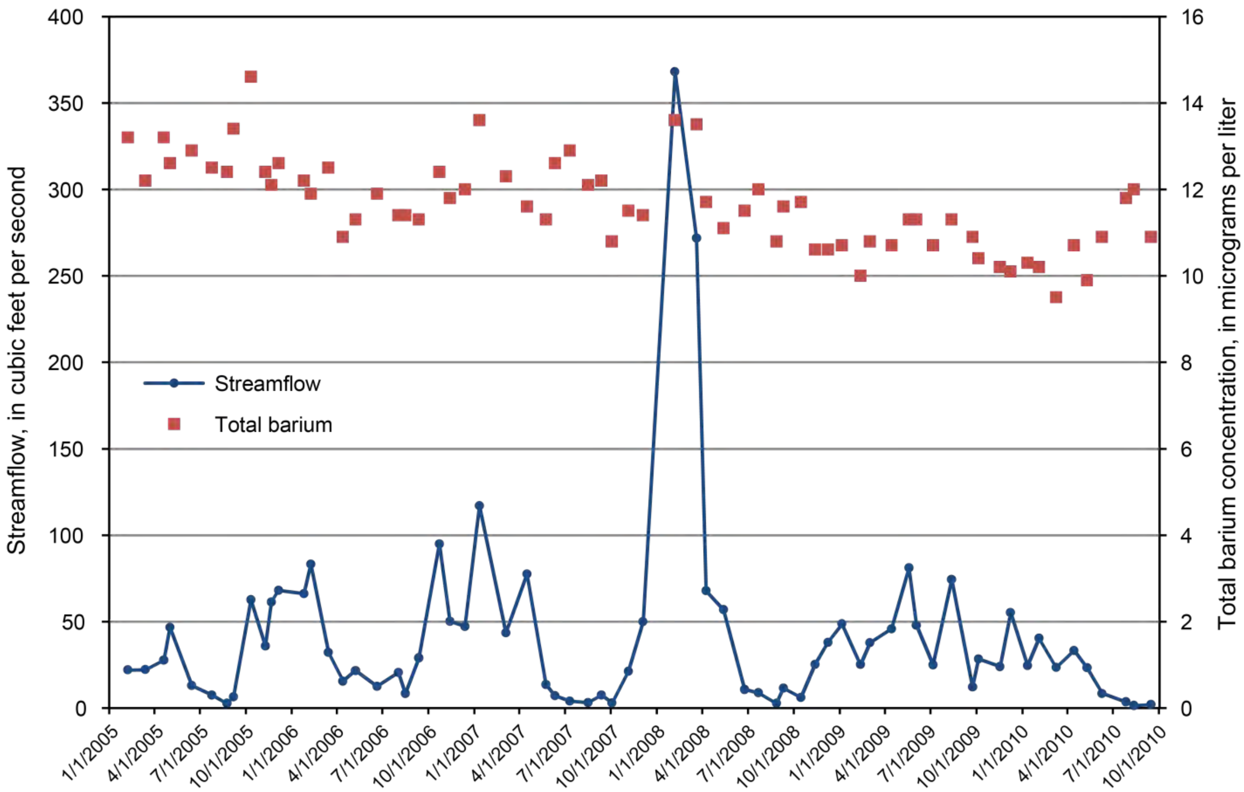


Figure 19. Time series of streamflow and total barium concentrations at the surface-water monitoring site West Branch Fishing Creek near Elk Grove, PA (01538709).

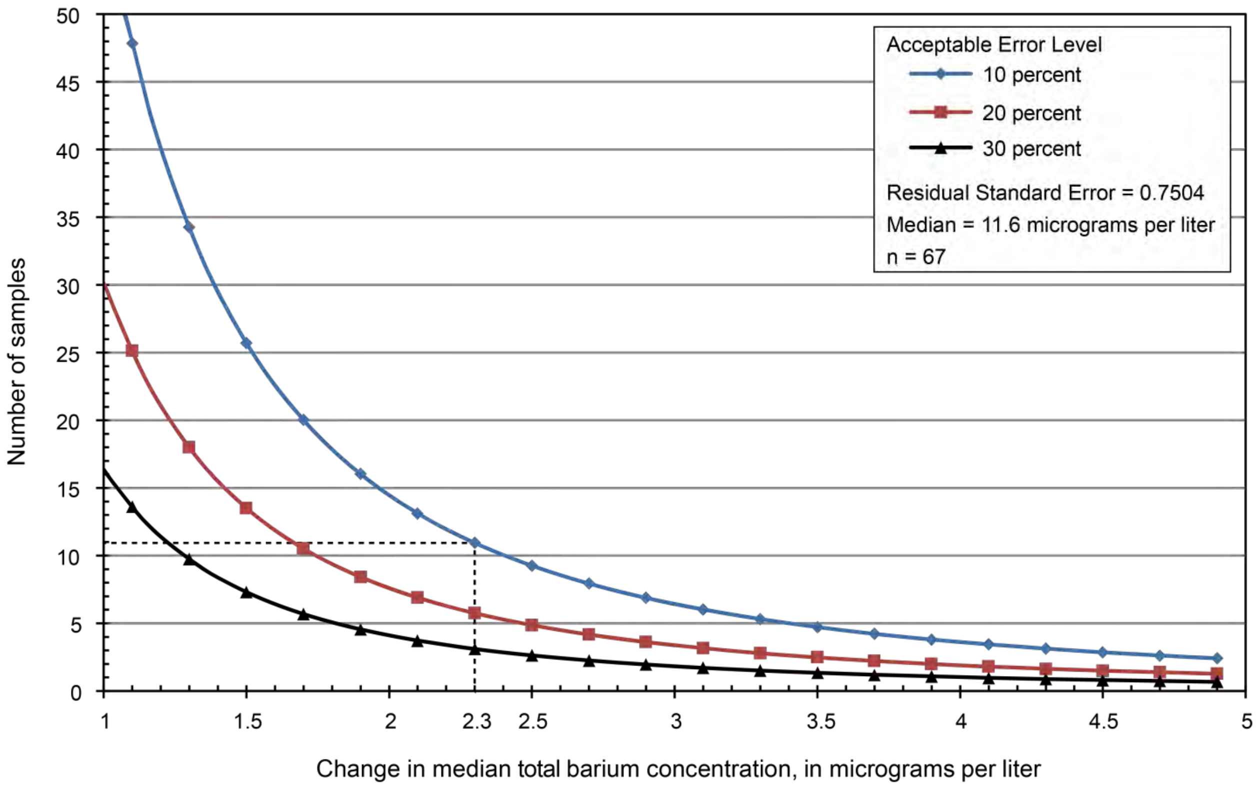


Figure 20. Power analysis estimates of the number of monthly samples needed to detect trends in median total barium concentration, for different error levels, for West Branch Fishing Creek near Elk Grove, PA (01538709). The power analysis method is described in Appendix B.

The power analysis gives insight into the number of samples that are needed to detect change for a given error level for each monitoring site. Figure 20 illustrates the number of samples needed to detect a change of 2.3 micrograms per liter (µg/L) (20 percent) in median total barium concentration for West Branch Fishing Creek. The three error curves in Figure 20 each represent both statistical significance (α) and power (β), which were set as equal and analyzed at levels of 10-, 20-, and 30-percent. In this example, the 10-percent error curve indicates that 11 samples would be required to detect a 2.3-µg/L (20-precent) change in median total barium concentration at this monitoring site. This means if the true trend was 20 percent there would be a 10-percent chance (β=0.1) that the test would fail to identify that trend, or a 90-percent chance of correctly identifying the trend; alternatively, if there was no true underlying trend, there would be a 10-percent chance (α=0.1) that the test would incorrectly identify a trend (due to random chance), or an 90-percent chance of correctly identifying the lack of a trend. If variability increases in the future or is greater than identified in the available records, a longer duration of monitoring would be needed. If variability decreases in the future or is less than identified in the records analyzed, a statistically significant trend may be identified sooner than estimated here. The number of samples required to detect change is a function of the background variability, magnitude of change, and the desired power and significance level. Fewer samples are required to detect a given amount of change for larger error levels or a larger magnitude change in concentration.

One must first have enough baseline data to characterize background variability after accounting for seasonal effects, streamflow effects, and pre-existing trends through the power analysis before conclusions can be made about whether change has occurred. The analysis for this case study used a minimum of 36 monthly samples to characterize background variability. In some cases, the power analysis may indicate that the number of samples needed to detect a given magnitude of change is less than the number of samples used to characterize the background variability as occurred in the Fishing Creek example in Figure 20. However, the minimum of 36 monthly samples is needed to characterize background variability and reflect a range of meteorological conditions to confirm that observed water-quality changes are not an artifact of short-term weather patterns. Therefore, in the case of the Fishing Creek monitoring site, a minimum of 36 monthly samples is needed.

As indicated in the power analysis results (Figure 20), the anticipated magnitude of change influences the duration of sampling needed to statistically discern trends. Not much is known about the potential cumulative effects of shale gas development in a small watershed or potential levels of contamination because intensive shale gas development is a relatively new phenomenon. The question of whether there is an impact in small watersheds has not been resolved although several related studies have been conducted. Olmstead et al. (2013) examined changes in water quality in large watersheds in Pennsylvania showing increased concentrations of chloride in watersheds with wastewater treatment plants that accept shale waste, and increased concentrations of total suspended solids in watersheds with shale gas development (average watershed size of 580 square miles). Warner et al. (2013b) studied local water quality and sediment near commercial wastewater treatment plants that process wastewater from oil and gas development. A USEPA review of state and industry spill data identified 32 hydraulic fracturing-related spills that reached surface water, but the USEPA did not determine whether spilled fluids affected the quality of surface-water resources (U.S. Environmental Protection Agency, 2015b).

Concentrations of contaminants in brines can be used to estimate an upper limit on the amount of change that might be observed in areas potentially affected by shale gas development. Average total dissolved solid concentrations in produced water range from 800 to 300,000 milligrams per liter (mg/L); dissolved solids concentrations range from 100-500 mg/L in freshwater and about 35,000 mg/L in ocean water (Olmstead et al., 2013). Additional comparisons can be made using data provided by the PADEP for Marcellus Shale brine (T. Shaw, Pennsylvania Department of Environmental Protection, written commun., November 2013). The plot in Figure 21 shows concentrations of parameters in produced water found in 30 samples within the Susquehanna River Basin obtained by the PADEP, plotted against baseline surface-water quality found in Young Womans Creek, a highly forested watershed in Pennsylvania. The plot is an illustration of the difference between brine and fresh water, and is not meant to characterize water quality in the Susquehanna River Basin. The parameters are arranged in order of decreasing median value for the PADEP data, and a log scale is used to help show individual points over the large range in values. Maximum contaminant levels, health advisory levels, and secondary drinking water standards are also shown for reference.

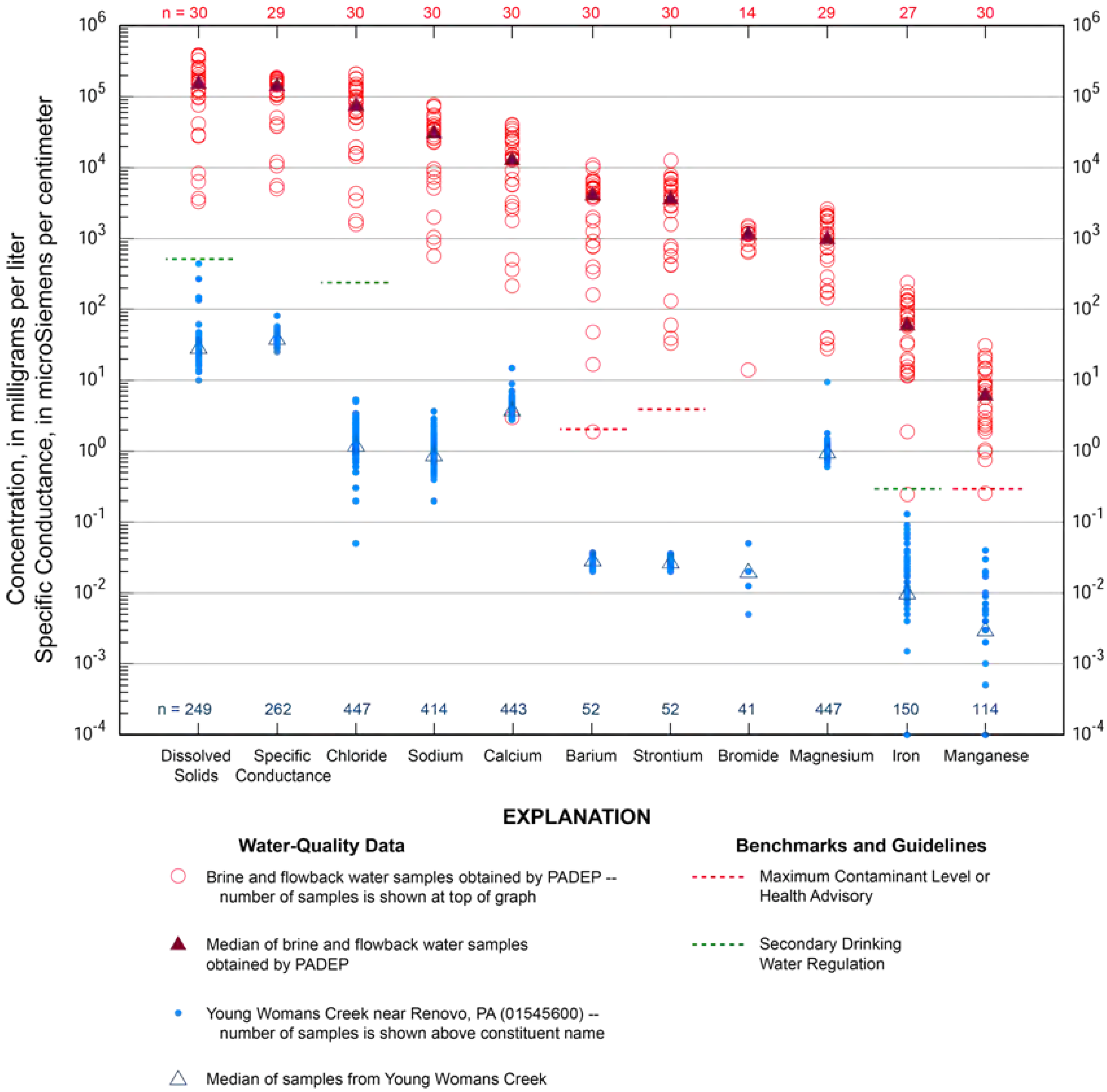


Figure 21. Concentrations of selected parameters in produced water from HVHF wells in the Susquehanna River Basin compared to background or baseline surface-water quality represented by the monitoring site Young Womans Creek near Renovo, PA (01545600).
[Abbreviations: PADEP, Pennsylvania Department of Environmental Protection; HVHF, high-volume hydraulic fracturing]

The amount of produced water required to change the chemistry of a stream in a watershed relatively unaffected by human activities such as shale gas development, agriculture, urbanization or other types of land disturbance, is small. For example, the amount of produced water required to double the median concentration of barium in 1 liter of water collected from Young Womans Creek (0.029 mg/L), assuming a median barium concentration of 4,155 mg/L in the produced water (Figure 21), would be approximately 0.07 milliliter of produced water—or 1.4 drops of produced water assuming there are 20 drops in 1 milliliter.

The magnitude and location of a contamination event relative to a monitoring site would also have a significant impact on contaminant concentrations measured at the site. Given the number and complexity of contamination pathways, the relative lack of information about the cumulative effects of shale gas development, the range of concentrations in produced waters, and the variation in land use and water chemistry, it is difficult to assign an expected magnitude of change in barium or specific conductance concentrations. A change of 20 percent from the median concentration was used here to illustrate the duration of monitoring needed to detect water-quality change in tributaries to the Susquehanna River.

The duration of monitoring required to detect change is also a function of acceptable error levels. Accepting a large error level increases the chance of indicating a trend when in fact no trend is actually occurring (or concluding that there is no trend when, in fact, a trend exists), and will require fewer samples than smaller acceptable error levels. Given monitoring results from a set of monitoring sites that have been in place for several years, managers might choose a large error level as a first-cut filter for identifying which watersheds warrant further investigation, especially for an indicator parameter like specific conductance, which indicates the possibility of upstream contamination but does not identify the cause of that contamination. The policy choice of what error level to choose is a decision for managers who must balance the resources available for further investigation against the number of sites that are monitored. The choice of error level can be iteratively and adaptively modified as ongoing investigations reveal what is actually happening in a watershed.

The estimated years of monthly monitoring needed to detect a 20-percent change is given in Table 9 for specific conductance and Table 10 for total barium. These tables summarize the number of years of sampling needed to detect a 20-percent change from median conditions, at 10- and 20-percent error levels, for monitoring sites that collect both concentration and streamflow data. The values are sorted by ecoregion. There are fewer monitoring sites with barium data available in the Susquehanna River Basin meeting the power analysis requirements than there are for specific conductance.

Table 9. Power analysis estimates of the minimum number of years of monthly sampling needed to detect a 20-percent change from median values for specific conductance, by ecoregion, for watersheds smaller than 70 square miles in the Susquehanna River Basin. The power analysis method is described in Appendix B.
[Abbreviations: ecoregion, U.S. Environmental Protection Agency Level III Ecoregion]

Ecoregion	Number of Sites Analyzed	Minimum years of monthly sampling to detect a 20-percent change in median value for specific conductance	
		20% error	10% error
Northern Alleghany Plateau	4	3	3-4
North Central Appalachians	4	3-5	3-10
Ridge and Valley	5	3-6	3-11

Table 10. Power analysis estimates of the minimum number of years of monthly sampling needed to detect a 20-percent change from median values for total barium, by ecoregion, for watersheds smaller than 70 square miles in the Susquehanna River Basin. The power analysis method is described in Appendix B.
[Abbreviations: ecoregion, U.S. Environmental Protection Agency Level III Ecoregion]

Ecoregion	Number of Sites Analyzed	Minimum years of monthly sampling to detect a 20-percent change in median concentration for total barium	
		20% error	10% error
Northern Alleghany Plateau	1	3	5
North Central Appalachians	2	3	3
Ridge and Valley	3[1]	3-5	3-9

[1] Results from Swatara Creek at Ravine, PA were not included here because one total barium sample that was not representative of the population was included in the power analysis, resulting in a very large estimate of the number of samples needed to detect a 20-percent change in concentration.

Table 9 shows there were differences in the ecoregions with respect to specific conductance. For most sites in the Northern Allegheny Plateau ecoregion, a minimum of 3 years of monthly specific conductance data are needed to determine whether a 20-percent change from the median has occurred at the 20-percent error level; 3 to 4 years are required for most sites at the 10-percent error level. Compared to the other ecoregions, the North Central Appalachian ecoregion had a larger range of background variability, which is reflected in the range of 3-5 years of sampling needed to detect a change of 20-percent at the 20 percent error level and 3-10 years at the 10-percent error level.

The Ridge and Valley ecoregion also had high background variability for specific conductance. For three of the five sites evaluated, 3 years of monthly specific conductance data are enough to detect a 20-percent change from median at the 20-percent error level, but the remaining 2 sites require roughly 5 to 6 years to detect change and up to 11 years for 10-percent error levels.

Table 10 shows differences in the ecoregions with respect to total barium. For monitoring sites in all 3 ecoregions, 3 years of monthly total barium data are usually the minimum to detect whether a 20-percent change from the median has occurred at the 10- and 20-percent error levels; the site in the Northern Alleghany Plateau ecoregion required 5 years at the 10-percent error level.

There were two sites that required substantially more time to detect change in the barium data set, both of which were located in the Ridge and Valley ecoregion. Sinking Run near Spruce Creek required 9 years of monthly samples to detect a 20-percent change at the 10-percent error level. Swatara Creek showed that thousands of years would be required to detect a change for total barium. The sample that caused the high background variability, leading to the result of thousands of years to detect change, was collected during a storm event—other samples were obtained during regular monthly monitoring. Total barium is often bound to sediments, and it is possible that storm events could cause greater variability for total barium as such events liberate bedload sediments during high-flow events. Alternatively, a higher frequency monitoring strategy that more accurately characterizes total barium concentrations during high-flow events may facilitate identification of a statistically significant trend in total barium concentrations over a shorter time frame (Spooner et al., 2011).

If a 20-percent error level is acceptable to decision makers, then 3 years of monthly monitoring is the minimum duration of monitoring needed to determine whether there has been a 20-percent change in the median specific conductance for the Northern Alleghany Plateau ecoregion in the Susquehanna River Basin. A minimum of 5 years of monthly monitoring may be needed to detect that change in the North Central Appalachians and Ridge and Valley ecoregions. If a 10-percent error rate is desired, substantially longer data records may be needed in the North Central Appalachians and Ridge and Valley ecosystems. For total barium, for both 10- and 20-percent error levels, 3 years of monthly monitoring is the minimum monitoring duration needed to determine whether there has been a 20-percent change in median barium concentrations for most sites. In the Ridge and Valley ecoregion for total barium, 5 years of monthly sampling might be needed to detect 20-percent change in total barium concentrations. Dissolved barium may be a more reliable measure of change than total barium, given that variability of total barium may be greater due to storm-related mobilization of sediments, but total barium data were more plentiful in the Susquehanna data set. If the sampling frequency at a monitoring site is less frequent than monthly, a longer period of record will be needed to collect the minimum number of samples to detect a change in water quality.

The objective for this monitoring design is to determine whether cumulative effects of shale gas development activities change surface-water quality over the long term. Therefore the minimum of 3 years of monthly monitoring is the minimum period of record needed to detect water-quality change. Ongoing, long-term monitoring of 10 years or longer is needed to detect whether water-quality change related to shale gas development is occurring as development continues in the Marcellus Shale. To be able to detect water-quality change, monitoring sites must have data collected after shale gas development began in the Susquehanna River Basin; 2007 is considered to be the start date for purposes of this report. The most useful monitoring sites will have baseline data pre-dating shale gas development and a current, uninterrupted monitoring record.

Table 11 summarizes the surface-water data needed to detect whether water quality is changing in the Susquehanna River Basin as a result of cumulative effects of shale gas development.

Table 11. Summary of surface-water data needed to determine whether cumulative shale gas development activities in the Susquehanna River Basin are resulting in surface-water quality change.

<table>
<tr><th>Criteria</th><th>Surface-water data needed</th></tr>
<tr><td>Monitoring parameters</td><td>• Suite of priority surface-water parameters from Table 3 and streamflow at each monitoring site</td></tr>
<tr><td>Sampling frequency</td><td>• Monthly</td></tr>
<tr><td>Locations of monitoring sites</td><td>• Monitoring sites in each of the ecoregions with active or predicted HVHF activity, including:
o Northern Allegheny Plateau,
o North Central Appalachians,
o Central Appalachians, and
o Ridge and Valley.</td></tr>
<tr><td>Watershed characteristics</td><td>• Watersheds smaller than 70 square miles.
• Medium and high density and reference watersheds:
o Watersheds with greater than 0.5 HVHF wells per square mile, and
o Watersheds with 0 HVHF wells per square mile and no significant shale gas development expected.
• Watersheds larger than 70 square miles that offer opportunities for nested monitoring (i.e. one or more small watersheds that are being monitored for change are nested within the larger watershed)</td></tr>
<tr><td>Number of monitoring sites</td><td>• Minimum of 1 monitoring site in a high density watershed per ecoregion.
• Minimum 1 reference monitoring site per ecoregion.</td></tr>
<tr><td>Duration and timing of monitoring</td><td>• At least 36 samples collected at monthly or longer intervals over 3-4 years including data collected after shale gas development (post-2007)
• Minimum duration of monitoring to detect water-quality change varies by ecoregion
• Ideal monitoring sites will have:
o Data collected before shale gas development (pre- 2007),
o An uninterrupted data record,
o Current/ongoing data collection (2009 or later), and
o Plans for long-term monitoring.</td></tr>
</table>

7. Surface-Water Data Availability

This chapter explores the availability of surface-water data that meet the monitoring criteria identified in Table 11.

7.1 Surface Water-Quality Data in the Susquehanna Data Set

Table 12 presents the number of monitoring sites in the Susquehanna data set with data for the priority surface-water parameters, for an overview of relevant data available in the Susquehanna data set. Table 13 lists the individual monitoring sites that meet the criteria in Table 11 for specific conductance or total barium, and Table 14 presents the total number of monitoring sites in the Susquehanna data set that meet additive criteria from Table 11. Table 14 reveals that there are no monitoring sites in the Susquehanna data set that meet all the monitoring criteria in Table 11.

As listed in Table 12, a total of 10 barium monitoring sites (8 sites with data for total barium and 2 with data for dissolved barium) and 22 specific conductance monitoring sites have a drainage area less than 70 square miles, have at least 36 samples for the focus parameter paired with streamflow, and have collected data as recently as 2009, thus meeting most of the criteria identified in Table 11 for those two parameters. Data were collected at all of these monitoring sites prior to 2007, which predates the beginning of shale gas development in the basin. Of these sites, one barium station and five specific conductance sites have not been active monitoring sites after 2007. Only two of the specific conductance sites, and no barium sites, are located in a watershed with greater than 0.5 HVHF well permits per square mile. The locations of these sites are shown in Figure 22. In some cases the HVHF well density shown in Figure 22 is different than the HVHF well density in Table 13 because the HVHF well density shown in Figure 22 is presented at the HUC12 watershed scale, whereas Table 13 lists the calculated HVHF well density for the monitored watershed given the specific location of the monitoring site, which might be larger or smaller than the nearest HUC12 watershed.

Table 12 provides more information regarding availability of specific parameter data. The field parameters of dissolved oxygen, pH, and water temperature were collected at 23 sites both before 2007 and in 2009 or later, about the same number of sites that met minimum criteria for specific conductance. Total barium samples meeting minimum criteria were collected at 10 monitoring sites. Samples collected at about 10 monitoring sites met the minimum criteria for alkalinity, total calcium, dissolved chloride, total magnesium, total nitrate, total phosphorus, dissolved and total sulfate, total dissolved solids, and total suspended solids. The remaining parameters have limited to no availability, including lithium, strontium, gross alpha, gross beta, potassium, all radium isotopes, bromide, suspended sediment concentration, total organic carbon, and uranium. Consequently, the data collected at the monitoring sites in the Susquehanna data set are not sufficient to determine whether cumulative shale gas development activities in the Susquehanna River Basin are resulting in surface-water quality change. The monitoring sites are not located in the right locations and data are not available for the right parameters.

Table 12. Number of surface-water monitoring sites in the Susquehanna data set that nearly meet case study criteria for the suite of priority surface-water parameters.
[Abbreviations: NO_3, nitrate; N, nitrogen; ROE, residue on evaporation; sum, sum of constituents; P, phosphorus; PO_4, phosphate]

Water-quality parameter[1]	Number of sites with 36 or more records paired with streamflow in watersheds of 70 square miles or less	Records	Number of sites by sample date				
			Before 2007	2007 and later	2009 and later	Both before 2007 and 2007 or later	Both before 2007 and 2009 or later
Alkalinity	22	1,163	21	20	13	19	12
Barium, dissolved	2	100	2	0	0	0	0
Barium, total	11	664	11	10	10	10	10
Calcium, dissolved	24	1,582	24	19	1	19	1
Calcium, total	28	1,632	28	28	10	28	10
Chloride, dissolved	23	1,486	23	18	10	18	10
Chloride, total	1	58	1	1	1	1	1
Dissolved oxygen	48	2,841	47	42	24	41	23
Gross beta	1	64	1	0	0	0	0
Lithium, dissolved	1	52	1	0	0	0	0
Magnesium, dissolved	24	1,583	24	19	1	19	1
Magnesium, total	28	1,644	28	28	10	28	10
Nitrate, dissolved, as NO_3	2	97	2	0	0	0	0
Nitrate, total, as N	13	764	13	11	11	11	11
Nitrate, total, as NO_3	5	283	5	5	5	5	5
Nitrite plus nitrate, dissolved, as N	3	189	3	0	0	0	0
Nitrite plus nitrate, total, as N	3	207	3	0	0	0	0
pH	53	3,229	52	42	24	41	23
Phosphorus, total, as P	14	981	14	11	11	11	11
Phosphorus, total, as PO_4	4	165	4	2	2	2	2

Water-quality parameter[1]	Number of sites with 36 or more records paired with streamflow in watersheds of 70 square miles or less	Records	Number of sites by sample date				
			Before 2007	2007 and later	2009 and later	Both before 2007 and 2007 or later	Both before 2007 and 2009 or later
Potassium, dissolved	23	1,326	23	17	0	17	0
Potassium, total	16	787	16	16	0	16	0
Radium-226	1	36	1	0	0	0	0
Sodium, dissolved	24	1,535	24	18	0	18	0
Sodium, total	18	1,007	18	18	0	18	0
Specific conductance	51	3,114	51	40	22	40	22
Strontium, dissolved	2	102	2	0	0	0	0
Sulfate, dissolved	36	2,192	36	27	10	27	10
Sulfate, total	10	416	9	10	10	9	9
Total dissolved solids, ROE	14	1,021	14	10	10	10	10
Total dissolved solids, sum	4	326	4	0	0	0	0
Total suspended solids	14	831	14	10	10	10	10
Turbidity	3	132	3	0	0	0	0
Water temperature	53	3,200	52	42	24	41	23

[1]Only those parameters are shown for which some data were available. Bromide, gross-alpha, radium-224 and radium-228, suspended sediment concentration, total organic carbon, and uranium had no sites meeting the study criteria.

Table 13. Surface-water monitoring sites in the Susquehanna data set in watersheds 70 square miles and smaller with 36 or more monthly specific conductance or total barium records, with streamflow, and data spanning pre- and post-shale gas development periods and data collected as recently as 2009. The Map IDs correspond to the sites shown in Figure 22.

[Abbreviations: mi^2, square miles; ecoregion, U.S. Environmental Protection Agency Level III Ecoregion; SC, Specific Conductance; Ba, barium]

Map ID	Monitoring site identification number	Site name	Water-quality parameter	Drainage area (mi^2)	Density of HVHF wells in monitored watershed (wells/mi^2)	Period of record in Susquehanna data set
		Northern Allegheny Plateau ecoregion				
1	BNTY000.9	Bentley Creek at Wellsburg, NY	SC	54.2	0.3	1988-2009
2	nySEEL010.3	Seeley Creek at State Line	SC	27.0	1.6	1986-2009
3	nyCASC001.6	Cascade Creek at State Line	SC	11.4	0	1988-2009
4	paLSNK007.6	Little Snake Creek near Brackney, PA	SC	7.7	0.5	1988-2009
5	paTRUP004.5	Troups Creek at Austinburg, PA	SC	53.7	0	1988-2009
		North Central Appalachians ecoregion				
6	01548476	Cedar Run above Mine Hole Run near Cedar Run, PA	SC, Ba	26.3	0	2005-10
7	01545600[1]	Young Womans Creek near Renovo, PA	SC, Ba[2]	46.2	0	1967-2012
8	01538709	West Branch Fishing Creek near Elk Grove, PA	SC, Ba	20.2	0	2005-10
		Central Appalachians ecoregion				
9	0154133098	Killbuck Run above Glendale Lake near St. Augustine	SC, Ba	6.61	0	2006-10
		Ridge and Valley ecoregion				
10	01548077	Cedar Run at Cedar Springs near Mill Hall, PA	SC, Ba	14.8	0	2005-10
11	CCPASEC_1216	Lick Run at Howard, PA	SC	11.4	0	2004-12
12	paCCPASEC_112	Little Fishing Creek-Bridge	SC	24.7	0	2003-13
13	paCCPASEC_1123	Little Fishing Creek-Tice	SC	16.1	0	2003-13
14	01554665	Elk Creek at Spring Bank near Millheim, PA	SC, Ba	31.0	0	2005-10
15	paCCPASEC_1017	Galbraith Gap Run	SC	5.14	0	2004-12
16	paCCPASEC_1018	Slab Cabin Run	SC	1.12	0	2004-12
17	01557990	Sinking Run near Spruce Creek, PA	SC, Ba	28.3	0	2005-10
18	pa1244	Wiconisco Creek above Porter Tunnel discharge[3]	SC	2.00	0	2002-12
19	pa1246	Wiconisco Creek below Porter Tunnel[3]	SC	2.10	0	2005-12
20	01566005	East Licking Creek above reservoir near Lewistown, PA	SC, Ba	21.9	0	2005-10

Map ID	Monitoring site identification number	Site name	Water-quality parameter	Drainage area (mi^2)	Density of HVHF wells in monitored watershed ($wells/mi^2$)	Period of record in Susquehanna data set
21	0155979602	Bobs Creek below Wallacks Branch at Pavia, PA	SC, Ba	22.1	0.1	2006-12
22	01569195	Conodoguinet Creek above Reservoir near Roxbury, PA	SC, Ba	27.2	0	2005-10

[1]USGS operates a continuous-record streamflow gaging station at this site.

[2]Total barium results at Young Womans Creek prior to 2007 are of limited use because of elevated reporting levels.

[3]Although they meet criteria for this table, the Porter Tunnel monitoring sites are not appropriate reference monitoring sites because they are designed to measure discharge from the Porter Tunnel.

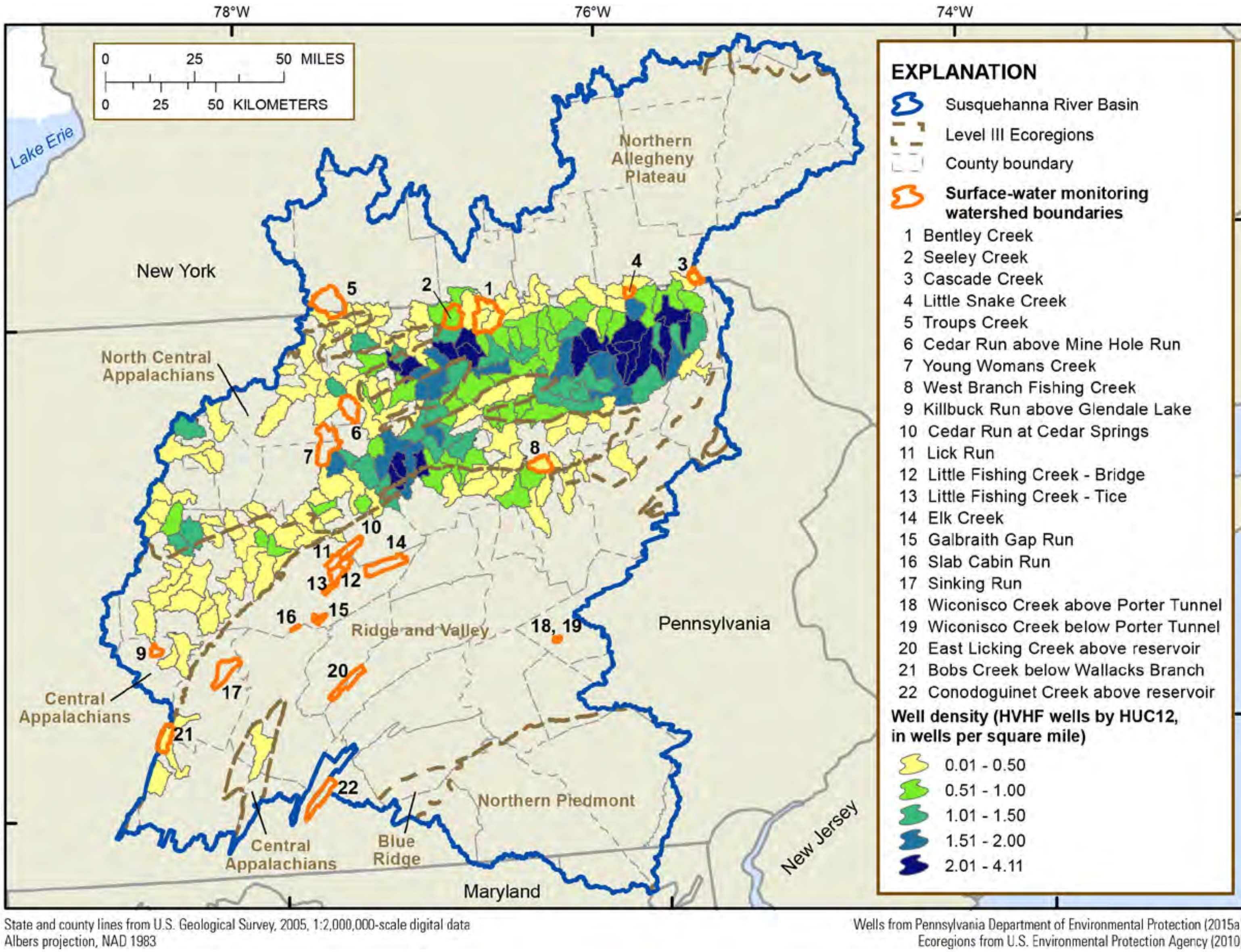


Figure 22. Watershed boundaries for surface-water monitoring sites in the Susquehanna data set that meet case study criteria for specific conductance or barium. Site labels are defined in Table 13.
[Abbreviations: HVHF, high-volume hydraulic fracturing]

Table 14. Number of surface-water monitoring sites in the Susquehanna data set that meet criteria for barium and specific conductance.

Criteria	Number of stations meeting criteria for each parameter		
	Barium, total	Barium, dissolved	Specific conductance
At least 36 records, matching streamflow, drainage area smaller than approximately 70 square miles	11	2	51
As above plus: monitoring precedes 2007	10	2	40
As above plus: most recent date of monitoring is 2009 or later	10	0	22
As above plus: located in a watershed with more than 0.5 HVHF well permits per square mile	0	0	2
As above plus: data available for most priority surface-water parameters	0	0	0

7.2 Susquehanna River Basin Commission's Remote Water Quality Monitoring Network

The SRBC deployed an extensive network of real-time monitors, called the Remote Water Quality Monitoring Network (RWQMN), starting in January 2010 in response to shale gas development activities in the basin (Susquehanna River Basin Commission, 2012a; Susquehanna River Basin Commission, 2015). This monitoring program is collecting appropriate water data for detecting water-quality change related to shale gas development in appropriate locations. This program uses real-time monitoring of field parameters at each monitoring site (pH, temperature, specific conductance, dissolved oxygen and turbidity), and SRBC staff collects water samples at each site four times a year for laboratory analysis including inorganics and metals. The SRBC analyzes for all the parameters identified in Table 3 with the exceptions of dissolved barium, uranium, and radium (Susquehanna River Basin Commission, 2012b). Corrected continuous data are available for all SRBC RWQMN sites through December 2014. The minimum number of records is available for the continuous monitoring non-specific field parameters, but the required amount of data for discrete sample parameters, which are analyzed quarterly, is not available yet.

Figure 23 shows the watersheds monitored by the SRBC as part of the RWQMN. As of July 2011 there were 37 monitoring sites in the watershed; the number of monitoring sites grew to 58 by October 2013. Minor modifications have been made to the network since 2013; one monitoring site was discontinued and one was added. Streamflow is measured during routine quarterly sampling, if conditions allow. The SRBC is developing stage-streamflow rating curves that can be applied to the stage data to derive estimates of streamflow. This network includes 32 monitoring sites in watersheds with HVHF wells. The SRBC monitoring sites in both reference and HVHF well watersheds can have impairments that are not related to shale gas development. This makes having the full suite of surface-water monitoring parameters available to interpret water-quality change relative to different sources and stressors that much more important.

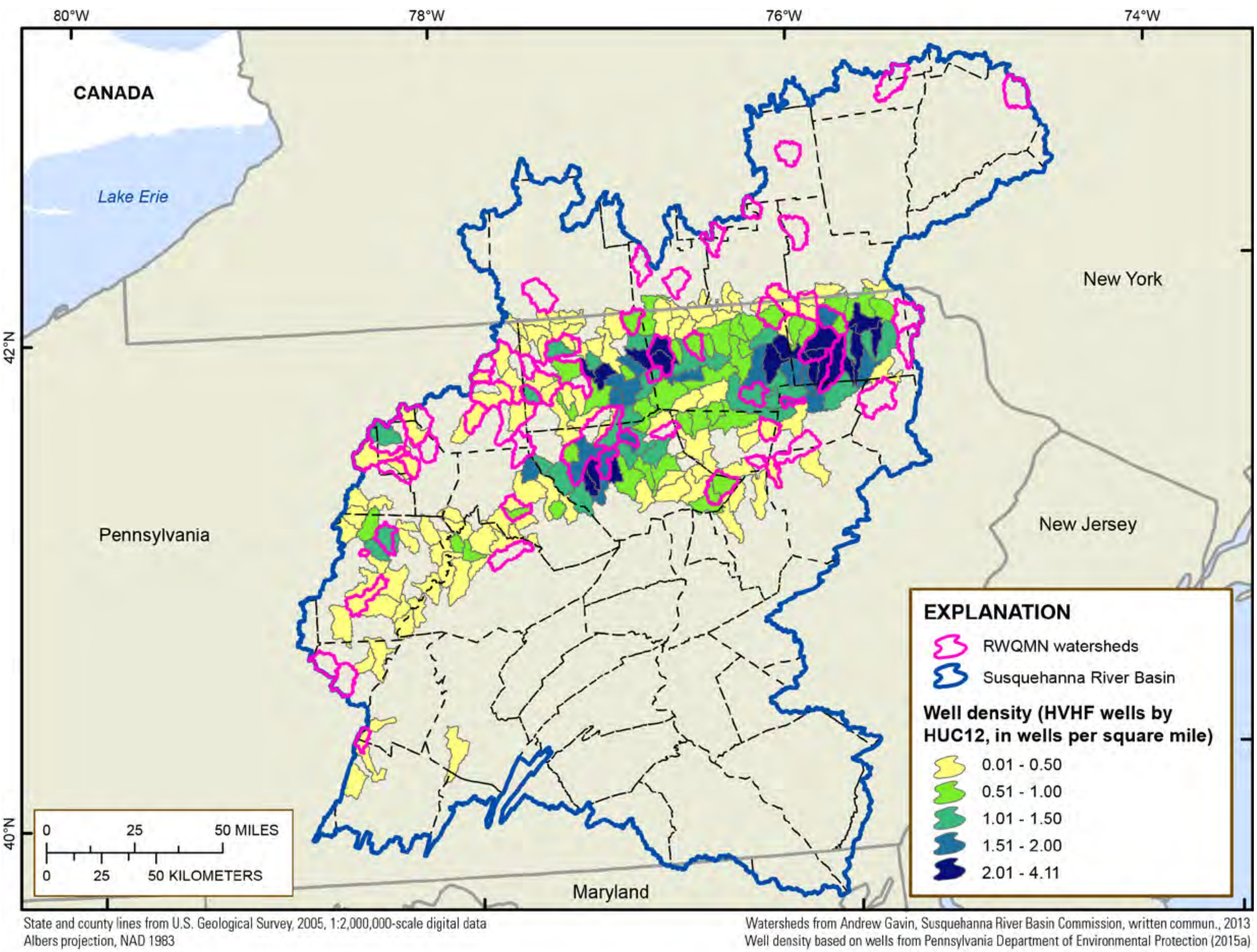


Figure 23. Watersheds monitored by the Susquehanna River Basin Commission as part of their Remote Water Quality Monitoring Network as of 2015, and density of HVHF wells by HUC12 watershed.

[Abbreviations: RWQMN, Remote Water Quality Monitoring Network; HVHF, high-volume hydraulic fracturing; HUC12, 12-digit hydrologic unit code boundaries]

Monitoring sites shown in Figure 23 have substantially fewer records for the parameters that are not measured in real-time, including barium. The SRBC typically collects barium data on a quarterly basis, which means that most sites have 20 or fewer barium samples. Substantially more samples will be needed to detect whether barium is changing at a given power and significance for these sites. Assuming that 20 samples have been collected and analyzed, at least 4 additional years of quarterly samples would be needed to meet the minimum criterion of 36 samples.

7.3 Pennsylvania Department of Environmental Protection Water Quality Network

Adjustments were made to the Susquehanna River Basin portion of the PADEP fixed water quality network (WQN) in the early 2000s as the PADEP cooperated with the SRBC and the USGS to implement a non-tidal nutrient monitoring program in the Chesapeake Bay watershed. Many of the Susquehanna River Basin WQN sites were included in this new monitoring initiative. Several inactive WQN sites were reactivated, several were moved, and all of the selected sites were targeted for storm flow sampling (Pennsylvania Department of Environmental Protection, 2005). Figure 24 shows the WQN in the Susquehanna River Basin, which includes "standard sites" which are monitored 6 times a year, "reference sites" in relatively undisturbed areas and "loading sites" that are monitored 12 times a year to assess nutrient loading to the Chesapeake Bay. Like the SRBC RWQMN program, the PADEP WQN collects appropriate water data for detecting water-quality change related to shale gas development in appropriate locations, but the minimum number of records for detecting water-quality change is not yet available.

In addition to the sites shown in Figure 24, the active streamgages were examined (Figure 25 and Table 15) to determine if there are available streamgages that could be used to assist in filling data gaps. A secondary purpose for examining streamgages was to identify water-quality monitoring sites co-located with streamgages in watersheds 70 square miles or larger in areas with a high density of HVHF wells that could serve as nested monitoring sites.

There are several PADEP WQN monitoring sites and data records that do not meet the case study criteria (Table 11) but have potentially useful water-quality data. Some monitoring sites are in watersheds larger than 70 square miles, have a high density of HVHF wells, and may provide nested monitoring opportunities. Parameters measured at the PADEP WQN monitoring sites include many of the priority parameters listed in Table 3 (U.S. Geological Survey, 2014b and 2014c; Pennsylvania Department of Environmental Protection, 2012b). However, these sites lack some critical parameters including dissolved barium, gross alpha and beta, lithium, potassium, suspended sediment concentration, and total organic carbon. The suite of priority surface-water parameters is even more important for these larger watersheds that are more likely to include multiple sources of potential contaminants. Many of the PADEP WQN sites in Figure 24 monitor watersheds that do not have a high density of HVHF wells, or are larger than 200 square miles. The most relevant monitoring sites in the PADEP WQN are described in greater detail in Chapter 8.

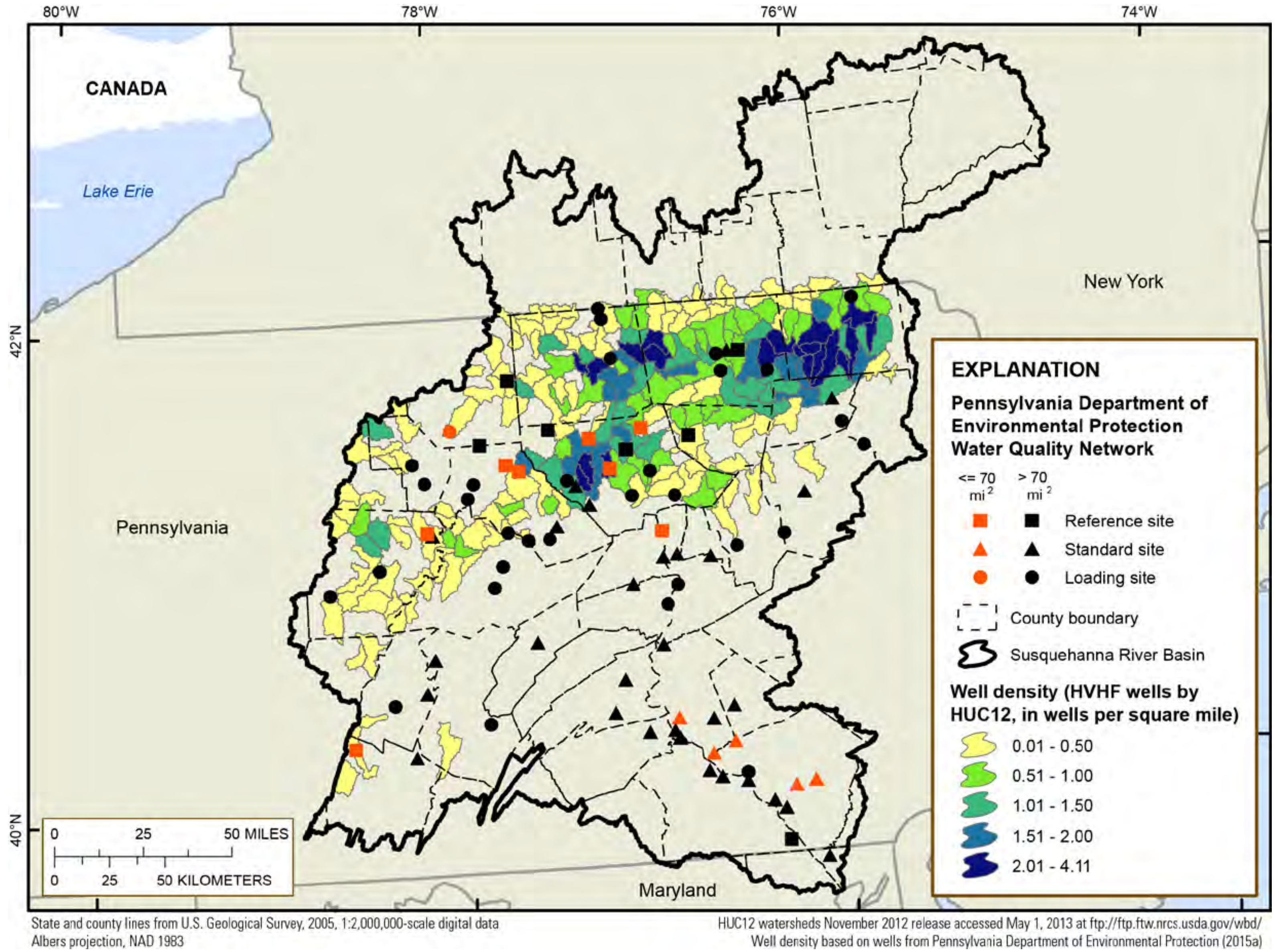
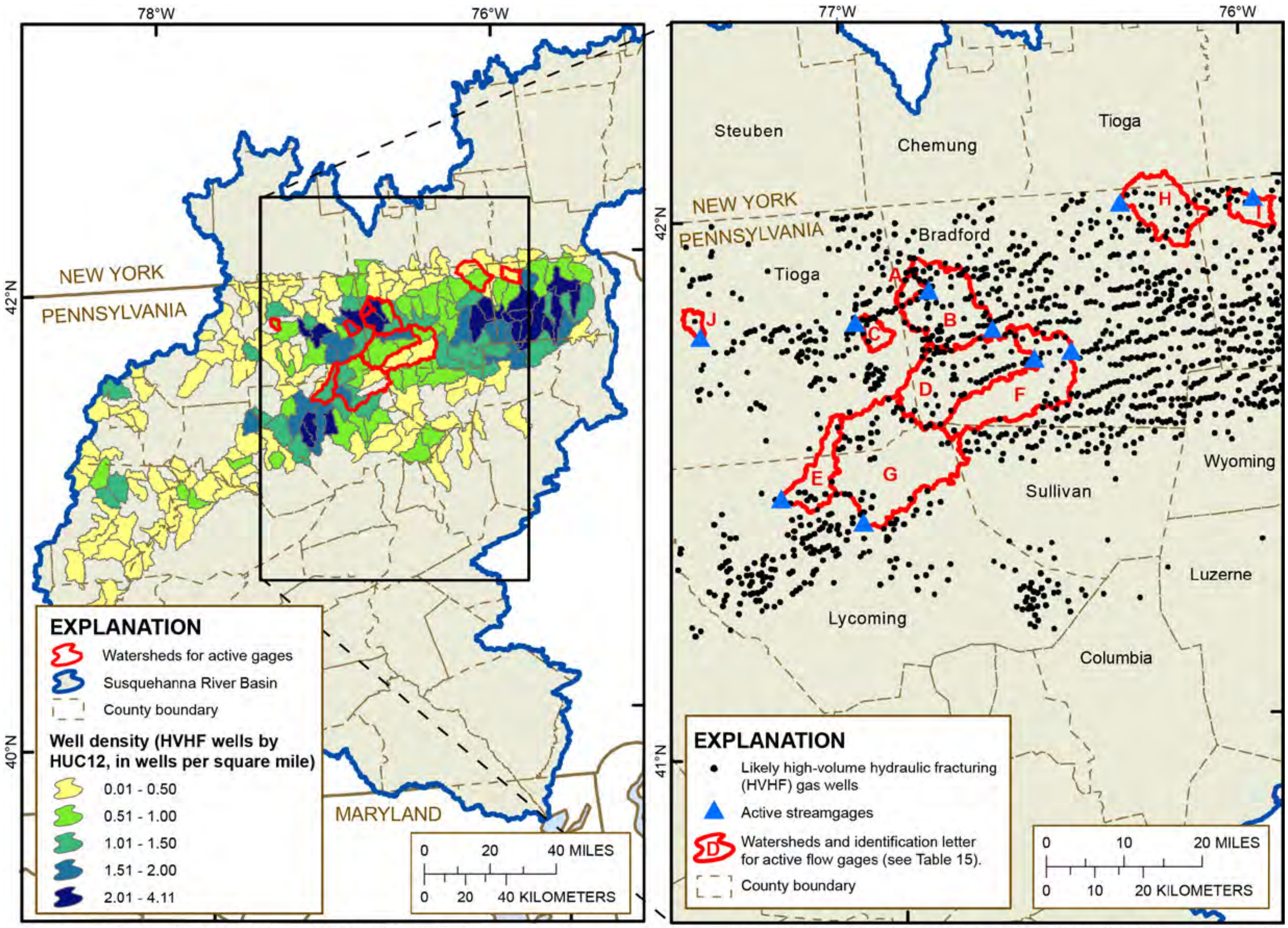


Figure 24. Pennsylvania Department of Environmental Protection Fixed Water Quality Network in the Susquehanna River Basin as of July 2015 and the density of HVHF wells by HUC12 watershed.

[Abbreviations: HVHF, high-volume hydraulic fracturing; HUC12, 12-digit hydrologic unit code boundaries]



State and county lines from U.S. Geological Survey, 2005, 1:2,000,000-scale digital data
Albers projection, NAD 1983

HUC12 watersheds November 2012 release accessed May 1, 2013 at ftp://ftp.ftw.nrcs.usda.gov/wbd/
Wells and well density based on wells from Pennsylvania Department of Environmental Protection (2015a)

Figure 25. Watersheds with active streamgages that are located within areas with a high density of HVHF wells. Site labels are defined in Table 15.
[Abbreviations: HVHF, high-volume hydraulic fracturing; HUC12, 12-digit hydrologic unit code boundaries]

Table 15. Active streamgages in areas with a high density of HVHF wells in the Susquehanna River Basin. The Map IDs correspond to the sites shown in Figure 25.
[Abbreviations: HVHF, high-volume hydraulic fracturing; mi^2, square miles; NB, North Branch; Trib., Tributary]

Map ID	Site number	Station name	Number of HVHF wells in monitored watershed	Drainage area (mi^2)	Density of HVHF wells in monitored watershed (wells/mi^2)
A	01531250	NB Sugar Creek Trib near Columbia Cross Roads, PA	36	8.83	4.1
B	01531325	Sugar Creek at West Burlington, PA	248	93.6	2.6
C	01516500	Corey Creek near Mainesburg, PA	19	12.2	1.6
D	01531908	Towanda Creek near Franklindale, PA	125	112	1.2
E	01549500	Blockhouse Creek near English Center, PA	33	37.7	0.9
F	01532000	Towanda Creek near Monroeton, PA	155	215	0.7
G	01550000	Lycoming Creek near Trout Run, PA	129	173	0.7
H	01514850	Wappasening Creek near Windham Center, PA	35	56.9	0.6
I	01513550	Choconut Creek near Choconut, PA	6	18.5	0.3
J	01548303	Straight Run, Site 1, at Marsh Creek, PA	1	6.48	0.2

7.4 Surface-Water Data Available

In summary, there are no small watershed monitoring sites collecting data for all, or most, of the priority surface-water parameters that are located in watersheds with a high density of HVHF wells with enough samples to identify whether water quality is changing in response to shale gas development activities. The sites that do have enough records in the areas with a high density of HVHF wells only have data for non-specific parameters. For these parameters, this quantity of data could potentially show areas in the Susquehanna River Basin in which further investigation is warranted, but they do not provide enough information to answer the case-study policy question.

There are a number of SRBC RWQMN and PADEP WQN sites collecting most of the needed parameters and in appropriate locations, but monitoring at these sites began only recently and not enough data has been collected yet to be able to detect change in these parameters. Modifications to these existing monitoring programs will likely provide the shortest path to obtaining the water data needed to answer the case-study policy question.

7.5 Surface-Water Data Consistency and Quality Assurance

Quality assuring data from multiple agencies and existing data sources is difficult, time-consuming, and expensive—but has to be done to prevent the use of unreliable data that could affect the statistical analysis of water-quality data. Water-quality data with results produced by different analytical methods, each with different reporting levels, can limit or preclude one's ability to perform valid statistical analysis.

The monitoring sites with sufficient data for detecting changes in specific conductance and barium concentrations over time (Table 13) were sampled primarily by the USGS and SRBC, and data from five sites were collected by the Centre County Pennsylvania Senior Environmental Corps (in collaboration with other groups). Additional relevant water-quality data are available from the PADEP WQN. Data quality assurance procedures are discussed for each agency or group below.

The USGS follows a set of documented methods, data analysis, and data quality control assurance procedures for the collection of streamflow and water-quality data. Overall, the USGS stresses objective and replicable data collection procedures that are transparent and peer reviewed. Data that are collected for publication must be documented to describe the methods or techniques used to collect, process, and analyze the data. For an overview of methods used to collect, analyze, and quality assure streamflow data see Rantz et al. (1982). For detailed descriptions of methods for collecting streamflow data, see individual chapters of USGS Techniques of Water Resources Investigations, Book 3: Applications of Hydraulics, Section A: Surface-Water Techniques, which are available at *http://pubs.usgs.gov/twri/index-last.html*. Individual chapters of Techniques of Water Resources Investigations may be superseded by updated and new chapters of USGS Techniques and Methods, which are available at *http://pubs.usgs.gov/tm/*. For information on methods used to collect and quality assure water-quality samples, see the USGS National Field Manual available at *http://water.usgs.gov/owq/FieldManual/*. For documentation of specific USGS laboratory or field analytical methods and related quality assurance/quality control procedures see *http://water.usgs.gov/owq/methods.html*.

The SRBC also follows documented data quality assurance procedures (Susquehanna River Basin Commission, 2012b). Samples are tested by accredited laboratories with documented quality assurance procedures. The SRBC submits duplicate samples at least once per 10 samples. For the RWQMN sites, the SRBC measures streamflow whenever discrete water-quality samples are collected—usually quarterly. All staff members who measure streamflow are required to participate in computer-assisted training provided by USGS, undergo 6 months of on-the-job training with an experienced staff member, and complete an annual field check (Susquehanna River Basin Commission, 2012b). All continuous data for field parameters, including stage data, are corrected quarterly using field verification procedures.

A majority of the SRBC monitoring sites have unvented pressure transducers that do not correct for changes in barometric pressure, which adversely affects the quality of the stage data at those sites. This may have an impact on the number of samples required to detect statistically significant change for field parameters for unvented sites. Fourteen sites have vented pressure transducers, and at those sites flow is measured more frequently than quarterly. The SRBC is developing rating curves relating streamflow to stage at these stations, and the SRBC is measuring flow more frequently than quarterly in order to develop these rating curves. The smaller subset of monitoring sites with the vented transducers will have higher data quality.

The PADEP WQN also follows a set of documented methods, data analysis, and data quality control assurance procedures for the collection of water-quality data. A quality assurance project plan describes the procedures used to implement the PADEP WQN in both the field and laboratory (Pennsylvania Department of Environmental Protection, 2014). One sample "blank" is submitted bimonthly by each field office involved in sample collection and duplicate or "split samples" are collected at the rate of one in every 20 regular WQN samples.

The Centre County Pennsylvania Senior Environmental Corps is a voluntary monitoring organization that has collected data that meet some of the case study criteria. The group strives for accurate and precise information and implements a water-quality control program. The data are intended to be backed by a level of confidence suitable for studies and decision making. The group’s most recent water-quality control report is available online and was published in March of 2012; the report discussed data error and performance standards that were met by the group for specific conductance (Centre County Pennsylvania Senior Environmental Corps, 2014).

7.6 Surface-Water Data Usability

Water-quality monitoring programs are usually designed to meet a stated objective or follow a historical precedent. Data collected for one monitoring objective may not be directly applicable to another objective, due to the location of monitoring sites, frequency of monitoring, parameters measured, and sample fraction analyzed. The surface-water quality records identified through this case study, summarized in Table 4 and Figure 12, were generated by 35 organizations that collected surface-water data in the Susquehanna River Basin. Significant time and effort were required to locate, obtain, and format water-quality records from multiple organizations that use different sampling plans and data documentation practices.

The Water Quality Portal (National Water Quality Monitoring Council, 2014a) is a cooperative service that integrates publicly available water-quality data from NWIS, STORET, and the U.S. Department of Agriculture (USDA) Agricultural Research Service (ARS) Sustaining The Earth’s Watersheds - Agricultural Research Database System (STEWARDS), and includes data collected by more than 400 State, Tribal, and local organizations. The Water Quality Portal provides an interface that allows a user to download water-quality data from these systems in a consistent format. Data collected at only 19 percent of the surface-water monitoring sites identified through this case study in the Susquehanna River Basin were available through the Water Quality Portal, but 85 percent of the water-quality data records are available through the Portal. This finding indicates that the monitoring sites from the Susquehanna data set that are available through the Water Quality Portal are sites with longer data records and are monitoring sites that are more likely to have the data necessary to identify a water-quality trend. The data that are missing from the Water Quality Portal are primarily collected by volunteer organizations, local governments, and academia, although some State and Federal data are also not available through the Water Quality Portal (e.g., the Pennsylvania Fish and Boat Commission, and the US Army Corps of Engineers).

Finally, for water data to be useful for addressing the policy question, they must be compatible in terms of sampling plans and protocols, analysis, and interpretation. Three primary agencies collect the water data that most nearly met the criteria for this case study: the PADEP, USGS and SRBC. There are differences in sampling frequency, sampling protocol, and sample analysis between the respective agencies that would need to be addressed to compare water-quality trends over time at these monitoring sites.

8. New Surface-Water Data Needed to Answer the Case-Study Policy Question and Associated Costs

As discussed in Chapter 7, no surface-water monitoring sites in the Susquehanna River Basin have data for the entire suite of priority surface-water parameters in watersheds with a medium or high density of HVHF wells, and for those sites that collect a substantial subset of the parameters, not enough data records are available to meet the case study criteria identified in Table 11. This chapter identifies the water data needed to fill the surface-water data gaps for answering the case-study policy question and provides rough cost estimates for collecting the data.

8.1 Surface-Water Data Needs

Candidate sites for new or increased monitoring to answer the case-study policy question include monitoring sites from the Susquehanna data set with greater than 36 unique records for one or more priority surface-water parameters, active SRBC RWQMN monitoring sites, and sites from the PADEP WQN monitoring program. Table 16 lists candidate reference watershed monitoring sites, and Table 17 lists candidate sites in watersheds with HVHF wells. Figure 26 (A) and (B) show the locations of these candidate monitoring sites within the Susquehanna River Basin.

As summarized in Table 11, monitoring sites are needed for medium or high density and reference watersheds in each of the four ecoregions with ongoing or future shale gas development. In addition to these minimum recommendations, additional monitoring sites in the areas of the densest shale gas development would improve the ability to answer the case-study policy question on a policy relevant time scale and this additional monitoring may be implemented through modifications at existing monitoring sites. There are many reasons to select individual monitoring sites from these candidate lists for increased monitoring to answer the case-study policy question. The following list includes several monitoring site characteristics that could make an individual monitoring site a priority site for increased monitoring:

- History of water-quality data available at the site
- Variety of parameter data available at the monitoring site
- Density of HVHF wells in the monitored watershed
- Availability of existing continuous monitors or streamgages at the monitoring site
- Availability of nested monitoring sites
- Current or long-term funding source available for the monitoring site
- History of cooperation with shale gas development companies that are active within the monitored watershed

Table 16. Candidate reference sites for surface-water monitoring in ecoregions with shale gas development activity in the Susquehanna River Basin that meet most monitoring criteria for determining whether cumulative shale gas development activities in the Susquehanna River Basin are resulting in surface-water quality change. The Map IDs correspond to the sites shown in Figure 26(A).

[**Abbreviations**: mi^2, square miles; ecoregion, U.S. Environmental Protection Agency Level III Ecoregion; SRBC, Susquehanna River Basin Commission; PADEP WQN, Pennsylvania Department of Environmental Protection Water Quality Network; USGS, U.S. Geological Survey; CCPASEC, Centre County Pennsylvania Senior Environmental Corps]

Map ID	Monitoring organization	Site name	Drainage area (mi^2)
		Northern Allegheny Plateau ecoregion	
R1	SRBC	South Branch Tunkhannock Creek near La Plume, PA	70.5
R2	SRBC	Troups Creek at Austinburg, PA	53.7
R3	SRBC	Upper Tuscarora Creek near Woodhull, NY	53.0
R4	SRBC	Sangerfield River near Poolville, NY	52.5
R5	SRBC	Starrucca Creek near Stevens Point, PA	52.0
R6	SRBC	Cherry Valley Creek near Middlefield, NY	51.0
R7	SRBC	Nanticoke Creek near Maine, NY	48.0
R8	SRBC	Lackawanna River near Forest City, PA	38.0
R9	SRBC	Trout Brook near McGraw, NY	35.7
R10	SRBC	Baldwin Creek near Lowman, NY	35.0
R11	SRBC	Sing Sing Creek near Big Flats, NY	35.0
R12	SRBC	Catatonk Creek near Spencer, NY	30.0
R13	SRBC	West Branch Owego Creek near Speedsville, NW	24.4
R14	SRBC	Cascade Creek at State Line	11.4
		North Central Appalachians ecoregion	
R15	PADEP WQN	Mosquito Creek near Karthaus, PA[2]	64.0
R16	SRBC	Bowman Creek near Bowman Creek, PA[1]	54.0
R17	PADEP WQN and USGS	Young Womans Creek near Renovo, PA	46.2
R18	SRBC	Young Womans Creek near North Bend, PA	41.0
R19	SRBC	Hunts Run near Cameron, PA	30.7
R20	SRBC	Loyalsock Creek near Ringdale, PA	27.0
R21	PADEP WQN	Cedar Run above Mine Hole Run near Cedar Run, PA	26.3
R22	SRBC	Long Run near Gaines, PA[1]	21.0
R23	SRBC	Pleasant Stream near Marsh Hill, PA	20.6
R24	PADEP WQN	West Branch Fishing Creek near Elk Grove, PA	20.2
R25	SRBC	Pine Creek (Upper) near Telescope, PA[1]	18.6
R26	SRBC	East Branch Fishing Creek near Jamison City, PA	12.6
		Central Appalachians ecoregion	
R27	SRBC	Chest Creek near Patton, PA	44.5
R28	SRBC	West Branch Susquehanna near Cherry Tree, PA	36.2

Map ID	Monitoring organization	Site name	Drainage area (mi²)
R29	PADEP WQN	Killbuck Run above Glendale Lake near St. Augustine	6.61
		Ridge and Valley ecoregion	
R30	PADEP WQN	White Deer Creek near Watsontown, PA	44.2
R31	SRBC	Marsh Creek near Blanchard, PA	44.0
R32	PADEP WQN	Elk Creek at Spring Bank near Millheim, PA	31.0
R33	PADEP WQN	Sinking Run near Spruce Creek, PA	28.3
R34	PADEP WQN	Conodoguinet Creek above Reservoir near Roxbury, PA	27.2
R35	CCPASEC	Little Fishing Creek-Bridge	24.7
R36	PADEP WQN	East Licking Creek above reservoir near Lewistown, PA	21.9
R37	CCPASEC	Little Fishing Creek-Tice	16.1
R38	PADEP WQN	Cedar Run at Cedar Springs near Mill Hall, PA	14.8
R39	CCPASEC	Lick Run at Howard, PA	11.4
R40	CCPASEC	Galbraith Gap Run	5.14
R41	CCPASEC	Slab Cabin Run	1.12

[1]Watershed is located in both Northern Allegheny Plateau and North Central Appalachians ecoregions
[2]Watershed is located in both North Central Appalachians and Central Appalachians ecoregions

Table 18 presents an example set of monitoring sites where increased monitoring could be used to measure whether cumulative shale gas development activities in the Susquehanna River Basin are resulting in surface-water quality change. The factors related to their utility for answering the case-study policy question are also summarized in the table. These sites include watersheds with the greatest density of HVHF wells, sites co-located with USGS streamgages, sites with monitoring data predating shale gas development, and larger watershed monitoring sites that have smaller watershed monitoring sites nested within. Two monitoring sites are presented for each of the two ecoregions with fewer HVHF wells, and additional monitoring sites are presented for the Northern Allegheny Plateau and North Central Appalachians ecoregions, which include substantially more shale gas development than the other two ecoregions.

Table 19 summarizes the additional data that would need to be collected at the set of monitoring sites described in Table 18 to answer the case-study policy question. Monthly monitoring is recommended for each of these monitoring sites. For SRBC monitoring sites, this means increasing sampling frequency from quarterly to monthly, which would reduce the time required to detect statistically significant change from about 4 additional years of monitoring to less than 2 additional years, assuming that 36 samples are sufficient to detect change per the power analysis method described in section 6.6. Samples collected at SRBC monitoring sites should be analyzed for three additional parameters to capture the entire suite of priority surface-water parameters: dissolved barium, uranium, and radium on an as-needed basis. Streamgages should be added for each of the SRBC monitoring sites selected for increased monitoring. The monitoring frequency varies at existing PADEP WQN monitoring sites, but should be consistent at all sites with a monthly sampling frequency. Analysis of dissolved barium, gross alpha and beta, uranium, lithium, potassium, suspended sediment concentration, total organic carbon, turbidity, and radium on an as-needed basis would need to be added for the PADEP monitoring sites.

Table 17. Candidate sites for increased surface-water monitoring in watersheds with HVHF wells in the ecoregions with shale gas development activity in the Susquehanna River Basin that meet most monitoring criteria for determining whether cumulative shale gas development activities in the Susquehanna River Basin are resulting in surface-water quality change. The Map IDs correspond to the sites shown in Figure 26(B).
[**Abbreviations**: mi^2, square miles; HVHF, high-volume hydraulic fracturing; ecoregion, U.S. Environmental Protection Agency Level III Ecoregion; SRBC, Susquehanna River Basin Commission; PADEP WQN, Pennsylvania Department of Environmental Protection Water Quality Network]

Map ID	Monitoring organization	Site name	Drainage area (mi^2)	Number of HVHF wells in monitored watershed	Density of HVHF wells in monitored watershed (wells/mi^2)	Site type (density)[1]
		Northern Allegheny Plateau ecoregion				
H1	USGS streamgage	NB Sugar Creek Trib near columbia Cross Roads	8.83	36	4.08	High
H2	SRBC	Sugar Creek near Troy, PA	56.2	209	3.72	High
H3	SRBC	Meshoppen Creek near Kaiserville, PA	51.8	157	3.03	High
H4	SRBC	East Branch Wyalusing Creek near Lawton, PA	69.3	166	2.40	High
H5	SRBC	Little Mehoopany Creek near North Mehoopany, PA	10.8	24	2.22	High
M1	SRBC	Seeley Creek at State Line	27.0	44	1.63	Medium
M2	USGS streamgage	Corey Creek near Mainesburg, PA[2]	12.2	19	1.56	Medium
M3	SRBC	Sugar Run near Sugar Run, PA[2]	33.5	48	1.43	Medium
M4	SRBC	Snake Creek near Lawsville Center, PA	45.1	61	1.35	Medium
M5	SRBC	Tomjack Creek near Burlington, PA	27.0	25	0.93	Medium
M6	SRBC	Hammond Creek near Millerton, PA	29.1	22	0.76	Medium
M7	PADEP WQN, USGS streamgage	Towanda Creek near Monroeton, PA[3]	215	155	0.72	Medium
M8	SRBC	Upper Crooked Creek near Keeneyville, PA	47.4	32	0.68	Medium
M9	SRBC	Wappasening Creek near Windham Center, PA	47.1	30	0.64	Medium
M10	USGS streamgage	Wappasening Creek near Windham Center, PA	56.9	35	0.62	Medium
M11	SRBC	Little Snake Creek near Brackney, PA	7.70	4	0.52	Medium
L1	SRBC	Apalachin Creek near Apalachin, NY	42.8	20	0.47	Low
L2	USGS streamgage	Choconut Creek near Choconut, PA	18.5	6	0.32	Low
L3	SRBC	Bentley Creek at Wellsburg, NY	54.2	14	0.26	Low

Map ID	Monitoring organization	Site name	Drainage area (mi^2)	Number of HVHF wells in monitored watershed	Density of HVHF wells in monitored watershed ($wells/mi^2$)	Site type (density)[1]
L4	SRBC	Choconut Creek near Vestal Center, NY	38.0	8	0.21	Low
L5	SRBC	Kitchen Creek near Huntington Mills, PA[4]	20.1	1	0.05	Low
		North Central Appalachians ecoregion				
H6	PADEP WQN	Hoagland Run near Quiggleville, PA	10.7	55	5.14	High
H7	SRBC	Larrys Creek near Salladasburg, PA	29.1	64	2.20	High
M12	SRBC	Tioga River near Fall Brook, PA	13.5	25	1.85	Medium
M13	PADEP WQN	Hyner Run near Hyner, PA	24.9	44	1.77	Medium
M14	SRBC	Trout Run near Shawville, PA	32.7	40	1.22	Medium
M15	SRBC	Elk Run near Watrous, PA	21.0	25	1.19	Medium
M16	SRBC	Grays Run near Gray, PA	16.2	16	0.99	Medium
M17	PADEP WQN, SRBC, USGS streamgage	Blockhouse Creek near English Center, PA	37.7	33	0.88	Medium
M18	SRBC	Little Pine Creek near Waterville, PA[3]	180	154	0.86	Medium
M19	PADEP WQN, USGS streamgage	Lycoming Creek near Trout Run, PA[2]	173	129	0.75	Medium
L6	SRBC	Baker Run near Glen Union, PA	35.0	16	0.46	Low
L7	SRBC	Sterling Run near Sterling Run, PA	28.9	9	0.31	Low
L8	SRBC	Moose Creek near Plymptonville, PA	3.32	1	0.30	Low
L9	PADEP WQN	Rock Run near Ralston, PA	28.0	7	0.25	Low
L10	USGS streamgage	Straight Run, Site 1, at marsh Creek, PA[2]	6.48	1	0.02	Low
L11	SRBC	Hicks Run near Hicks Run, PA	34.1	5	0.15	Low
L12	SRBC	Ninemile Run near Walton, PA	15.7	2	0.13	Low
L13	PADEP WQN	East Fork Sinnemahoning Creek near Wharton, PA	47.6	6	0.13	Low
L14	SRBC	East Fork First Fork Sinnemahoning Creek near Logue, PA	32.8	4	0.12	Low
L15	SRBC	West Creek near Weber City, PA	59.4	3	0.05	Low
L16	SRBC	Portage Creek near Emporium, PA	70.5	2	0.03	Low
L17	SRBC	West Branch Pine Creek near Galeton, PA	70.0	1	0.01	Low

Map ID	Monitoring organization	Site name	Drainage area (mi²)	Number of HVHF wells in monitored watershed	Density of HVHF wells in monitored watershed (wells/mi²)	Site type (density)[1]
		Central Appalachians ecoregion				
L18	SRBC	Little Clearfield Creek near Dimeling, PA	44.4	7	0.16	Low
		Ridge and Valley ecoregion				
M20	SRBC	Little Muncy Creek near Moreland, PA	51.0	77	1.51	Medium
L19	SRBC	Bobs Creek near Pavia, PA	17.0	3	0.18	Low
L20	PADEP WQN	Bobs Creek below Wallacks Branch at Pavia, PA	22.1	3	0.14	Low

[1]High Density: >2.00, Medium Density: 0.51 to 2.00, Low Density: 0.01 to 0.50
[2]Watershed is located in both Northern Allegheny Plateau and North Central Appalachians ecoregions
[3]Watershed is located in both North Central Appalachians and Central Appalachians ecoregions
[4]Watershed is located in the Northern Allegheny Plateau, North Central Appalachians, and Ridge and Valley ecoregions

The minimum duration of monitoring to detect change depends on the ecoregion of the monitoring site. Increased monitoring to answer the case-study policy question should be considered a long-term investment. Ongoing monitoring at these locations (10 or more years) would identify whether cumulative effects of shale gas development result in detectable contamination of surface water in these watersheds over the long term.

The set of monitoring sites described in Table 19 includes a total of 17 long-term monitoring sites, 13 sites in watersheds with HVHF wells and 4 reference sites in watersheds that currently have no HVHF wells (Figure 27). This set of monitoring sites includes portions of 25 HUC12 watersheds with HVHF wells in their catchments, out of a total of 193 HUC12 watersheds with HVHF wells. There are 863 HUC12 watersheds in the Susquehanna River Basin. Table 18 and Table 19 present just one set of monitoring sites where increased monitoring could answer the case-study policy question. Monitoring agencies may want to consider other combinations of monitoring sites for any of the reasons listed above.

To identify whether cumulative impacts of shale gas development are changing surface-water quality, statistical analysis would be applied to data collected at both watersheds with a high density of HVHF wells and reference watersheds to see if there are differences in detected trends. If a control watershed shows the same trend as the watershed with a high density of HVHF wells, then shale gas development may not be the cause of detected change in water quality.

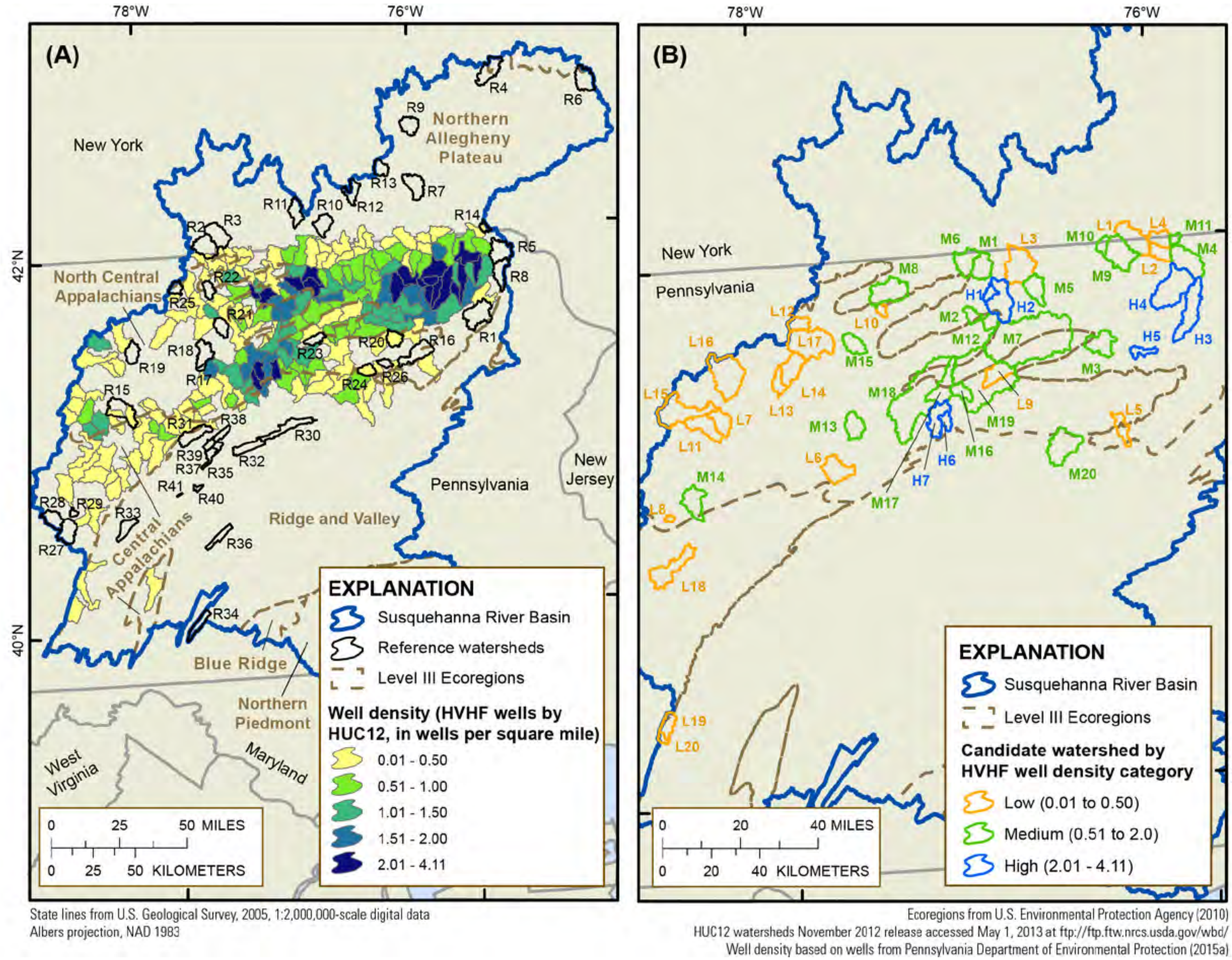


Figure 26. Candidate sites for surface-water monitoring in ecoregions with shale gas development activity in the Susquehanna River Basin that meet most monitoring criteria for determining whether cumulative shale gas development activities are resulting in surface-water quality change. (A) Candidate reference sites, and (B) Candidate monitoring sites in watersheds with HVHF wells. Site labels are defined in Table 16 and Table 17.
[Abbreviations: HVHF, high-volume hydraulic fracturing; HUC12, 12-digit hydrologic unit code boundaries]

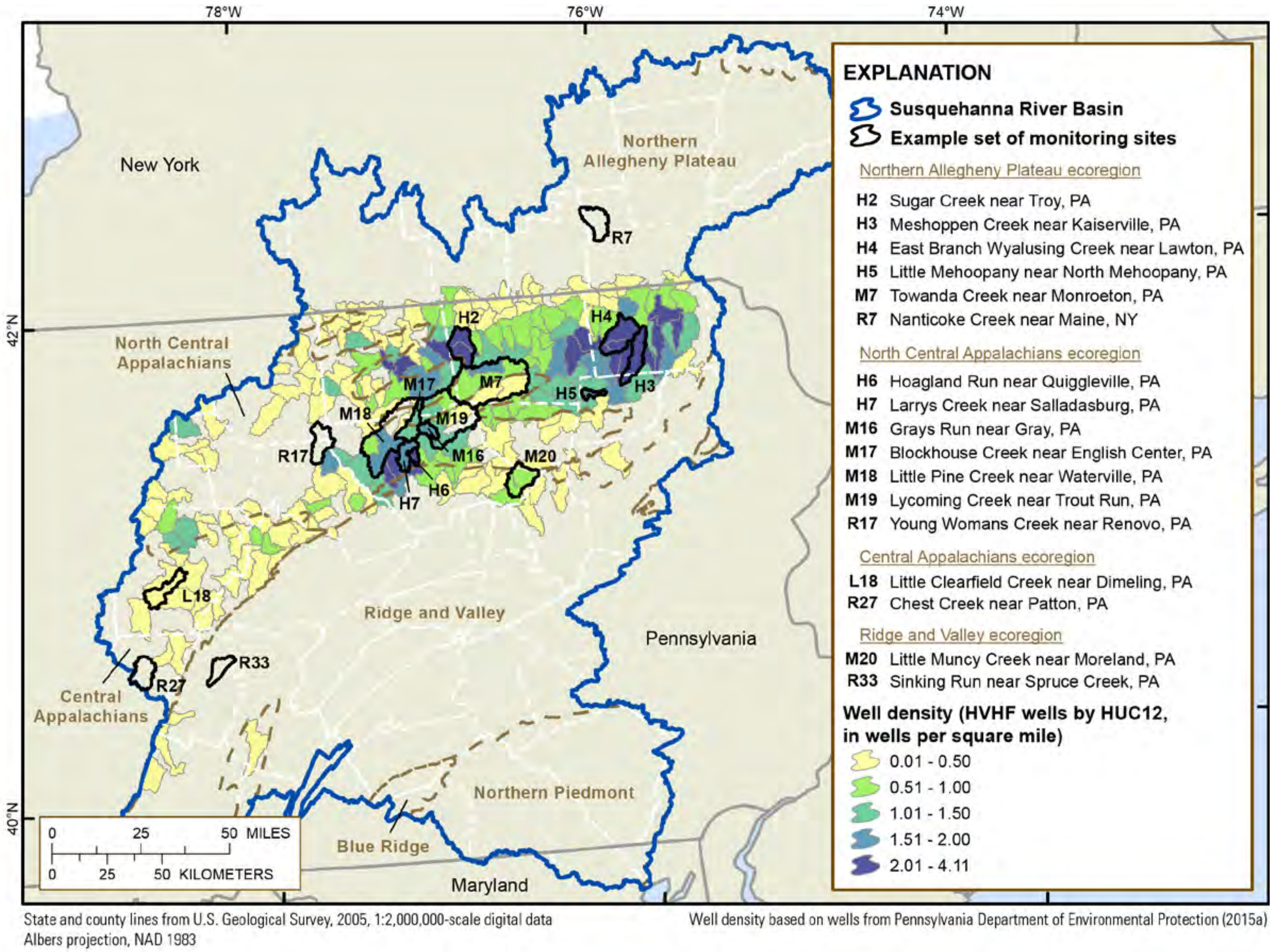


Figure 27. Watersheds for the example set of monitoring sites that could be used to measure whether cumulative shale gas development activities in the Susquehanna River Basin are resulting in surface-water quality change. Site labels are defined in Table 18.
[Abbreviations: SRBC, Susquehanna River Basin Commission; PADEP, Pennsylvania Department of Environmental Protection; USGS, U.S. Geological Survey; HVHF, high-volume hydraulic fracturing; HUC12, 12-digit hydrologic unit code boundaries]

Table 18. Example set of monitoring sites that could be used to measure whether cumulative shale gas development activities in the Susquehanna River Basin are resulting in surface-water quality change and factors related to their selection. The Map IDs correspond to the sites shown in Figure 26(A) and (B), and Figure 27.
[Abbreviations: HVHF, high-volume hydraulic fracturing; ecoregion, U.S. Environmental Protection Agency Level III Ecoregion; Diss. Ba, dissolved barium; U, uranium; SSC, suspended sediment concentration; Ra, radium; Li, lithium; K, potassium; TOC, total organic carbon; SRBC, Susquehanna River Basin Commission; PADEP WQN, Pennsylvania Department of Environmental Protection Water Quality Network; NWIS, National Water Information System]

Map ID	Monitoring organization	Short site name	Site type[1]	Pre-2007 data	Active monitoring site	Needed parameters	Meets sampling frequency criteria	Continuous Monitor[2]	Streamgage present	Nested monitoring
					Northern Allegheny Plateau ecoregion					
H2	SRBC	Sugar Creek, PA	High	No	Yes	Diss. Ba, U, Ra	No	SRBC	No	No
H3	SRBC	Meshoppen Creek, PA	High	No	Yes	Diss. Ba, U, Ra	No	SRBC	No	No
H4	SRBC	East Branch Wyalusing Creek, PA	High	No	Yes	Diss. Ba, U, Ra	No	SRBC	No	No
H5	SRBC	Little Mehoopany Creek, PA	High	No	Yes	Diss. Ba, U, Ra	No	SRBC	No	No
M7	PADEP WQN, USGS streamgage	Towanda Creek, PA	Medium	Yes	Yes	Diss. Ba, gross alpha and beta, U, Li, K, SSC, TOC, turbidity, Ra	No	No	Yes	No but potential exists
R7	SRBC	Nanticoke Creek, NY	Reference	No	Yes	Diss. Ba, U, Ra	No	SRBC	No	No
					North Central Appalachians ecoregion					
H6	PADEP WQN	Hoagland Run, PA	High	No	Yes	Diss. Ba, gross alpha and beta, U, Li, K, SSC, TOC, turbidity, Ra	Yes	No	No	No
H7	SRBC	Larrys Creek, PA	High	No	Yes	Diss. Ba, U, Ra	No	SRBC	No	No
M16	SRBC	Grays Run, PA	Medium	No	Yes	Diss. Ba, U, Ra	No	SRBC	No	Yes with Lycoming Creek
M17	PADEP WQN, SRBC, USGS streamgage	Blockhouse Creek, PA	Medium	No	Yes	Diss. Ba, gross alpha and beta, U, Li, K, SSC, TOC, turbidity, Ra[3]	Yes	SRBC	Yes	Yes with Little Pine Creek
M18	SRBC	Little Pine Creek, PA[4]	Medium	No	Yes	Diss. Ba, U, Ra	No	SRBC	No	Yes with Blockhouse Creek

Map ID	Monitoring organization	Short site name	Site type[1]	Pre-2007 data	Active monitoring site	Needed parameters	Meets sampling frequency criteria	Continuous Monitor[2]	Streamgage present	Nested monitoring
M19	PADEP WQN, USGS streamgage	Lycoming Creek, PA[5]	Medium	Yes	Yes	Diss. Ba, gross alpha and beta, U, Li, K, SSC, TOC, turbidity, Ra	Yes	No	Yes	Yes with Grays Run
R17	PADEP WQN, USGS streamgage	Young Womans Creek near Renovo, PA	Reference	Yes	Yes	Diss. Ba, gross alpha and beta, U, Li, K, SSC, TOC, turbidity, Ra	No	No	Yes	No
				Central Appalachians ecoregion						
L18	SRBC	Little Clearfield Creek, PA	Low	No	Yes	Diss. Ba, U, Ra	No	SRBC	No	No
R27	SRBC	Chest Creek, PA	Reference	No	Yes	Diss. Ba, U, Ra	No	SRBC	No	No
				Ridge and Valley ecoregion						
M20	SRBC	Little Muncy Creek, PA	Medium	No	Yes	Diss. Ba, U, Ra	No	SRBC	No	No
R33	Discontinued PADEP WQN	Sinking Run, PA	Reference	Yes	No	All surface-water parameters from Table 3	No	No	No	No

[1]High Density: >2.00 HVHF wells/mi^2, Medium Density: 0.51 to 2.00 HVHF wells/mi^2, Low Density: 0.01 to 0.50 HVHF wells/mi^2, Reference Site: 0.
[2]SRBC continuous monitors measure pH, temperature, specific conductance, dissolved oxygen and turbidity
[3]PADEP samples 12x per year excluding the listed parameters; SRBC samples 4x per year excluding Diss. Ba, U, and Ra.
[4]Watershed is located in both North Central Appalachians and Central Appalachians ecoregions
[5]Watershed is located in both Northern Allegheny Plateau and North Central Appalachians ecoregions

Table 19. Summary of water monitoring data needed at the example set of monitoring sites described in Table 18 to answer the case-study policy question "Do shale gas development activities contaminate surface water or groundwater?" based on criteria in Table 11. The Map IDs correspond to the sites shown in Figure 260(A) and (B), and Figure 27.

[**Abbreviations:** HVHF, high-volume hydraulic fracturing; ecoregion, U.S. Environmental Protection Agency Level III Ecoregion; Diss. Ba, dissolved barium; U, uranium; SSC, suspended sediment concentration; Ra, radium; Li, lithium; K, potassium; TOC, total organic carbon]

Map ID	Short site name	Site type[1]	Additional site visits per year	Streamgage needed	Additional parameters[2]	Minimum years of monitoring to detect change, based on monthly sampling[3]	Ancillary data
				Northern Allegheny Plateau ecoregion			
H2	Sugar Creek, PA	High	8	Yes	Diss. Ba, U, and Ra	3	Annual HVHF permit and other parameter source data within monitored watershed
H3	Meshoppen Creek, PA	High	8	Yes	Diss. Ba, U, and Ra	3	Annual HVHF permit and other parameter source data within monitored watershed
H4	Little Mehoopany Creek, PA	High	8	Yes	Diss. Ba, U, and Ra	3	Annual HVHF permit and other parameter source data within monitored watershed
H5	East Branch Wyalusing Creek, PA	High	8	Yes	Diss. Ba, U, and Ra	3	Annual HVHF permit and other parameter source data within monitored watershed
M7	Towanda Creek, PA	Medium	6	No	Diss. Ba, gross alpha and beta, U, Li, K, SSC, TOC, turbidity, and Ra	3	Annual HVHF permit and other parameter source data within monitored watershed
R7	Nanticoke Creek, NY	Reference	8	Yes	Diss. Ba, U, and Ra on an as-needed basis	3	Annual HVHF permit and other parameter source data within monitored watershed
				North Central Appalachians ecoregion			
H6	Hoagland Run, PA	High	0	Yes	Diss. Ba, gross alpha and beta, U, Li, K, SSC, TOC, turbidity, and Ra	3-5	Annual HVHF permit and other parameter source data within monitored watershed
H7	Larrys Creek, PA	High	8	Yes	Diss. Ba, U, and Ra	3-5	Annual HVHF permit and other parameter source data within monitored watershed
M16	Grays Run, PA	Medium	8	Yes	Diss. Ba, U, and Ra	3-5	Annual HVHF permit and other parameter source data within monitored watershed

Map ID	Short site name	Site type[1]	Additional site visits per year	Streamgage needed	Additional parameters[2]	Minimum years of monitoring to detect change, based on monthly sampling[3]	Ancillary data
M17	Blockhouse Creek, PA	Medium	0	No	Diss. Ba, gross alpha and beta, U, Li, K, SSC, TOC, turbidity, and Ra[4]	3-5	Annual HVHF permit and other parameter source data within monitored watershed
M18	Little Pine Creek, PA[5]	Medium	8	Yes	Diss. Ba, U, and Ra	3-5	Annual HVHF permit and other parameter source data within monitored watershed
M19	Lycoming Creek, PA[6]	Medium	0	No	Diss. Ba, gross alpha and beta, U, Li, K, SSC, TOC, turbidity, and Ra	3-5	Annual HVHF permit and other parameter source data within monitored watershed
R17	Young Womans Creek near Renovo, PA	Reference	0	No	Diss. Ba, gross alpha and beta, U, Li, K, SSC, TOC, turbidity, and Ra	3-5	Annual HVHF permit and other parameter source data within monitored watershed
				Central Appalachians ecoregion			
L18	Little Clearfield Creek, PA	Low	8	Yes	Diss. Ba, U, and Ra	3[7]	Annual HVHF permit and other parameter source data within monitored watershed
R27	Chest Creek, PA	Reference	8	Yes	Diss. Ba, U, and Ra	3[7]	Annual HVHF permit and other parameter source data within monitored watershed
				Ridge and Valley ecoregion			
M20	Little Muncy Creek, PA	Medium	8	Yes	Diss. Ba, U, and Ra	3-6	Annual HVHF permit and other parameter source data within monitored watershed
R33	Sinking Run, PA	Reference	12	Yes	All surface-water parameters from Table 3	3-6	Annual HVHF permit and other parameter source data within monitored watershed

[1]High Density: >2.00 HVHF wells/mi^2, Medium Density: 0.51 to 2.00 HVHF wells/mi^2, Low Density: 0.01 to 0.50 HVHF wells/mi^2, Reference Site: 0.
[2]Radium on an as-needed basis
[3]Based on power analysis results in Table 9 and Table 10
[4]PADEP samples 12x per year excluding the listed parameters; SRBC samples 4x per year excluding Diss. Ba, U, and Ra.
[5]Watershed is located in both North Central Appalachians and Central Appalachians ecoregions
[6]Watershed is located in both Northern Allegheny Plateau and North Central Appalachians ecoregions
[7]A minimum of 3 years of monthly data are needed to complete a power analysis, but no sites in the Central Appalachians ecoregion were analyzed

8.2 Cost of Surface Water Monitoring

The cost of new monitoring to answer the case-study policy question is not small. However, these monitoring costs should be considered in the context of the economic value of the pristine watersheds and cold-water fisheries that are prevalent in the Marcellus and Utica Shale area of the Susquehanna River Basin; in Pennsylvania, $2.8 billion was spent on fishing, hunting, and wildlife viewing in 2011 (U.S. Fish and Wildlife Service, 2014). In addition, some tributaries in the Susquehanna River Basin are drinking water sources; the degradation of these waters is a substantial public health, cost, and quality of life issue for the people who live there. The potential for shale gas development activities to contaminate these environments is a major concern to many individuals who live in the Susquehanna River Basin. A lack of monitoring sufficient for detecting water-quality trends related to shale gas development delays the ability to identify potential contamination associated with shale gas development and to prevent or minimize the consequences of such contamination. The annual cost of new monitoring, as estimated in the following paragraphs, is small in comparison to the economic value of the headwaters of the Susquehanna River.

The cost of new monitoring presented here relies on data provided by the USGS. Table 20 shows average surface-water related sampling costs for 43 stream and river sites monitored as part of the USGS NAWQA Program, exclusive of laboratory costs. Table 20 presents these costs based on Fiscal Year (FY) 2012 data adjusted by 3.8 percent to account for inflation to October 2015 (U.S. Department of Labor Bureau of Labor Statistics, 2015). Water-quality samples are collected between 6 and 26 times per year to meet specific monitoring objectives (median = 18). Average costs were calculated on a per sample basis for each site. Costs include salary, travel, supplies, equipment, miscellaneous and administrative services. Note that salary costs include field labor, office labor including sample preparation, clean-up, sample shipping, data entry, and quality control and quality assurance management costs. The average October 2015 cost per sample, not including laboratory costs, was about $4,300 and the cost per site ranged from $2,900 to $6,000.

Table 20. Estimated cost per sample for operating a typical U.S. Geological Survey stream water-quality monitoring site, exclusive of laboratory analytical costs or streamgages. Average costs are presented from 43 stream and river monitoring sites as of October, 2015.

Cost category	Average cost per sample as of October, 2015	Percent of total cost by category
Salary cost estimate (2 person crew)	$2,271	53%
Vehicle cost estimate	$185	4%
Travel cost estimate	$39	1%
Supply cost estimate	$118	3%
Equipment cost estimate	$222	5%
Total plus administrative services	**$4,261**	**100%**

The range in costs reflects factors such as differences in distances required to travel between an office and monitoring sites as well as differences in stream size and the requisite sampling time, equipment, and sampling methods. In addition, these costs reflect the data-quality requirements of the USGS NAWQA Program.

Table 21 shows the total cost of analysis for the suite of priority surface-water parameters based on laboratory costs that would be incurred by use of the USGS National Water Quality Laboratory in Denver, Colorado, using Fiscal Year 2016 prices. The total cost of analyzing the parameters listed in Table 21 is $442, excluding radium. Testing for Ra-226 and Ra-228 in surface water is recommended only if there is a change in gross alpha and gross beta measurements. If measurements of gross alpha and gross beta indicate possible radioactive contamination, the TAC recommends enhanced monitoring for Ra-226 and Ra-228 isotopes. Adding Ra-226 and Ra-228 increases the total analytical cost to $821 per sample.

Total costs in Table 21 do not include the cost of collecting and analyzing quality-assurance/quality-control (QA/QC) samples, which can include blank, replicate, spike, and blind reference-material samples. Quality control requirements vary by constituent group (for example, "nutrients") because different groups require different types and numbers of QA/QC samples. Since the suite of parameters shown in Table 21 includes parameters from various constituent groups, an average of the cost of QA/QC samples for the NAWQA Program was used. The cost associated with QA/QC samples was 13.7 percent of the laboratory analysis cost (G. Rowe, U.S. Geological Survey, written commun., March 2014). Adding in a factor of 13.7 percent ($61) to account for the cost of QA/QC samples yields a total cost of $503 ($442 + $61) for the laboratory analytical costs and QA/QC sample costs, for an average total cost of $4,800 per site visit.

The analytical costs for laboratory analysis of environmental and related QA/QC samples were added to the costs from Table 20 to estimate the total cost of monitoring in streams or rivers for the priority surface-water parameters on an annual basis, assuming monthly sampling. The average annual cost for the suite of priority surface-water parameters sampled 12 times per year would be $57,200, with a minimum of $41,200 and maximum of $78,500 depending on the geographic proximity of the office to field sites, the size of streams and other factors.

As a point of comparison, the SRBC provided cost information for their RWQMN sites. More than 50 water-quality monitoring sites are currently in operation through this program. Each station continuously monitors the following field parameters: pH, water temperature, specific conductance, dissolved oxygen, turbidity, and relative water depth. In addition, SRBC staff collects quarterly water samples, which are analyzed for 22 parameters at a Pennsylvania State-certified laboratory. Each SRBC monitoring site costs about $20,000 to purchase and install, and about $8,000 annually to maintain; this annual maintenance cost includes labor, equipment servicing, laboratory costs, shipping samples to the laboratory, travel and vehicle, overhead, supply costs, and data management (A. Gavin, Susquehanna River Basin Commission, written commun., July 2013).

Table 21. Parameter analysis costs at USGS laboratories and reporting levels for the suite of priority surface-water parameters, Fiscal Year 2016.

[Abbreviations: µg/L, micrograms per liter; mg/L, milligrams per liter, pCi/L, picocuries per liter; TDS, total dissolved solids]

Parameter Name	Parameter Code	CAS Number	Reporting level	Unit	Cost per analysis[1]
Barium, dissolved	01005	7440-39-3	0.25	µg/L	$27[2]
Barium, total	01005	7440-39-3	0.25	µg/L	$7[2]
Bromide	71870	24959-67-9	0.03	mg/L	$22
Calcium	00915	7440-70-2	0.022	mg/L	$8
Chloride	00940	16887-00-6	0.02	mg/L	$14
Gross-alpha radioactivity	62636	12587-46-1	3	pCi/L	$150
Gross-beta radioactivity	62642	12587-47-2	4	pCi/L	
Lithium	01130	7439-93-2	0.1	µg/L	$8
Magnesium	00925	7439-95-4	0.011	mg/L	$8
Nitrogen-nitrate + nitrite	00631	NA	0.04	mg/L	$9
Potassium	00935	7440-09-7	0.004	mg/L	$8
Radium-224[1]	50833	13233-32-4	1	pCi/L	$245
Radium-226	09503	13982-63-3	1	pCi/L	
Radium-228[1]	81366	15262-20-1	1	pCi/L	$134
Sodium	00930	7440-23-5	0.06	mg/L	$8
Strontium	01080	7440-24-6	0.2	µg/L	$8
Sulfate	00945	14808-79-8	0.02	mg/L	$14
Suspended sediment concentration (SSC)	80154	NA	1	mg/L	$50
Residue, 180 degrees Celsius (TDS)	70300	NA	20	mg/L	$15
Organic carbon (Total)	00680	NA	0.7	mg/L	$57
Phosphorus (Total)	00665	7723-14-0	0.004	mg/L	$21
Uranium, natural	22703	7440-61-1	0.014	µg/L	$8
Total					$442[1]

[1] Total sample cost is $442, excluding radium. Cost including radium is $821. Prices apply to Fiscal Year 2016 (Oct 1, 2015-Sept 30, 2016). Numbers may not add due to rounding error.

[2] The price shown for barium analysis includes measurement of the specific conductance measurement ($10.58; this measurement is used to screen for high salinity which could bias the results of this or any other trace element analysis) and also includes the ICP-MS instrument set up cost ($9.08). These are one-time charges that are not applied to any other major inorganic or trace element analyses requested.

These estimated costs for collecting and analyzing water samples do not include costs related to measuring streamflow or operating a streamgage. Cost information provided by the USGS indicates that operation and maintenance of a standard streamgage is $14,000 per year (Greg Koltun, U.S. Geological Survey, written commun., May 15, 2014), or $14,100 adjusted for October, 2015. This cost does not include initial installation and equipment costs, which can be as low as $15,000 for a simple installation to $20-22,000 for a walk-in shelter (G. Koltun, U.S. Geological Survey, written commun., May 15, 2014). To estimate costs for this case study, $17,600 was used as the average cost to represent one-time start-up costs for a new streamgage.

As shown in Table 22, the 10-year cost for a new monitoring site with a new streamgage would be $834,600, for an average annual cost of $83,500. The incremental cost for an existing SRBC monitoring site reflects an increase in sampling frequency from quarterly to monthly sampling, additional parameter analysis as specified in Table 19, and new streamgages. The incremental cost would be $616,900 over 10 years, for an annual average of $61,700 over 10 years. The annual average incremental cost at PADEP monitoring sites is different at each monitoring site listed in Table 19: $35,200 at Towanda Creek where 6 additional site visits are needed each year; $22,400 for Hoagland Run where a new streamgage is needed; $4,500 at Lycoming Creek where additional sample analysis is needed; and $3,100 at Blockhouse Creek where there is an existing streamgage and both the SRBC and the PADEP collect samples but additional sample analysis is still needed.

Table 23 provides estimates of the annual cost for a surface-water monitoring program to detect whether cumulative shale gas development activities in the Susquehanna River Basin are resulting in surface-water quality change. The surface-water monitoring identified here should continue for at least 10 years to detect whether cumulative effects of shale gas development change surface-water quality over the long-term. As shown in Table 23, the cost of option 1 to initiate monitoring at 17 new sites, if there were no appropriate existing monitoring sites, would be an annual average of $1.4 million for 10 years. The cost for option 2, the set of monitoring sites presented in Table 19, includes increased monitoring plus streamgages at 11 SBRC monitoring sites, increased monitoring at 5 PADEP WQN sites, and new monitoring at one site (R33, Sinking Run, PA in Table 19), which results in an average annual cost of $831,600 over 10 years in addition to the cost of the ongoing SRBC and PADEP WQN monitoring programs. Option 2 (the set of monitoring sites in Table 19) includes a total of 17 monitoring sites, more than the minimum number of monitoring sites specified in section 6.5. This selection builds on the highest HVHF density watersheds with existing monitoring sites, plus additional nested watershed monitoring sites that will provide critical information regarding the scope and magnitude of potential water-quality change associated with shale gas development.

In Table 23, option 3 represents the cost of the minimum monitoring program to answer the case-study policy question, if there were no appropriate existing monitoring sites. The average annual cost for this program is $667,600 for 10 years. Option 4 represents the incremental cost of the minimum monitoring program to answer the case-study policy question, building on existing monitoring sites in the Susquehanna River Basin. This option includes increased monitoring plus streamgages at 5 SBRC monitoring sites, increased monitoring at 3 PADEP WQN sites, and two additional streamgages at the PADEP WQN sites, for an annual average cost of $357,800 for 10 years in addition to the cost of the ongoing SRBC and PADEP WQN monitoring programs.

These generalized cost estimates are based on historical, national average monitoring costs for the USGS NAWQA Program and those provided by the SRBC. A more rigorous estimate would include multiple estimates of cost from additional monitoring programs, and could be further refined by more rigorously estimating the travel and other costs for the specific monitoring program recommended here.

Table 22. Estimated water-quality and streamflow monitoring costs per monitoring site sampled for 10 years, assuming 3-percent annual inflation.
[**Abbreviations:** FY, fiscal year; %, percent; NA, Not Applicable; NY, New York; PA, Pennsylvania, SRBC, Susquehanna River Basin Commission, PADEP WQN, Pennsylvania Department of Environmental Protection Water Quality Network]

Type of Monitoring site	A. FY 2016 cost per sample	B. Number of new site visits per year	C. Total sample cost per year (A*B)	D. Additional analysis for current site visits	E. Streamgage installation cost (one-time cost)	F. FY 2016 annual streamgage operation cost	G. Ten-year total cost assuming 3% annual inflation	H. Average annual cost per monitoring site
Streamgage only	NA	NA	NA	NA	$17,600	$14,100	$179,200	$17,900
New monitoring site	$4,764	12	$57,163	NA	$17,600	$14,100	$834,600	$83,500
SRBC incremental cost	$4,764	8	$38,112	$68	$17,600	$14,100	$616,900	$61,700
PADEP WQN incremental cost (standard site, 6x/year)	$4,786	6	$28,716	$1,965	NA	NA	$351,700	$35,200
PADEP WQN incremental cost (reference site, 12x/year)	NA	NA	NA	$3,929	NA	NA	$45,000	$4,500

Table 23. Estimated cost for a surface-water monitoring program to detect whether cumulative shale gas development activities in the Susquehanna River Basin are resulting in surface-water quality change.

Monitoring program description	A. FY 2016 cost of new monitoring	B. Cost of streamgage installation	C. FY 2016 annual streamgage operation cost	D. Ten-year total cost assuming 3% annual inflation	E. Average annual cost per monitoring program
Option 1: 17 new monitoring sites	$971,800	$299,200	$239,700	$14,187,000	$1,419,000
Option 2: Incremental cost of 17 monitoring sites building on existing sites in SRBC and PADEP WQN networks (Table 19)	$522,100	$228,800	$183,300	$8,316,000	$831,600
Option 3: Minimum monitoring program (8 new monitoring sites)	$457,300	$140,800	$112,800	$6,676,000	$667,600
Option 4: Incremental cost of 8 monitoring sites building on existing sites in SRBC and PADEP WQN networks	$202,700	$123,200	$98,700	$3,578,000	$357,800

8.3 Cost of Data Analysis

Water-quality monitoring data on their own do not provide information to answer policy questions; data analysis is the critical step of converting data to information. The cost of data analysis must be factored into the cost of new monitoring because data without analysis provide very little value. Data analysis plays a critical role in quality control and is essential for adaptive management. The USGS NAWQA Program, based on more than two decades of running multiple data collection activities and analyzing, modeling, and reporting on those data, uses the cost of data collection as the starting point for budgeting the cost of data analysis (Gary Rowe, U.S. Geological Survey, written commun., 2014). New data collected through the program described here should be analyzed annually to understand ongoing trends in the data and to support efficient use of monitoring resources through adaptive management. The total dollar value obtained by doubling the estimates provided above would cover data collection, data management, and data analysis. Using Table 23 Option 2, it would cost approximately $1.7 million per year to fill the data gaps and analyze water-quality trends related to shale gas development in the Susquehanna River Basin, in addition to the cost of the ongoing SRBC and PADEP WQN monitoring programs. The annual cost of the new surface-water monitoring recommended in this report represents only 0.06 percent of the estimated $2.8 billion generated annually by fishing, hunting and wildlife viewing in the Susquehanna River Basin (U.S. Fish and Wildlife Service, 2014).

9. Quantities of Groundwater Data Needed to Answer the Case-Study Policy Question

The next three chapters focus on groundwater and describe the water data needed and available to answer “Do shale gas development activities contaminate groundwater?” in the Susquehanna River Basin, building on the criteria discussed in Chapter 4. The objectives for answering this question are to be able to detect whether groundwater concentrations of parameters associated with shale gas development have changed over time, and to identify whether any change detected is a result of shale gas development. These objectives, plus a testable hypothesis, were used to define a sampling plan in terms of the number and location of appropriate sampling sites, the number of sampling events, and an appropriate statistical design for interpreting monitoring results.

9.1 Focus Parameters – Groundwater

As for surface water, the TAC identified focus parameters from the suite of priority groundwater parameters to be used as examples quantifying the data needed to answer the case-study policy question. The focus parameters chosen by the TAC are methane, barium (dissolved), specific conductance, and bromide (dissolved).

The TAC selected methane as a focus parameter for groundwater due to the publicity of reports of methane contamination in domestic wells (Lustgarten, 2009) and because these incidents may result from shale gas development (Jackson et al., 2013). Barium, specific conductance, and bromide were selected for further analysis because these parameters had among the longest records for individual sampling sites or the most data records in the Susquehanna data set. They are the parameters most likely to have enough data to detect whether a statistically significant change in concentration has occurred. In addition, barium and bromide are highly specific signatures of produced waters (Vidic et al., 2013) and may be indicators of contamination related to shale gas development. Specific conductance may show potential changes in water quality and has the most records, but it is not specific to water-quality changes resulting from shale gas development.

The range of existing concentrations of focus parameters in both brine and in groundwater that is presumed unaffected by shale gas development can be used to estimate outer bounds on the potential magnitude of change that could result from shale gas development. Given the number and complexity of contamination pathways and the relative lack of information about the cumulative effects of shale gas development, it is difficult to estimate an expected magnitude of change in parameter concentrations at any given well. Brine and background water-quality data were used to illustrate an upper and lower bound for potential contaminant concentrations in the Susquehanna River Basin.

The PADEP provided water-quality data for produced waters from some Marcellus Shale HVHF wells (T. Shaw, Pennsylvania Department of Environmental Protection, written commun., November 2013). The plot in Figure 28 shows concentrations of several priority groundwater parameters in 30 samples of produced water obtained by the PADEP, and baseline groundwater quality for 20 samples from domestic wells in Sullivan County, PA (Sloto, 2013). The plot is an illustration of the difference between brine and fresh groundwater, and is not meant to characterize groundwater quality in the Susquehanna River Basin. Although it is small, the Sloto (2013) data set was chosen because it has information for a broad suite of parameters, is considered 'baseline' for shale gas development, and is derived from domestic wells. The parameters are arranged in order of decreasing median value for the PADEP data; maximum contaminant levels, health advisory levels, and secondary drinking water standards are also shown for reference.

The amount of produced water required to double the concentration of barium in 1 liter of water compared to background levels, assuming a barium concentration of 4,155 mg/L in the produced water (which is the median value of barium concentration in produced water; Figure 28) and a background concentration of barium of 0.14 mg/L (which is equal to the median value found in groundwater wells in Sullivan County (Sloto, 2013)) would be approximately 0.034 milliliter of produced water—0.068 drop (assuming 20 drops in 1 milliliter).

9.2 Hypothesis – Groundwater

As for surface water, clear identification of monitoring objectives and hypotheses to be tested are critical for identifying a successful water-quality monitoring design (Sanders and others, 1983; Gilbert, 1987; U.S. Environmental Protection Agency, 2002 and 2006). The following hypothesis was selected for groundwater, reflecting a possible scenario of shale gas development activities resulting in water-quality change:

- Cumulative effects of shale gas development using HVHF will cause impacts to groundwater quality near HVHF wells over the long term, resulting in statistically significant increases of methane, specific conductance, bromide and/or barium concentrations in groundwater sampling wells located within 1 mile of an HVHF well.

This hypothesis is referred to as the "groundwater hypothesis" throughout this document.

9.3 Statistical Design to Test the Groundwater Hypothesis

Trend analysis is typically used for determining if long-term and/or low-level changes in water quality are occurring. To test the groundwater hypothesis, sufficient data must be collected using an appropriate sampling design such that statistically significant trends in water quality can be detected through analysis. Two statistical design options were considered for their applicability for testing the groundwater hypothesis: monotonic trend analysis and "network" or step-trend analysis. These methods are based on temporal and spatial analysis, respectively.

As for surface water, monotonic trend analysis was examined for its applicability to the existing groundwater monitoring data in the Susquehanna River Basin. Monotonic trend analysis requires multiple samples taken at an individual monitoring site over time. For surface water, 36 samples were used as a starting point for trend detection; fewer samples are needed for groundwater, which does not experience the same magnitude of annual and seasonal variability as surface water does due to streamflow. Ten

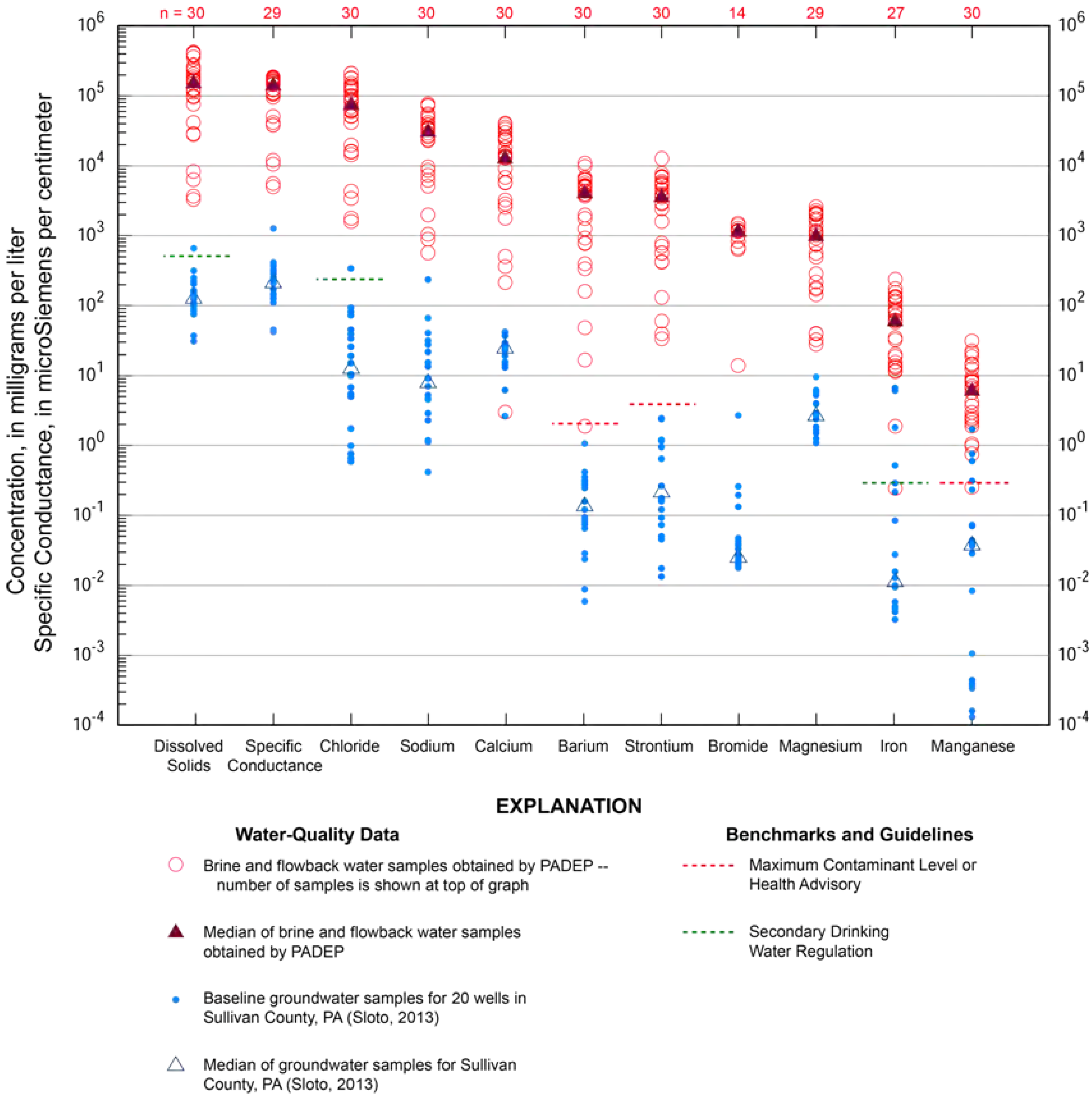


Figure 28. Concentrations of selected parameters in produced water from HVHF wells in the Susquehanna River Basin compared to background or baseline groundwater quality in domestic wells of Sullivan County, PA.
[Abbreviations: PADEP, Pennsylvania Department of Environmental Protection; HVHF, high-volume hydraulic fracturing]

samples were used as a threshold for identifying available groundwater data. While there are numerous sampling sites with specific conductance data in or near the Marcellus and Utica Shale area of the Susquehanna River Basin (n=3,899), only 43 sites have 10 or more specific conductance measurements (Table 24). Out of these 43 sites, only 8 have data before and after 2007, and none of the 8 sites are within 1 mile of an HVHF well. Specific conductance was the parameter with the most long-term records in the Susquehanna data set; the other focus parameters had even fewer long-term records.

Figure 29 shows the groundwater sampling sites in the Susquehanna data set with multiple records for barium, bromide, specific conductance, and methane. Given the lack of long-term water-quality records near HVHF wells, monotonic trend analysis was not selected for testing the groundwater hypothesis.

Table 24. Availability of groundwater sampling sites in or near the Marcellus and Utica Shale area with specific conductance records.
[**Abbreviations**: SC, specific conductance; PADEP, Pennsylvania Department of Environmental Protection; NYDEC, New York State Department of Conservation; HVHF, high-volume hydraulic fracturing]

Criteria	Number of sites meeting selection criteria for each site type and sample date					
	Site type			Sample date		
	All sites	Springs	Wells	2006 or earlier	2007 or later	Both dates
Site in/near the Marcellus and Utica Shale area	3,899	1,112	2,787	3,596	333	30
No oil-gas wells[1] within 1 mile of site[2]	3,199	827	2,372	2,930	296	27
An oil-gas well[1] within 1 mile of site	700	285	415	666	37	3
An HVHF well within 1 mile of site	290	133	157	283	7	0
10 or more SC records	43	7	36	43	8	8
10 or more SC records, and no oil-gas well[1] within 1 mile of site[2]	23	7	16	23	8	8
10 or more SC records, and an oil-gas well[1] within 1 mile of site	20	0	20	20	0	0
10 or more SC records, and an HVHF well within 1 mile of site	0	0	0	0	0	0

[1]Any oil and gas well in the inventory compiled from the online databases maintained by the PADEP and NYDEC.
[2]For potential use as a reference site.

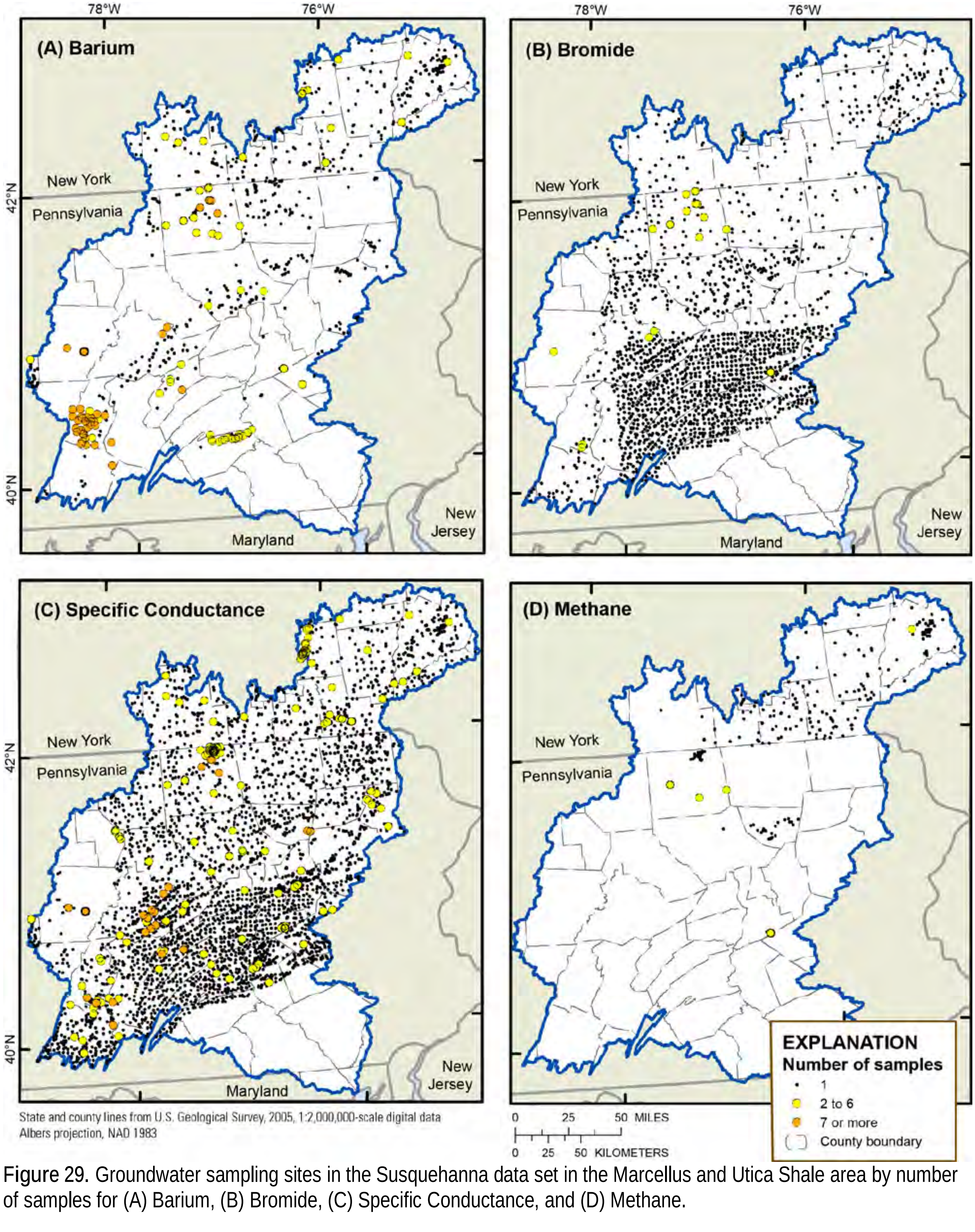


Figure 29. Groundwater sampling sites in the Susquehanna data set in the Marcellus and Utica Shale area by number of samples for (A) Barium, (B) Bromide, (C) Specific Conductance, and (D) Methane.

The second statistical design considered, called network analysis, has been used in the USGS NAWQA Program to assess changes in groundwater quality for specific parameters across networks of wells screened in specific aquifers or near the water table below specific land-use settings. Such networks have been sampled on a decadal time interval to assess changes in groundwater quality, which tend to be much slower than those observed in streams and rivers. This approach requires 25 to 30 sampling sites per hydrologic or hydro-geographic unit, such as an aquifer or ecoregion. The network analysis approach uses a two-sample comparison between two sampling events separated by a period of approximately 10 years. This method was applied and is described in the national USGS study of changes in chloride, total dissolved solids, and nitrate in groundwater (Lindsey and Rupert, 2012). In that study, samples of groundwater collected from wells during 1988-2000 were compared on a pair-wise basis to samples from the same wells collected during 2001-2010. The data set used by the USGS included over 1,000 groundwater sampling wells in 56 networks that represented major aquifers and urban and agricultural land-use areas. Statistical analysis was completed for each network rather than for individual wells. Control networks are not needed for network analysis because the first sampling event serves as the baseline to which the second sampling event is compared.

An advantage of the network analysis approach is that it can be used with as few as two samples for a given sampling site. This is particularly useful where long-term records of shale-gas parameters at sampled wells are rare. A disadvantage is that some existing data in the Susquehanna data set lack information about the source aquifer that is needed to apply the method as used in the NAWQA studies. However, it is possible to perform a network analysis if the sampling sites have similar well types, similar lithology (e.g., sandstone or carbonate), and similar depths and screened intervals. To use existing data for a network analysis, the filtered/unfiltered status of samples must be specified to be able to complete the trend analysis. Filtered samples are preferred because particulates in unfiltered samples can obscure the source of metals and make it difficult to determine whether shale gas development is the source of a change in groundwater quality, particularly for samples taken at domestic wells (Siegel et al., 2015b). On the other hand, turbidity in groundwater is a common complaint from homeowners near HVHF well pads. Another consideration is that "before" and "after" samples should be collected in the same general time frame so that the interval between sampling is relatively consistent for all sampling sites within a network. The network analysis method was selected as the most appropriate statistical design for testing the groundwater hypothesis given the lack of individual sites with data suitable for monotonic trend analysis.

The minimum data set needed to perform a network analysis can be summarized by:

- Defined network hydro-geographic unit, such as aquifer, ecoregion, or land-use type
- 25-30 sampling sites per network, including well location information and filter status of groundwater samples
- Two sampling events at each sampling site, separated by approximately 10 years

9.4 Spatial Networks

Groundwater quality is closely linked to underlying geology; each geological formation can have different ranges of naturally occurring contaminants. In order to make comparisons between data collected from different locations and times, one must first understand the hydrogeology and physical properties of the formation in which the groundwater sampling wells are located.

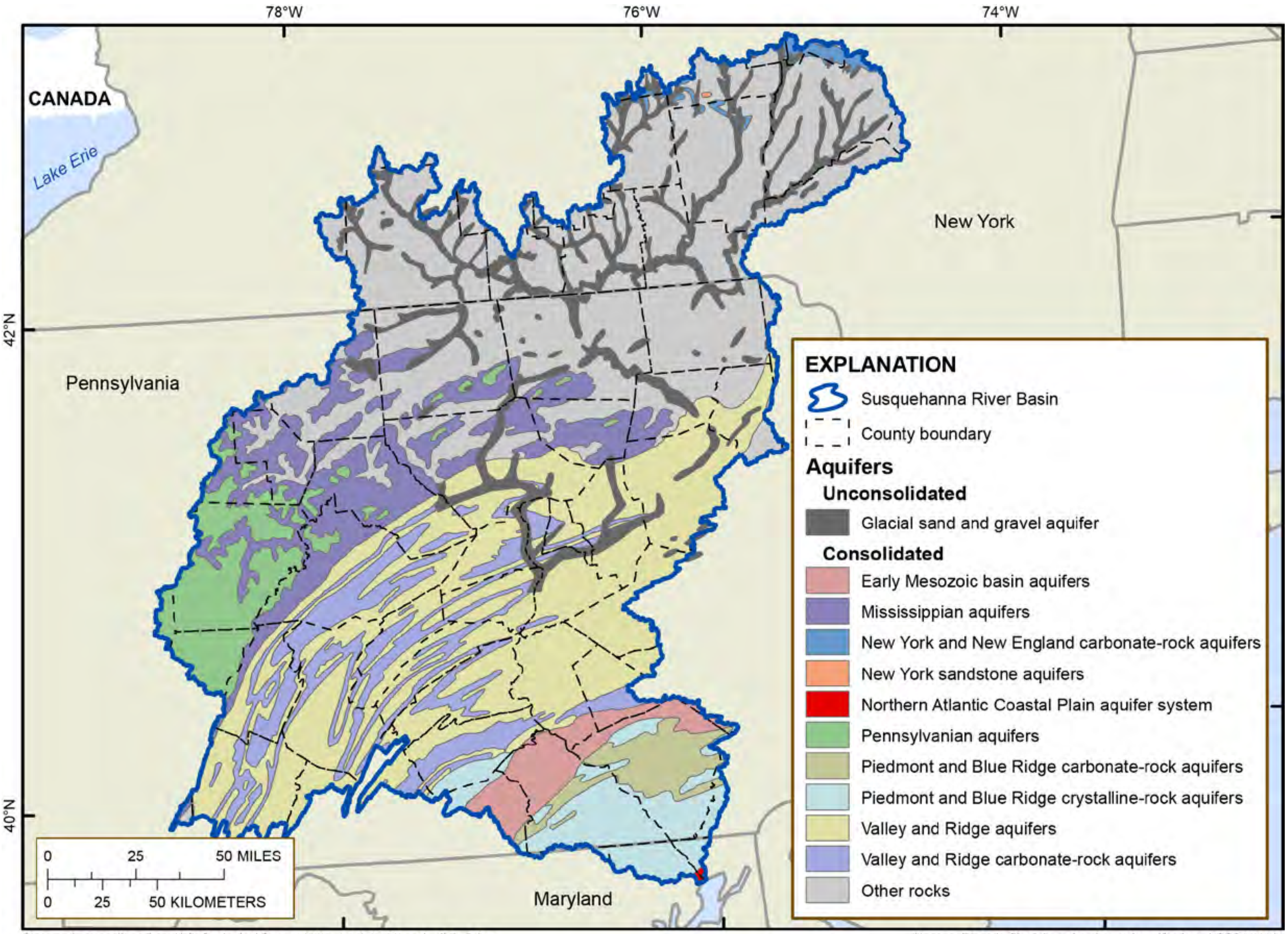
A principal aquifer is "…a regionally extensive aquifer or aquifer system that could be used as a source of potable water. An aquifer is a geologic formation, a group of formations, or part of a formation that contains sufficient saturated permeable material to yield significant quantities of water to wells and springs. Aquifers are often combined into aquifer systems (U.S. Geological Survey, 2014a)." Aquifers and principal aquifers can be used to group sampling sites from similar geologic formations to characterize groundwater quality. Figure 30 shows the principal bedrock and glacial aquifers in the Susquehanna River Basin. Many groundwater sampling sites in the Susquehanna data set did not include metadata identifying the aquifer, as shown in Figure 31, which limited the study team's ability to assess these sampling sites. In some cases, missing aquifer information was estimated based on available coordinate and well depth information (see section 5.3).

In order to interpret the origin of a parameter measured in groundwater, aquifer properties must be characterized. Aquifer properties determine the susceptibility of an individual aquifer to contamination, and include depth, porosity, hydraulic conductivity, and fracture incidence as well as water sources and stresses to the system including recharge rates, topography, rainfall, interactions with surface water, travel through the unsaturated zone, water age, water use, and well discharge. All of these factors affect the physical flow of water through the aquifer (Focazio et al., 2002). For example, recharge due to rainfall will dilute and disperse contaminants in shallow aquifers, especially in highly permeable soils.

An understanding of topographic position (uplands versus valleys) and water chemistry can also be important. For example, in northeastern Pennsylvania, different concentrations of methane can be associated with different groundwater residence and/or rock-water interaction times, and correlate with different water-types, which in turn can correlate with topographic position (Molofsky et al., 2013; Siegel et al., 2015a).

Groundwater chemistry in the Susquehanna River Basin is characterized in Appendix A of this report. Appendix A (Chapter 15) shows how methane, specific conductance, bromide, or barium vary by the main aquifers and principal aquifers in the Marcellus and Utica Shale area of the Susquehanna River Basin. An additional section on methane variability is included because numerous studies have found varied and naturally changing background levels of methane in domestic wells prior to shale gas development (Whisman et al., 2012, Coleman and McElreath, 2012).

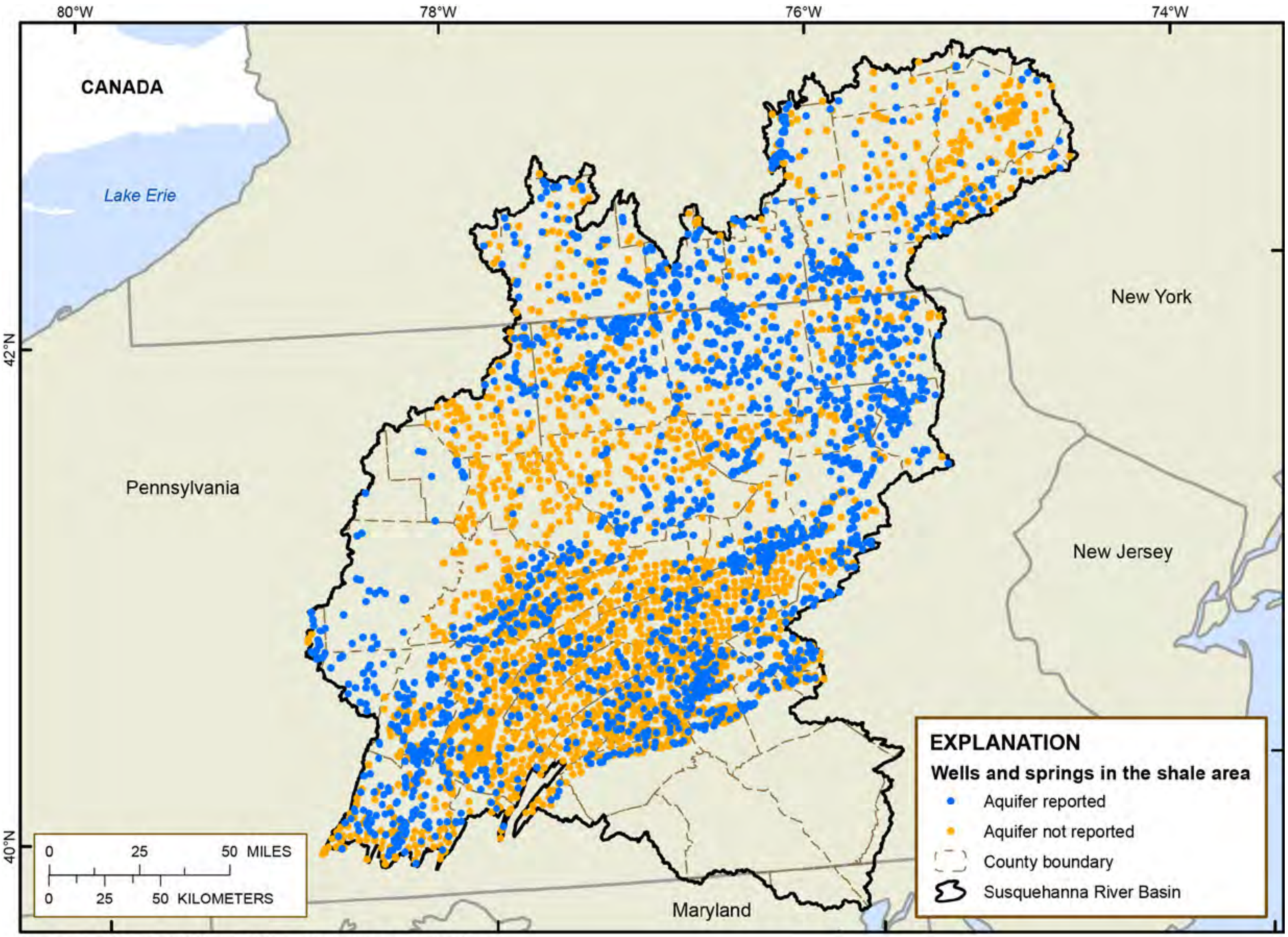
Given the variation in water quality by aquifer and topography in the Susquehanna River Basin, the combination of aquifer and topography was selected as the appropriate network hydro-geographic unit for using network analysis to determine whether groundwater quality is changing as a result of shale gas development.



State and county lines from U.S. Geological Survey, 2005, 1:2,000,000-scale digital data
Albers projection, NAD 1983

Unconsolidated: Glacial sand and gravel aquifer from USGS (2002b)
Consolidated: Principal Aquifers from USGS (2003)

Figure 30. Principal aquifers in the Susquehanna River Basin.



State and county lines from U.S. Geological Survey, 2005, 1:2,000,000-scale digital data
Albers projection, NAD 1983

Figure 31. Groundwater sampling sites in the Susquehanna data set in the Marcellus and Utica Shale area, with sites distinguished by whether or not the aquifer is reported.

Groundwater monitoring networks should be established in the most common drinking water aquifers in northeastern Pennsylvania, which include the Upper Devonian Catskill aquifer and the Upper Devonian Lock Haven aquifer (Molofsky et al., 2013; Perry et al., 2012). The Catskill and Lock Haven aquifers should each be separated into two networks, one for upland and one for valley topography, as the water chemistry is different in each type of topography (Mollofsky et al., 2013; Siegel et al., 2015a). In addition to these four networks, a fifth network is recommended for the Pleistocene deposit aquifer (also referred to as glacial sand and gravel or aquifers of alluvial and glacial origin). Warner et al., (2012) examined three aquifers including alluvium, Catskill, and Lock Haven in their study of shale gas impacts in northeastern Pennsylvania. This monitoring strategy yields a total of five networks. To the degree that this is possible, wells of the same type, preferably monitoring or domestic wells screened at similar depths, should be chosen for each network as water quality can vary between depths and across well types.

9.5 Number and Locations of Sampling Sites

The magnitude and location of a contamination event in relation to a groundwater sampling site will have an impact on concentrations measured at the site, as would the groundwater flow paths in the vicinity of the contamination source to the well. The distance, type of soil, and aquifer materials the water flows through from the contamination source to the sampling site could affect the attenuation and dilution of a spill. The direction of groundwater flow also affects whether a sampling site is in a position to detect a contamination event.

Given variability in aquifers and aquifer susceptibility, the range of contamination scenarios, and differences in how methane and brine would travel through the subsurface, there is no precise definition of the potential area of influence for a single HVHF well. The TAC selected a range of within 1 mile from a HVHF well as the potential area of influence for a HVHF well, based on professional experience and based on what other researchers have reported in the literature (Table 25). The value of 1 mile represents a middle-range estimate; the distances found in the literature range from 0.5 to 2 miles. In addition, a physical justification for the 1-mile area of influence is that horizontal laterals typically extend for distances of 3,000 to as much as 10,000 feet from the well pad (U.S. Department of Energy, 2013).

To meet the criteria for the network analyses described here, 25-30 sampling sites are needed per aquifer/topography combination, and each of those sites within the network must be within 1 mile of a HVHF well.

Table 25. Distances used by researchers to define the potential influence of a high-volume hydraulic fracturing (HVHF) well on groundwater.

Study	Comments	Region	Zone of influence cited in study – radial distance from the HVHF well (miles)
(Heisig et al., 2013)	Area beyond which a HVHF well pad is unlikely to influence groundwater chemistry.	New York	1.0
Jackson et al., 2013	Area within which a HVHF well pad may influence groundwater chemistry.	Northeastern Pennsylvania	0.6
Molofsky et al., 2013;	Area within which a HVHF well pad may influence groundwater chemistry.	Northeastern Pennsylvania	0.6
Boyer et al., 2012)	Area within which a HVHF well pad may influence groundwater chemistry.	Pennsylvania	0.5 (Phase 1; 2,500 feet) 0.9 (Phase 2: 5,000 feet)
Kresse et al., 2010	Area within which a HVHF well pad may influence groundwater chemistry.	Arkansas	2.0
Osborn et al., 2011	Area within which a HVHF well pad may influence groundwater chemistry.	Northeastern Pennsylvania and New York	0.6
Fontenot et al., 2013	Area within which a HVHF well pad may influence groundwater chemistry.	Texas	1.8

9.6 Duration, Frequency, and Timing of Monitoring

Samples for the full suite of priority groundwater parameters must be collected from time periods both before and after shale gas development at each sampling site, which is assumed to have begun in 2007 for this case study. Sampling events should be separated by approximately 10-year increments. To detect whether long-term change is occurring as a result of shale gas development, a third sampling event should be planned about 10 years in the future. Additional rounds of network sampling should be conducted beyond the 10-year window to detect long-term water-quality change. Each of the 25-30 sampling sites within each network must have the first two sampling events occur before and after shale gas development, and separated by 10 years, to meet the case study criteria.

If statistical change is detected using the network analysis, additional investigation would be appropriate to further characterize changes at individual sampling sites, including an analysis of hydrogeologic factors discussed in Appendix A (Chapter 15) and additional water-quality sampling. Additional water-quality sampling would determine whether changes found using network analysis were also evident when using monotonic trend analysis. Such an analysis may be more useful for understanding water-quality changes in site-specific conditions rather than identifying change for the entire network. The USGS samples a subset of the sampling sites used for its network analyses on a biennial (every other year) basis to determine whether or not changes identified in the network analysis are continuous on a smaller time scale (Lindsey and Rupert, 2012). A subset of five sampling sites per network is recommended for biennial sampling as part of this network analysis monitoring design.

Table 26 summarizes the groundwater data needed to use network analysis to test the groundwater hypothesis and answer the case-study policy question.

Table 26. Summary of groundwater data needed to use network analysis to determine whether cumulative shale gas development activities in the Susquehanna River Basin are resulting in water-quality change.

Criteria	Groundwater data needed
Monitoring parameters	• Suite of priority groundwater parameters from Table 3 at each monitoring site.
Spatial networks	• Minimum of 5 networks in each of the major drinking water aquifers with shale gas development, distinguished by topography: o Upper Devonian Lock Haven aquifer with upland topography, o Upper Devonian Lock Haven aquifer with valley topography, o Upper Devonian Catskill aquifer with upland topography, o Upper Devonian Catskill aquifer with valley topography, and o Pleistocene deposits aquifer.
Number and location of sampling sites	• For each network: o 25-30 sampling sites o Each site within 1 mile of a HVHF well
Duration, frequency, and timing of monitoring	• Two samples at each site, separated by approximately 10 years and taken: o before shale gas development, and o after shale gas development • Additional long-term monitoring, in subsequent 10 year increments, • A subset of 5 sites per network sampled every 2 years.

10. Groundwater Data Availability

This chapter explores the extent to which available groundwater data meet the monitoring criteria for network analysis.

10.1 Groundwater-Quality Data in the Susquehanna data set

Figure 32 and Table 27 show the number of focus parameter sampling sites from the Susquehanna data set within 1 mile of a HVHF well in the Marcellus and Utica Shale area, and the timing of data collection. Table 27 shows that while many sites met the criteria for being located within 1 mile of a HVHF well, no sites met the criteria for both pre- and post-2007 data. The sites sampled in 2007 or later were not sampled prior to that year. Methane had the least amount of existing data available. None of the focus parameters had enough “before” or “after” data to meet the criteria for network analysis.

Of the thousands of existing sampling sites, several hundred sites were identified that are located within 1 mile of HVHF wells. For these sampling sites, there are potentially enough “before” data to cover the first sampling event of a network analysis for specific conductance, barium and bromide for some aquifers. Table 28 presents the number of existing focus parameter sampling sites by the primary drinking water aquifers. One of the major challenges for monitoring agencies is selecting appropriate groundwater sampling sites; monitoring agencies do not have access to information on the location of future HVHF wells, so sampling sites with “before” data within 1 mile of HVHF wells are due more to coincidence than to planning.

Table 27. Groundwater sampling sites in the Susquehanna data set in or near the Marcellus and Utica Shale area meeting selected criteria for barium, bromide, methane, and specific conductance.
[Abbreviations: HVHF, high-volume hydraulic fracturing]

Criteria	Number of sites meeting data criteria for each parameter			
	Barium	Bromide	Methane	Specific conductance
Sites in/near the shale area	902	1,686	302	3,899
Sites within 1 mile of a HVHF well[1]	59	74	4	290
Sample date of 2006 or earlier	57	72	3	283
Sample date of 2007 or later	2	2	1	7
Sample dates of before and after 2007	0	0	0	0

[1]HVHF well in Marcellus or Utica Shale (through August 21, 2015).

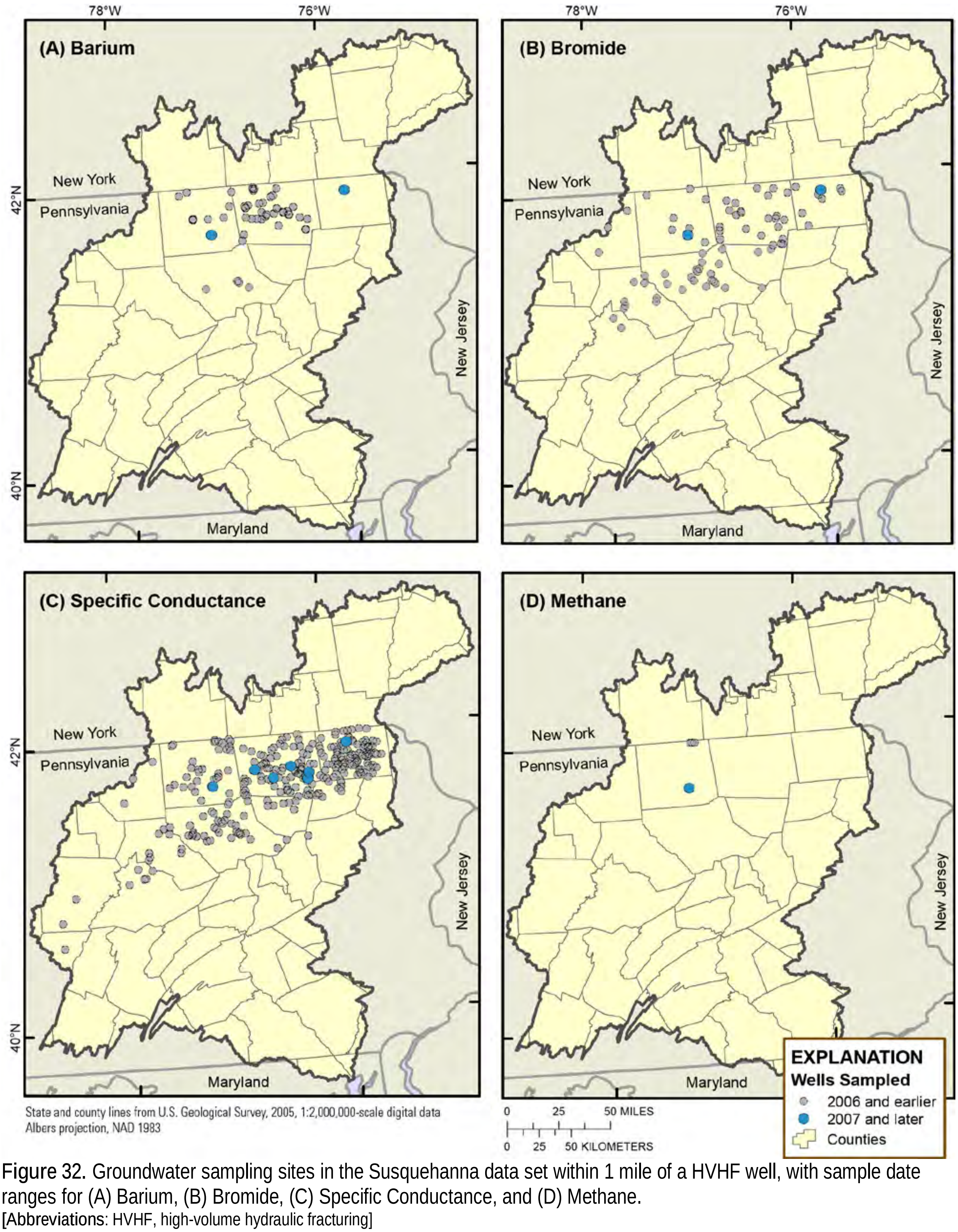


Figure 32. Groundwater sampling sites in the Susquehanna data set within 1 mile of a HVHF well, with sample date ranges for (A) Barium, (B) Bromide, (C) Specific Conductance, and (D) Methane.
[**Abbreviations**: HVHF, high-volume hydraulic fracturing]

Table 28. Groundwater sampling sites in the Susquehanna data set that are in or near the Marcellus and Utica Shale area with data for the focus parameters, available for network analysis in the primary drinking water aquifers.
[**Abbreviations:** SC, specific conductance; --, not applicable; HVHF, high-volume hydraulic fracturing]

Aquifer	Counts of groundwater sites and records for the focus parameters											
	Number of records				Number of sites				Number of sites within 1 mile of a HVHF well[1]			
	Barium	Bromide	Methane	SC	Barium	Bromide	Methane	SC	Barium	Bromide	Methane	SC
Pleistocene deposits	275	64	54	454	206	36	50	275	**43**	0	0	15
Upper Devonian Catskill aquifer	121	221	38	727	90	207	30	657	2	**31**	0	**156**
Upper Devonian Lock Haven aquifer	36	44	28	201	32	41	28	173	12	17	3	**72**

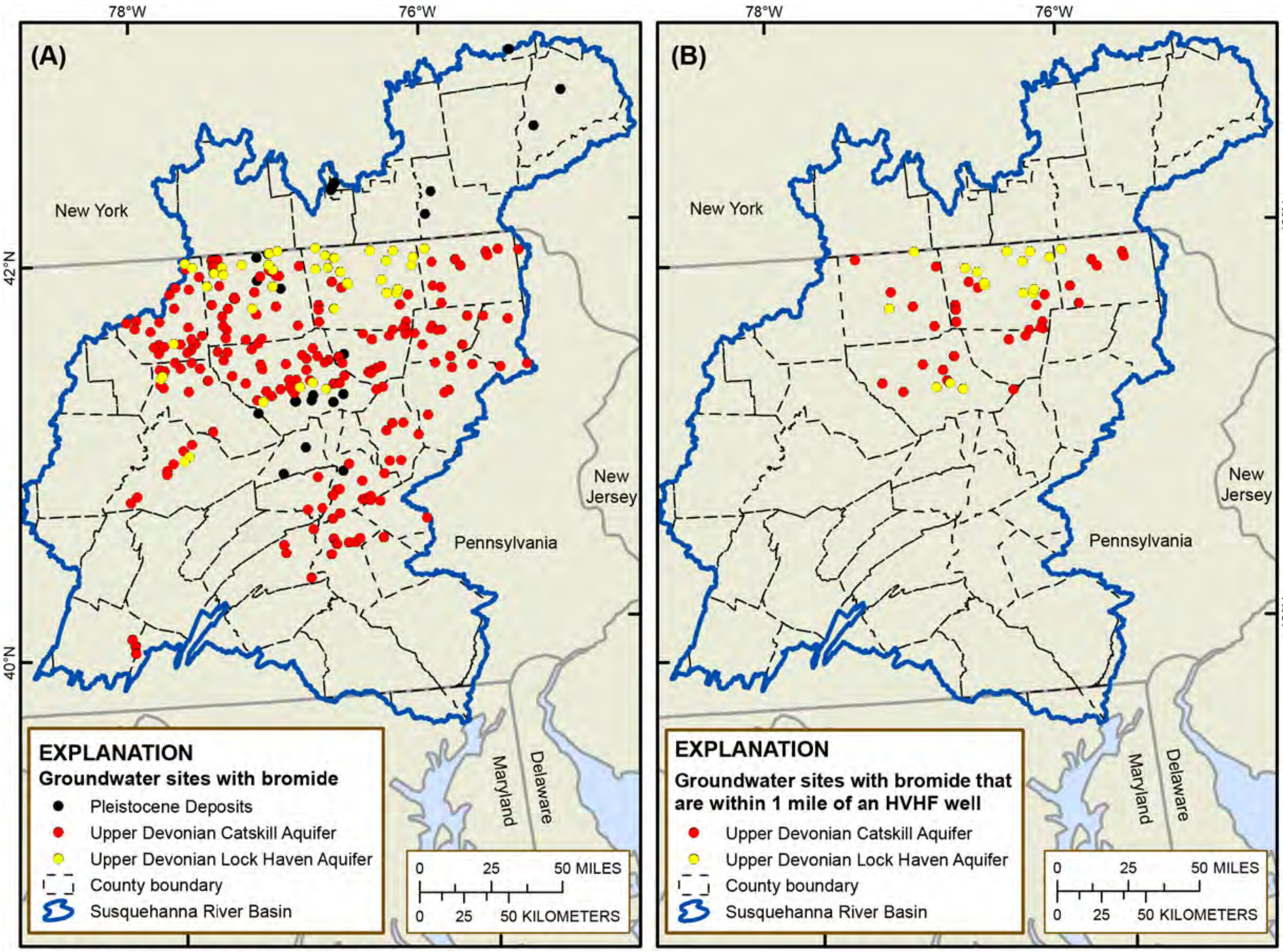
[1]Aquifers with 25 or more groundwater sampling sites are shown in bold and may have enough data for a network analysis.

There are potentially more than 25 “before” data records in the three primary drinking water aquifers that could be used as the first sampling event of a network analysis. The Upper Devonian Catskill aquifer has the most data, with 31 bromide and 156 specific conductance sites within 1 mile of a HVHF well (Table 28). The Pleistocene deposits have 43 sites with barium measurements, and the Upper Devonian Lock Haven aquifer has 72 sites with specific conductance measurements. For the recommended network analyses (Table 26), the available sites in the Upper Devonian Catskill and Upper Devonian Lock Haven aquifers would need to be further separated into upland and valley topography, making it unlikely that enough data would be available for a network analysis, even for these limited parameters. Figure 33(A) and (B) show the distribution of bromide sampling sites in the three aquifers.

As summarized in Table 26, the entire suite of priority groundwater parameters are needed to interpret water-quality data and correlate water-quality change with shale gas development. Given that the needed data are not available for all the focus parameters for any of the primary drinking water aquifers, additional analysis was not completed on the groundwater data in the Susquehanna data set. To be used to answer the case-study policy question, further scrutiny of groundwater data would be needed to determine whether samples were collected in the right decade or if the appropriate minimum detection limits are available to be used in a network analysis. Data for the full suite of priority groundwater parameters are less available than data for the four focus parameters. Although several parameters were collected at sampling sites before and after 2007, none of the individual sites in the Susquehanna data set located within 1 mile of an HVHF well have data for most of the suite of priority groundwater parameters. As shown in Figure 32, there are almost no currently available data in the Susquehanna data set collected after 2007 that can be used for the second sampling event of a network analysis.

Although there have been many domestic wells sampled for methane by the shale gas industry prior to drilling HVHF wells nearby, methane has the least amount of available data of the focus parameters in the Susquehanna data set. Methane is especially problematic given that there is significant temporal variability in methane concentrations at individual sampling sites (Whisman et al., 2012; see Appendix A (Chapter 15)) so characterizing the “before” and “after” conditions present additional statistical challenges.

In summary, none of the sampling sites in the Susquehanna data set have enough data for the first sampling event of a network analysis. Consequently, the available data in the Susquehanna data set cannot be used to support a network analysis and cannot be used to answer the case-study policy question.



State and county lines from U.S. Geological Survey, 2005, 1:2,000,000-scale digital data
Albers projection, NAD 1983

Figure 33. Groundwater sampling sites in the Susquehanna data set in the primary drinking water aquifers showing (A) All sampling sites with bromide data, and (B) Sampling sites with bromide data within 1 mile of a HVHF well.
[Abbreviations: HVHF, high-volume hydraulic fracturing]

10.2 Groundwater Data Sources Limited by Privacy Concerns

Data collected at domestic wells are often subject to confidentiality agreements and can be unavailable to researchers. This can be an issue for data collected by both public entities and private industry.

For groundwater data collected by public entities, a recurring issue is the lack of access to spatial coordinates for sites sampled by public institutions at private homes; privacy issues and confidentiality agreements severely restrict the ability to use data collected from domestic wells for scientific purposes. The Susquehanna data set includes barium samples from 2011 collected in and near the Marcellus and Utica Shale area by State University of New York Oneonta, the Penn State Marcellus Center for Outreach and Research, and Penn State University. These samples were collected with the intention of establishing baseline conditions prior to shale gas development; however, these data were provided to the project with degraded coordinates to protect the privacy of homeowners. A degraded coordinate masks the real coordinates; for example, degraded coordinates might be given to the nearest center of a 1-mile grid, not at the actual site where samples were collected. The lack of unique coordinates means these data cannot be used in this format for any analysis beyond summary statistics for the whole data set. The scientist is unable to use these data with any grouping variable, such as aquifer, which relies on a precise location. As a consequence, the Susquehanna data set includes only 35 barium records instead of 100 barium records for 2011 that can be used for analyzing spatial attributes such as aquifer or proximity to HVHF wells.

Similar issues preclude access to groundwater data collected by industry at private homes; again, privacy issues and confidentiality agreements severely restrict the ability to use data collected from domestic wells for scientific purposes.

Pennsylvania assumes oil and gas companies are liable for water contamination, unless they can demonstrate otherwise (Swistock and Rizzo, 2012). Industry typically takes water samples from nearby domestic wells prior to drilling a HVHF well to document pre-existing water-quality conditions. Industry has collected tens of thousands of samples from domestic wells prior to drilling HVHF wells near them, potentially representing a robust groundwater data set in the Susquehanna River Basin. Chesapeake Energy has collected and summarized data from over 13,000 data records from domestic wells in northeastern Pennsylvania (Siegel et al., 2015a). Molofsky et al. (2013) cite industry sampling of over 1,700 domestic wells in one northeastern Pennsylvania county.

Industry data sets include an extensive list of parameters and may resemble the suite of priority groundwater parameters recommended in this study (Table 3). Table 29 lists the parameters that are included in the Chesapeake Energy pre-drilling sampling program (Siegel et al., 2015b). Industry data sets can include sites that have been sampled both before and after shale gas development. At least one large company samples any nearby water-supply wells whenever a new HVHF well is drilled, even if the well has been previously sampled in connection with earlier drilling and hydraulic fracturing. Pre-drill data from Chesapeake Energy was published in Reese et al. (2014). Chesapeake Energy is just one of multiple companies that have collected groundwater data in the Susquehanna River Basin; this publication may lead the way for other companies to begin sharing the groundwater data they have collected.

The PADEP maintains paper copies of each groundwater monitoring report that is collected from domestic wells by industry. However, there are practical obstacles preventing access to these data. The PADEP currently does not have the data available in an electronic database, although PADEP is updating its regulations to require that future submittals be made electronically. To obtain the data, a formal request must be submitted under Pennsylvania's Right-to-Know Law, Section 404, 65 P.S. §67.101 et seq. To comply with the request, PADEP staff makes photocopies of individual reports, and the homeowner name and address are redacted due to privacy concerns. Latitude and longitude, when provided in the report, would be provided to the person making the Right-to-Know-Law request. Latitude and longitude may be included in the field notes section of the report, but it is not clear how frequently those data are available. Collaboration with industry may be an alternative to going through the PADEP to obtain access to industry data. Precedents for such collaboration exist, and several groups have established confidentiality agreements to use industry data sets for characterizing groundwater conditions in the region prior to shale gas development (Baldassare et al., 2014; Reese et al., 2014; Boyer et al., 2012; Molofsky et al., 2013; Siegel et al., 2013; Siegel et al., 2015a).

Figure 34 represents an unusually high density of HVHF wells in proximity to domestic wells in the Susquehanna River Basin. The blue triangle in the middle is a sampling site from the Susquehanna data set and the blue circles are domestic wells, for which monitoring data are not available in the Susquehanna data set. Figure 34 illustrates the small percentage of groundwater sampling sites with water-quality data in the Susquehanna data set; presumably many of these domestic wells would have been sampled by industry prior to drilling some of the HVHF wells.

Table 29. Parameters included in the Chesapeake Energy pre-drilling sampling program (Siegel et al., 2015b).

Parameter	In Table 3 Suite of priority groundwater monitoring parameters	In Table 3 but not Chesapeake Energy pre-drilling sampling program
Alkalinity	X	
Arsenic		
Barium, dissolved	X	
Barium, total		
Benzene	X	
Bromide	X	
Cadmium		
Calcium	X	
Chloride	X	
Chromium		
Ethane		
Ethylbenzene	X	
Hardness		
Iron		
Lead		
Lithium	X	
Magnesium	X	
Manganese, dissolved		
Manganese, total		
Mercury		
Methane		
Oil and Grease Hexane Extractable Material (HEM)		
pH	X	
Potassium	X	
Propane		
Selenium		
Silver		
Sodium	X	
Specific Conductance	X	
Strontium	X	
Temperature	X	
Sulfate	X	
Sulfur		
Surfactants		
Toluene	X	
Total Dissolved Solids	X	
Total Suspended Solids		
Turbidity	X	
Xylenes, total	X	
Dissolved Oxygen		X
Gross alpha		X
Gross beta		X
Nitrate		X
Radium-226, -228		X
Uranium		X

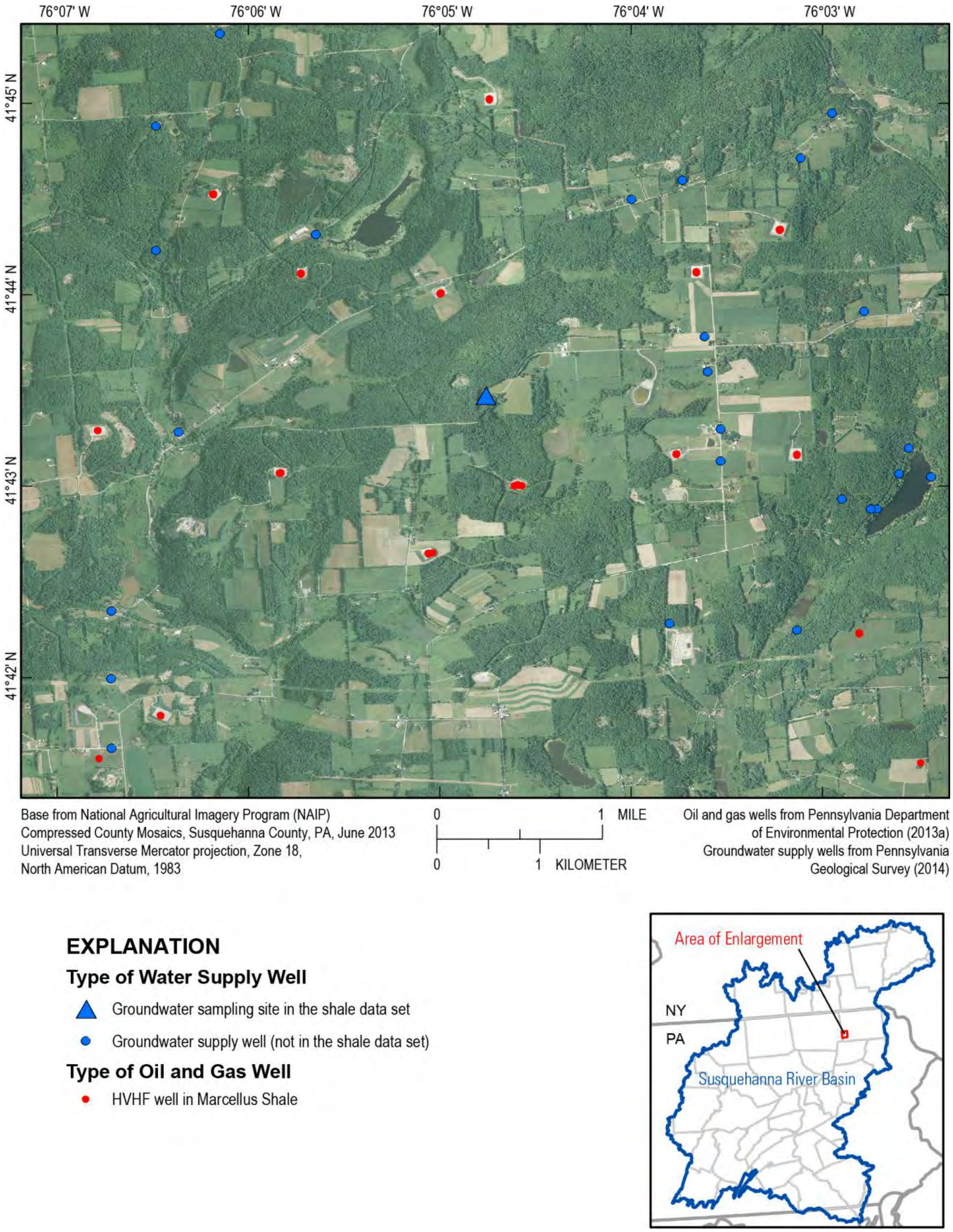


Figure 34. Example of a groundwater sampling site and nearby HVHF wells in the Marcellus Shale (as of 8/21/2015), southwestern Susquehanna County, PA.
[Abbreviations: HVHF, high-volume hydraulic fracturing]

10.3 Data Consistency and Quality Assurance – Groundwater

The groundwater sampling sites with relevant data for a limited number of parameters are collected primarily by the USGS. USGS data from NWIS usually have the aquifer identified, and the USGS Office of Groundwater recommends that the code that indicates the aquifer(s) of the sampling site be entered into the NWIS database when a new site is established (U.S. Geological Survey, 1998b).

Overall, the USGS stresses objective and replicable data collection procedures that are transparent and peer reviewed. Data that are collected for publication must be documented to describe the methods or techniques used to collect, process, and analyze the data. For information on methods used to collect and quality assure water-quality samples see the USGS National Field Manual available at *http://water.usgs.gov/owq/FieldManual/*. For documentation of specific USGS laboratory or field analytical methods and related quality assurance/quality control procedures see *http://water.usgs.gov/owq/methods.html*.

10.4 Groundwater Data Usability

As stated in section 7.6, data collected for one monitoring objective may not be directly applicable to another objective, due to the location of monitoring sites, frequency of monitoring, parameters measured, and sample fraction analyzed. The groundwater quality records identified through this case study were generated by 10 organizations that collected groundwater data in the Susquehanna River Basin. Significant time and effort were required to locate, obtain, and format water-quality records from multiple organizations that use different sampling plans and data documentation practices. Substantial project time and effort over the course of this multi-year project were required to locate, obtain, and consistently format data. Information on the aquifer from where groundwater samples are taken is particularly important for this case study and is missing for a substantial number of groundwater sampling sites (Figure 31).

The Water Quality Portal (National Water Quality Monitoring Council, 2014a) includes data collected at 45 percent of the groundwater monitoring sites identified through this case study in the Susquehanna River Basin, and 61 percent of the water-quality data records are available through the Portal. The only groundwater data in the Susquehanna data set that are available through the Water Quality Portal are available through NWIS (U.S. Geological Survey, 2002a). In the case of groundwater data, the most important data sets are generally not available, but recent publications indicate this might be changing (Reese et al., 2014). Any groundwater data analysis to identify contamination from shale gas development will benefit from greater accessibility to these data.

11. New Groundwater Data Needed to Answer the Case-Study Policy Question and Associated Costs

As discussed in Chapter 10, no set of groundwater sampling sites in the Susquehanna data set has data that meet the case study criteria for groundwater (Table 26). This chapter explores options for filling data gaps and provides estimates of costs to fill those gaps. The following recommendations for filling data gaps explore how to make the most of available data, and represents an assessment of the minimum data needed to fill the groundwater data gaps.

11.1 Groundwater Data Needs

For groundwater, none of the existing sampling sites in the Susquehanna data set have the needed water-quality data to support or complete a network analysis for the suite of priority groundwater parameters. Every sampling site is missing at least one of parameters, location, or timing. Building on existing data from the Susquehanna data set would ideally be the first option for filling the data gaps, but this is not a viable option because there are no existing sampling sites with data for the full suite of priority groundwater parameters prior to HVHF development (Table 3). Sampling for more parameters at existing sampling sites will not fill the gap for sampling prior to shale gas development. Consequently, two options for working with industry data were explored as viable means for filling groundwater data gaps. The collection of relevant groundwater data depends on accurate information regarding future HVHF well development plans that can only be obtained from the shale gas industry. Any groundwater monitoring plan that is developed without the involvement of the shale gas industry will be inefficient and leave to chance whether the needed groundwater data will be collected.

The first groundwater monitoring option relies on existing industry water-quality data. The industry data set includes the most comprehensive existing set of “before” shale gas development water-quality data located close to HVHF wells, and it includes many of the priority groundwater parameters (Table 29). If industry data sets include most of the suite of priority groundwater parameters, specify whether collected samples are filtered or unfiltered, and include well location information, it may be possible to use them for the “before” sampling event of a network analysis. In this case, 25 to 30 “before” sampling sites within 1 mile of HVHF wells would need to be identified for each of the priority networks (Table 26) from the industry data set. A second sampling event after shale gas development could be conducted at the same group of sampling sites approximately 10 years after the first sampling event. The network analysis of the two data sets would show whether or not there were changes in concentrations for the suite of groundwater parameters between the sampling events. As specified in Table 26, a subset of five sites per network would be monitored every 2 years, and another network-wide sampling event would be planned 10 years in the future.

If access to existing industry data cannot be obtained or if the available data are not adequately documented for the analysis described in section 9.3, the second option for filling the groundwater data gaps would be to collect new “before” and “after” data working both with industry to identify appropriate sampling sites, and with homeowners to collect data from private wells. In addition to working with industry to identify homeowners that are near HVHF wells that will be developed soon, industry would also be involved in the development of appropriate disclosure agreements. This option would involve collecting both “before” and “after” samples for the suite of priority groundwater parameters at 25 to 30 sampling sites within 1 mile of HVHF wells for each network analysis. If this option is used it is likely that both industry and the monitoring agency would both be collecting “before” samples, because industry is likely to continue sampling before drilling for liability protection (see section 10.2) even if other monitoring agencies collect samples at the same location. This second monitoring option would collect the same data identified in the first option, plus the new set of “before” data. Table 26 summarizes the data that must be collected to fill the groundwater data gaps. The first option includes the entire table except for the first sampling event prior to shale gas development; the second option includes the entire contents of the table.

Confidentiality agreements would need to be worked out with industry, and new agreements developed with homeowners located near existing or soon-to-be-drilled HVHF wells for both of these options. Working with industry-collected water-quality data would require significant effort in negotiations and legal agreements with both industry and homeowners. Both industry and homeowners need to be willing to share their data and coordinates so the wells can be mapped, for groundwater monitoring to answer the policy question to be successful.

The source aquifer of each sampling site must also be known for industry data to be used as part of a network analysis. A drilling record (driller’s log) typically provides the information needed to determine the rock formation from which the well produces water, which could be available from the homeowner or the Pennsylvania Geological Survey. Information obtained from a driller’s log can include items such as well depth, depth to the open interval, type of open interval (open-hole or screened; a screened interval would correspond to unconsolidated deposits), and descriptions of the sediment and rock formations that were drilled through. In some instances, the aquifer may even be identified. If a well log with aquifer information cannot be located for a domestic well, it is recommended that the well not be used for network analysis.

11.2 Cost of Groundwater Monitoring

Costs were estimated for filling the groundwater data gaps, based on the set of monitoring recommendations from section 11.1. Decision makers may decide they want additional sampling sites to support their policy decisions if they decide the minimum data set is not adequate to inform their decisions. As with surface water, groundwater monitoring costs should be considered in the context of the regional importance of groundwater as the primary drinking water source in the Marcellus and Utica Shale area of the Susquehanna River Basin. The economic value of groundwater is difficult to estimate, and once groundwater is contaminated it is very difficult to treat. Proactive monitoring to detect low-levels of potential contamination can identify concerns and remedial options prior to significant impact to freshwater aquifers.

The cost of monitoring to fill groundwater data gaps presented here relies on data provided by the USGS. Table 30 shows average groundwater-related sampling costs for 17 groundwater networks that were sampled as part of the USGS NAWQA Program in Fiscal Year 2013 (October 1, 2012-September 30, 2013), and adjusted by 2.3 percent for inflation to October 2015 (U.S. Department of Labor Bureau of Labor Statistics, 2015). Each of the 17 NAWQA networks contains 25 to 30 wells; the total number of sampling sites is about 500. The number of wells in a network varies because sometimes the goal of 30 wells cannot be met. The studies were a mix of NAWQA Land-Use Studies (monitoring sites) and Study Unit Survey networks (primarily domestic wells), and all sampling was conducted by USGS personnel. Sampling costs include salary, travel, supplies, postage, equipment, miscellaneous and administrative services. The average cost per well per sampling event as of October 2015 is $4,300 and the range of cost per site is $3,300 to $5,400.

Table 30. Average sampling costs associated with groundwater monitoring of 17 USGS National Water-Quality Assessment network studies comprised of about 500 monitoring sites[1], as of October, 2015.

Type of cost	Average cost per well per sampling event	Average cost per network per sampling event	Percent of total cost by category
Salary cost estimate	$2,312	$64,741	53.3%
Vehicle cost estimate	$93	$2,593	2.1%
Travel cost estimate	$232	$6,498	5.3%
Postage cost estimate	$71	$1,973	1.6%
Supply cost estimate	$249	$6,958	5.7%
Miscellaneous cost estimate	$24	$665	0.5%
Equipment cost estimate	$34	$939	0.8%
Total, plus administrative services	**$4,341**	**$121,544**	**100.0%**

[1]Trends network, 25-30 wells per network; average number of wells = 28

Table 31 shows the total laboratory cost for the suite of priority groundwater parameters (Table 3) for Fiscal Year 2016, based on USGS National Water-Quality Laboratory costs. The total laboratory cost is $921 for these parameters. Adding in a factor of 14.8 percent ($136) to account for the cost of quality control samples yields a total laboratory cost of $1,057 per sample. These costs are higher, but not substantially different than costs provided by Penn State Extension, which provides an estimate of costs to measure a suite of parameters that are specific to shale gas development in groundwater by independent laboratories: “Packages typically range from about $200 for the most basic tests to over $900 for extensive lists (Swistock and Rizzo, 2012).”

Table 31. Parameter analysis costs at USGS laboratories and reporting levels for the suite of priority groundwater parameters in Table 3, Fiscal Year 2016.
[Abbreviations: µg/L, micrograms per liter; mg/L, milligrams per liter, pCi/L, picocuries per liter; TDS, total dissolved solids]

Parameter Name	Parameter Code	CAS Number	Reporting level	Unit	Cost per analysis[1]
Barium, dissolved	01005	7440-39-3	0.25	ug/L	$27[2]
Barium, total	01005	7440-39-3	0.25	ug/L	$7
Bromide	71870	24959-67-9	0.03	mg/L	$22
Calcium	00915	7440-70-2	0.022	mg/L	$8
Chloride	00940	16887-00-6	0.02	mg/L	$14
Gross-alpha radioactivity	62636	12587-46-1	3	pCi/L	$75
Gross-beta radioactivity	62642	12587-47-2	4	pCi/L	$75
Lithium	01130	7439-93-2	0.1	ug/L	$8
Magnesium	00925	7439-95-4	0.011	mg/L	$8
Methane, dissolved	68831	74-82-8	0.001	mg/L	$170
Nitrogen-nitrate + nitrite	631	NA	0.04	mg/L	$9
Potassium	00935	7440-09-7	0.004	mg/L	$8
Radium-224	50833	13233-32-4	1	pCi/L	$245
Radium-226	09503	13982-63-3	1	pCi/L	
Radium-228	81366	15262-20-1	1	pCi/L	$135
Sodium	00930	7440-23-5	0.06	mg/L	$8
Strontium	01080	7440-24-6	0.2	ug/L	$8
Sulfate	00945	14808-79-8	0.02	mg/L	$14
Residue on evaporation, 180 degrees Celsius (TDS)	70300	NA	20	mg/L	$15
Uranium, natural	22703	7440-61-1	0.014	ug/L	$8
Benzene	34030	71-43-2	0.04	ug/L	$227
Ethylbenzene	34371	100-41-4	0.03	ug/L	
Toluene	34010	108-88-3	0.02	ug/L	
o-xylene	77135	95-47-6	0.03	ug/L	
m- and p-xylene	85795	179601-23-1	0.05	ug/L	
Total					**$921**

[1] Total sample cost is $921. Prices apply to Fiscal Year 2016 (Oct 1, 2015-Sept 30, 2016)
[2] The price shown for barium analysis includes a specific conductance measurement ($10.58; this measurement is used to screen for high salinity which could bias the results of this or any other trace element analysis) and also includes the ICP-MS instrument set up cost ($9.08). These are one-time charges that are not applied to any other major inorganic or trace element analyses requested.

The total laboratory cost of $1,057 was added to the average sampling costs per well from Table 30 ($4,341) to arrive at an estimate of cost per sample and the total cost per sampling event per network for the suite of priority groundwater parameters (Table 32). Minimum and maximum costs are also presented.

Table 32. Estimates of average, minimum and maximum total cost per groundwater sample and cost per sampling event per network as of October 2015.

Scenarios	A. Water Science Center Monitoring Costs per sample	B. Lab costs and QC per sample	C. Total per sample (A+B)	D. Network total per sampling event (30 samples)
Average	$4,341	$1,057	$5,398	$161,900
Minimum	$3,307	$1,057	$4,365	$130,900
Maximum	$5,384	$1,057	$6,441	$193,200

Table 32 shows that the average total cost per sample would be about $5,400. This cost translates into an estimate of $161,900 for one sampling event in one network of sampling sites, assuming 30 groundwater sampling sites in a network. The range of costs may vary per sample depending on the network and on salary and overhead-related costs; the range of costs for different networks in different aquifer systems would be roughly $4,400 to $6,400 per sample based on the specific sites that were analyzed in Table 30.

Additional costs are necessary for network design, mainly salary and travel, to include tasks such as contacting homeowners and obtaining permission to sample, collecting and evaluating homeowner well construction data to find suitable sampling sites, and field reconnaissance to locate private wells and make sure the plumbing is suitable for water-quality sampling. These tasks could take several months and could add another $20,000 to $25,000 to the overall cost per network (T. Arnold, U.S. Geological Survey, written commun., August 2014).

The costs for filling the groundwater data gaps were estimated by applying the average costs from Table 32 to the monitoring program options identified in section 11.1; these costs are summarized in Table 33. The first option uses industry data to populate the "before" sampling event of the network analyses. The cost for this option includes the design cost for the network, biennial monitoring, and the "after" network sampling event that would happen in year 10. The cost presented in Table 33 assumes 3-percent annual inflation, for a 10-year total cost of $1,809,000 and average annual cost of $180,900.

The second option includes conducting both "before" and "after" network sampling events for the suite of priority groundwater parameters for each network, design cost, and biennial monitoring. The 10-year total cost of option 2 is $2,618,000 and the average annual cost is $261,800.

It should also be noted that sufficient characterization of methane temporal variability (see Appendix A (Chapter 15)) would likely require multiple "before" and "after" samples. It was impossible to quantify the number of methane samples needed to characterize baseline variability due to the lack of available data. The costs presented here represent the minimum sampling costs to fill the groundwater data gaps. Additional methane samples are likely needed to fill this data gap at each sampling site selected for a network analysis and methane should be included in the biennial sampling.

Table 33. Estimated cost for a groundwater monitoring program to detect whether cumulative shale gas development activities in the Susquehanna River Basin are resulting in groundwater-quality change, using average costs from Table 32 and assuming 3-percent annual inflation.

Monitoring program description	Design cost	Sampling event cost for five networks, including inflation	Biennial sampling cost over 10 years	Ten-year total cost	Average annual cost
Option 1: Conduct biennial monitoring and "after" sampling event. Rely on industry data for "before" sampling event.	$125,000	$1,056,000	$627,100	$1,809,000	$180,900
Option 2: Conduct "before," biennial, and "after" sampling.	$125,000	$1,866,000	$627,100	$2,618,000	$261,800

11.3 Cost of Data Analysis

As mentioned in section 8.3, water-quality data on their own do not provide information to answer policy questions; data analysis is the critical step of converting data to information. The cost of data analysis must be factored into the cost of new monitoring because data without analysis provide very little value. New data collected through the monitoring options described here should be analyzed biennially to understand ongoing trends in the data and to support efficient use of monitoring resources through adaptive management. As described in section 8.3, the dollar value obtained by doubling the costs in Table 33 would cover data collection, data management, and data analysis. Based on using Table 33 option 2, it would cost an average of $523,700 per year to fill the data gaps and analyze groundwater quality trends related to shale gas development in the Susquehanna River Basin. Although the estimated annual cost of filling the data gaps and associated data analysis for groundwater are substantially lower than for surface water, the challenges of accessing industry shale gas development plans and developing disclosure agreements with both industry and homeowners may be equivalent if not more substantial barriers compared to the cost for collecting the groundwater data needed to answer the case-study policy question.

12. Recent Developments

There are a number of recent developments in the Susquehanna River Basin that should be mentioned, including data collection and monitoring programs. Some of these developments are exploring existing data and some are collecting new data. The new data sets described here were not available in time to be included in the Susquehanna data set (Chapter 5). Emerging technologies that may improve trend detection are also described.

12.1 Data Collections

The Shale Network is a project funded by the National Science Foundation to help scientists and citizens store data for water resources that may be affected by shale gas development. The primary focus of the Shale Network currently is the Marcellus and Utica Shales in the northeastern United States. The Shale Network's goal is to find, organize, and upload data for water resources for online publication (Shale Network, 2014). The Shale Network exists to enable the generation of knowledge from water chemistry and flow data collected in areas of shale gas development. Project managers are working with the Consortium of Universities for the Advancement of Hydrologic Sciences, Inc. (CUAHSI) to create this database.

12.2 Surface-Water Quality Monitoring Programs

Lock Haven University's Geology Department has partnered with several volunteer groups and organizations to monitor surface water in the vicinity of HVHF wells. The groups include the Clearfield and Centre County chapters of the Pennsylvania Senior Environmental Corps, Centre County Conservation District, the Beech Creek Watershed Organization, and the South Renovo Borough Water Supply System. The groups monitor several parameters, including specific conductance, pH, total suspended solids, total dissolved solids, barium, copper, iron, and sulfate, and have several multi-year data sets approaching 4 years in length (M. Khalequzzamen, Lock Haven University, written commun., May, 2014).

12.3 Groundwater Quality Monitoring Programs

An ongoing USEPA effort, *Assessment of the Potential Impacts of Hydraulic Fracturing for Oil and Gas on Drinking Water Resources,* explores the potential impacts of shale gas development on drinking water and groundwater (U.S. Environmental Protection Agency, 2015a). In this study, the USEPA completed five "retrospective case studies," in which there had been reports of drinking water contamination in close proximity to past hydraulic fracturing operations. One of the retrospective case studies was located in Bradford and Susquehanna Counties in Pennsylvania where domestic wells were reported to be contaminated (U.S. Environmental Protection Agency, 2015c). The USEPA collected water-quality data in support of these case studies for most of the parameters listed in Table 3, but those data are not available through the Susquehanna data set.

Syracuse University has developed a baseline sampling program to characterize potable groundwater in the Southern Tier of New York State and plans to develop a publicly accessible database of the results (Lautz et al., 2013).

In late fall 2014, the USGS and the PADEP initiated a new long-term ambient groundwater-quality monitoring network. Samples will be collected twice a year from monitoring sites with dedicated sampling pumps. Two rounds of samples have been collected to date at 17 monitoring sites, several of which are in the Marcellus and Utica Shale area of the Susquehanna River Basin. There are plans to add 8 more monitoring sites in 2015 (D. Risser, U.S. Geological Survey, written commun., Aug, 2015).

In summer 2014 the USGS collected about 75 groundwater samples at randomly selected domestic wells in Lycoming County, Pennsylvania to characterize the quality of groundwater from aquifers used by private domestic wells (U.S. Geological Survey, 2015).

The USGS is planning a groundwater study in the Marcellus Shale for summer of 2017. This study will use a variation on the network analysis approach that is being explored for feasibility for future groundwater studies. Thirty domestic wells will be sampled; 20 located within 1 kilometer (0.62 miles) of oil and gas wells, and 10 domestic wells located farther than 1 kilometer from oil and gas wells to represent background water quality (G. Rowe, U.S. Geological Survey, written commun., November 2015).

12.4 New Technology and Monitoring Techniques

Heilweil et al. (2013) show that impacts on groundwater quality from shale gas development may be evaluated at the watershed scale by sampling dissolved methane along “gaining” streams—streams that have a net inflow of groundwater. Gaining streams can provide an integrated signal of relatively large groundwater capture areas, in contrast to the point-specific nature of groundwater sampling wells. The authors have demonstrated the utility of stream-gas sampling as a reconnaissance tool for quantifying methane inputs into streams from groundwater sources.

New technology may allow for innovations in how water sampling is conducted. For example, researchers are attempting to build a continuous groundwater monitor that would allow real-time monitoring of methane gas (Barton, 2014). This monitor would be useful for obtaining spatial, temporal, and vertical profiles on methane variability representing pre- and post-shale gas development conditions, and may also be useful as a tool for testing well-casing integrity. Another example is the development of continuous samplers using enzymes for organic parameters: a methanol sensor is being developed using an optical enzymatic approach for quantifying the target constituent (Reardon, 2014). Because a selective enzyme is used, a sensor using this technology can only measure one target constituent at a time. These technologies are still under development, but they have the potential to solve a considerable challenge in detecting variability in groundwater quality.

The continuous and discrete data collected by the SRBC may be able to support new analytical techniques for identifying water-quality impacts of shale gas development. Measurement of major ions and shale gas indicator parameters could be used to develop regression models that relate changes in ion chemistry (both individual and/or combinations of specific ions) to changes in specific conductance. The incorporation of HVHF well density as an independent variable in the regression could determine if shale gas development is potentially a causative factor. Long-term, multi-year monitoring could potentially discern cumulative effects of shale gas development via trends in specific conductance and related chemistry, particularly in the smaller watersheds where changes in specific contaminant sources or human activities are more readily tracked.

12.5 Data Sharing and Data Standards

There are data-sharing efforts underway to bring water-quality databases from multiple agencies together at the regional to national scale and to standardize reporting and metadata. The STORET Data Warehouse is USEPA's repository of water-quality-monitoring data (U.S. Environmental Protection Agency, 2014). Data in the STORET Data Warehouse can come from States, Tribes, watershed groups, Federal agencies, volunteer groups, and universities. These organizations can upload data to STORET using the Water Quality Exchange (WQX) framework, which defines a standard set of data elements and internet protocols for submitting data to the USEPA. The USGS NWIS database (U.S. Geological Survey, 1998a, 2002a, 2014c) stores water-quality data collected by the USGS.

The Water Quality Portal (National Water Quality Monitoring Council, 2014a) is a cooperative service that integrates publicly available water-quality data from NWIS, STORET, and the USDA ARS Sustaining The Earth's Watersheds - Agricultural Research Database System (STEWARDS). The Water Quality Portal provides an interface that allows a user to download water-quality data from these systems in a consistent format. These efforts should continue to be supported and enhanced. There are also efforts to aid in comparison of analytical methods used by different organizations; the National Environmental Methods Index (NEMI) website offers a searchable database that allows scientists and managers to find and compare analytical and field methods for all phases of environmental monitoring (National Water Quality Monitoring Council, 2014b). The CUAHSI Water Data Center (Consortium of Universities for the Advancement of Hydrologic Sciences, Inc., 2014) is another data sharing effort that is collecting water-quality data that relate to shale gas development in the Marcellus Shale region.

The National Ground-Water Monitoring Network (NGWMN) is a cooperative groundwater data collection, management, and reporting system based on data from selected groundwater monitoring wells from Federal, State and local groundwater monitoring networks across the nation (Advisory Committee on Water Information, 2015). The USGS is working to increase participation in the NGWMN during 2016.

There are challenges and costs for developing and maintaining integrated data systems, as well as for monitoring organizations that choose to participate in these systems. Individual monitoring agencies must define, adopt, and maintain data-management standards and move legacy data systems to information-management systems that are compatible with WQX in order to participate.

13. Findings and Recommendations

This case study explored water data collected by monitoring agencies throughout the Susquehanna River Basin to determine whether water-quality data are available to answer the case-study policy question, "Do shale gas development activities contaminate surface water or groundwater?" Overall, this case study found that the water-quality data needed to answer the case-study policy question are not available. Although more than 960,000 surface-water records were collected at about 14,700 monitoring sites over the last 85 years in the Susquehanna River Basin, there are no surface monitoring sites that met all the case-study criteria for monitoring location, parameters, streamflow, sampling frequency, and sampling longevity. There are surface-water monitoring programs that can be enhanced to collect the needed water-quality data; on the other hand, substantial new groundwater monitoring is needed to answer the case-study policy question.

The case-study findings relative to water data needed, water data available and usable, and approaches for filling the data gaps are summarized below.

13.1 Water Data Needed to Answer the Policy Question

- **Water data must satisfy study design criteria to be used to answer the case-study policy question.**

The right water data must be available in the right locations with the right supporting information to detect water-quality change and identify the cause of that change, as presented in Figure 9. The selection of appropriate monitoring sites is critical for answering the case-study policy question. For surface water, monitoring sites must be located in watersheds with HVHF wells and in reference watersheds in each ecoregion with active or planned shale gas development. Monitoring sites in these types of watersheds allow for the detection of water-quality changes in watersheds with HVHF well development and comparison with undeveloped watersheds to identify whether water-quality changes are resulting from HVHF development. The study design for groundwater again requires monitoring sites in the right places; networks of groundwater sampling sites are needed with each sampling site located within 1 mile of an HVHF well.

Next, water-quality and streamflow data at the appropriate monitoring sites must be available with sufficient sampling frequency and duration to evaluate trends in concentration over time. Finally, data on shale gas development, geology, climate, and other changes in land use in the monitored area must be available to correlate water-quality change with shale gas development activity. Without this information, the relationship between shale gas development and water quality cannot be evaluated, even if shale gas development is causing water-quality change.

- **Multiple lines of evidence, using a suite of priority monitoring parameters, are needed to identify shale gas development as the source of water-quality change.**

No one parameter can identify whether shale gas development is the source of contamination if a change in that parameter concentration is detected. By monitoring a suite of parameters, more information is available to identify the likely source of contamination. Each shale gas development activity, from initial well pad development to production of gas from a completed well is associated with different pathways that contribute different potential contaminants, so monitoring for just one of the parameters would miss multiple types of potential contamination.

Consequently, a suite of water-quality parameters is needed to determine if contamination from the cumulative impact of shale gas development activities has occurred in the Susquehanna River Basin. The suites of priority parameters for surface water and groundwater are based on the specific hydrology, geology, past and current land use, and other environmental concerns expressed in the Susquehanna River Basin.

- **The sampling frequency and duration of monitoring must meet minimum requirements to adequately characterize and detect changes in priority monitoring parameters related to shale gas development.**

Assuming that monitoring sites located in watersheds with HVHF wells and in reference watersheds and the needed ancillary data are available (Figure 9), surface-water quality and streamflow data requirements can be characterized as shown in Table 11. As mentioned above, the entire suite of surface-water parameters are needed to identify whether shale gas development is the source of water-quality change. Monthly sampling frequency is needed to detect changes in water quality year round and to minimize the time needed to detect statistically significant water-quality change at each monitoring site (see section 6.4). A minimum of eight surface-water monitoring sites are needed: one monitoring site in a watershed with HVHF wells and one reference watershed monitoring site is needed in each of the four ecoregions with active or planned HVHF development. Additional monitoring sites will provide critical information regarding the scope and magnitude of potential water-quality change associated with shale gas development, especially in watersheds with the highest density of HVHF wells and nested watershed monitoring sites.

The minimum monitoring duration must be sufficient to characterize background surface-water concentrations for each parameter and detect statistically significant change over normal background fluctuations. The minimum monitoring duration to detect change varies by ecoregion and was discussed in detail in section 6.6. Because the purpose of the monitoring described in Table 11 is to detect whether the cumulative effects of shale gas development are resulting in water-quality change, monitoring should continue at selected monitoring sites for the long-term, as long as shale gas development activities continue in the Susquehanna River Basin. The magnitude of water-quality change that could occur from contamination related to shale gas development is unknown, but it would take 3-6 years of monthly monitoring to detect a 20-percent change in median specific conductance or total barium in the Susquehanna River Basin.

Groundwater data requirements can be characterized as shown in Table 26. Again, the entire suite of groundwater parameters is needed to identify whether shale gas development is the source of groundwater quality change. The study design for groundwater that is most applicable for existing groundwater data requires networks of 25-30 groundwater sampling sites in each of the primary drinking water aquifers in the Susquehanna River Basin, each site located within 1 mile of an HVHF well. Five priority spatial networks were identified for groundwater monitoring. Water-quality data should be collected at those sampling sites before and after shale gas development, with sampling events separated by approximately 10 years, and a repeat sampling event an additional 10 years into the future. A subset of 5 sampling sites in each network should be sampled every 2 years to identify interim water-quality changes.

13.2 Availability and Usability of Existing Water Data to Answer the Policy Question

This investigation found more than 960,000 surface-water records collected at about 14,700 monitoring sites over the last 85 years in the Susquehanna River Basin (Figure 13). However, there are no surface-water monitoring sites that meet all the criteria in Table 11.

- **The surface-water data needed for answering the case-study policy question are not currently available in the Susquehanna River Basin.**

While there are some applicable surface-water data available, each of the existing monitoring sites does not meet at least one of the criteria for location, parameters analyzed, frequency of monitoring, or duration of monitoring to detect statistically significant change associated with cumulative effects of shale gas development. Monitoring sites that have used an appropriate sampling plan for answering the case-study policy question are not in the right locations for detecting water-quality change related to shale gas development. Only 4 of 22 surface-water monitoring sites in the Susquehanna River Basin with enough existing data for a water-quality trend analysis for barium or specific conductance are located in watersheds with active HVHF wells, and few of the 26 recommended surface-water monitoring parameters are available for those sites. Only one of those monitoring sites is in a watershed with an HVHF well density greater than 0.5 HVHF wells per square mile. The existing surface-water data in the Susquehanna data set are not sufficient to detect whether the cumulative effects of shale gas development are resulting in water-quality change.

- **The groundwater data needed for answering the case-study policy question are not being collected.**

There is no systematic, large-scale, long-term monitoring effort underway to assess the effects of shale gas development on groundwater quality in the Susquehanna River Basin, and from the data sources that do exist, Figure 33 shows that only limited groundwater data are publicly available to help answer the policy question. The groundwater sampling sites with existing data are rarely located within 1 mile of an HVHF well, but even when they are in the right locations they lack data for most of the priority groundwater parameters. The available groundwater data lack the sampling frequency needed for a water-quality trend analysis and lack the number and location of sampling sites needed for a spatial water-quality network analysis. Selecting appropriate groundwater sampling sites is a major challenge for monitoring agencies because they do not have access to information on the location of future HVHF wells. Sampling sites within 1 mile of an HVHF well are due more to coincidence than to planning.

- **Current water data usability for answering the policy question is limited by insufficient data documentation and availability.**

The surface-water data identified through this case study were generated by 35 organizations, and groundwater data were collected by 10 organizations that collect water-quality data for parameters related to shale gas development in the Susquehanna River Basin. Insufficient and inconsistent documentation of available data limited the utility of these existing data sets. Substantial project time and effort over the course of this multi-year project were required to locate, obtain, and consistently format data. Missing information that is particularly important for this case study includes specification of whether water-quality samples were filtered or unfiltered, and information on the aquifer from which groundwater samples were taken.

Data sharing and data accessibility were limiting factors in data availability in this case study. The Water Quality Portal (National Water Quality Monitoring Council, 2014a), a cooperative service that provides publicly available water-quality data from Federal databases, including data collected by more than 400 State, Federal, Tribal, and local organizations, was established to facilitate water data sharing. Data collected at only 19 percent of the surface-water monitoring sites identified through this case study in the Susquehanna River Basin are available through the Water Quality Portal, but 85 percent of the water-quality data records are available through the Portal. This finding indicates that the monitoring sites from the Susquehanna data set that are available through the Water Quality Portal are sites with long data records, monitoring sites that are more likely to have the data necessary to identify a water-quality trend. Still, important data sets are missing from this collection including data collected by volunteer organizations, local governments, and academia. The Water Quality Portal includes data collected at 45 percent of the groundwater monitoring sites identified through this case study in the Susquehanna River Basin, and 61 percent of the groundwater quality records are available through the Portal. The most important groundwater data sets, those collected by industry, are not available through data sharing systems and access to those data sets is very limited.

13.3 Approaches for Filling Data Gaps to Answer the Policy Question

This section presents the study approaches for filling the data gaps to address the case-study policy question in the Susquehanna River basin.

- **Increase monitoring at a minimum of 8 targeted surface-water monitoring sites; additional monitoring sites are highly recommended. The water data identified in Table 11 must be collected for each of these sites.**

A recently established surface-water monitoring program, the SRBC RWQMN, is designed to collect surface-water quality data related to shale gas development, and recent updates to a monitoring program at the PADEP collect data more closely associated with shale gas development than previous monitoring efforts in the Susquehanna data set. Many of the monitoring sites for these programs are in the right locations, but additional sampling frequency, parameters, and streamflow data are needed. For SRBC monitoring sites, a minimum of 4 additional years of monitoring are needed at current sampling frequencies before water-quality trends can begin to be detected. Increased monitoring at a subset of priority monitoring sites would start generating answers to the case-study policy question in a time-frame that is useful to decision makers and would provide the water data needed to identify whether shale gas development is contaminating surface water in the monitored watersheds. A minimum of 8 monitoring sites are needed, 2 in each of the ecoregions with active or planned HVHF development. Additional monitoring sites will provide critical information regarding the scope and magnitude of potential water-quality change associated with shale gas development, especially in watersheds with the highest density of HVHF wells and nested watershed monitoring sites. An example set of 17 priority surface-water monitoring sites was presented in Chapter 8. The increased monitoring must include analysis for the full suite of priority surface-water parameters, increased frequency to monthly sampling, and addition of streamgages where needed. An example of data needs was presented in Table 19.

- **Maintain data collection and analysis at enhanced surface-water monitoring sites for a minimum of 10 years and as long as shale gas development activities continue in the Susquehanna River Basin.**

Monitoring at enhanced priority sites should continue for more than 10 years to determine whether cumulative shale gas development activities are resulting in water-quality change over the long-term. Monitoring to fill the data gaps should be implemented using an adaptive-monitoring approach and coordinated among monitoring agencies. Care should be taken to coordinate closely among participating monitoring programs so that data collection can be planned for compatibility, sharing, and easy analysis. The data should be evaluated on a regular basis so the monitoring program can be adjusted as necessary to adapt to a changing understanding of shale gas development and water quality.

- **Design and implement a systematic, long-term groundwater monitoring program for detecting groundwater quality change related to shale gas development in the Susquehanna River Basin, building on data collected by the shale gas industry, if appropriate.**

There is no groundwater monitoring equivalent to the SRBC RWQMN that is investigating the potential for shale gas development to change groundwater quality across the Susquehanna River Basin, and the groundwater data compiled for this study cannot be used as the foundation of a new groundwater monitoring program as described above. The groundwater data summarized in Table 26 must be collected through a systematic monitoring program to be able to answer the case-study policy question.

The shale gas industry has collected the most comprehensive set of groundwater data that pre-dates development at HVHF wells in the Susquehanna River Basin. The use of these existing data sets as the foundation of a new long-term groundwater monitoring program would result in the most cost-effective and most timely approach for collecting groundwater-quality data to answer the case-study policy question if the data sets meet minimum requirements for statistical analysis. Most of the priority groundwater parameters would need to be available for each sampling site (Table 3), the source aquifer identified, and the filtered/unfiltered status specified for each parameter.

If access to the industry data cannot be obtained, a completely new groundwater monitoring program must be initiated that would also require industry participation to identify appropriate sampling sites based on plans for future shale gas development.

- **Establish a coordinating entity to develop and implement surface-water and groundwater monitoring plans in the Susquehanna River Basin, with representation from water monitoring organizations, the shale gas industry, domestic well owners, and public citizens.**

To answer the case-study policy question, the right water-quality data must be collected in the right locations with the right supporting information. Water monitoring organizations and academia are tasked with collecting the right water-quality data, but industry involvement is necessary to identify the right monitoring locations and provide the right supporting information. Industry cooperation is needed to provide access to existing data and to identify locations of new HVHF wells so appropriate sampling sites can be identified. Ongoing coordination between water monitoring agencies and the shale gas industry will provide the necessary updates on locations of active production and technological advances in shale gas extraction practices that may affect water monitoring strategies.

Because domestic wells are the most accessible locations for monitoring groundwater quality related to shale gas development, it is critical to include domestic well owners in the process to obtain access to domestic wells for sampling before and after new shale gas development. To gain public trust, water monitoring programs should engage the people living in affected areas in addition to independent experts. According to the Council of Canadian Academies (2014), citizens will have greater faith in water monitoring results if they can influence the design, access the results, and comment throughout the process.

Engagement of all the critical stakeholders will improve water data coordination, sharing, and analysis to better understand the water-quality impacts of shale gas development in the Susquehanna River Basin. Stakeholder participation will promote confidence in the analytical results and lead to a better understanding of the risks to water resources in the Susquehanna River Basin. With this information, real policies and regulations can be implemented to protect water resources in the Susquehanna River Basin.

A coordinating entity is needed to facilitate coordination of sampling plans among water monitoring organizations so data collection, analysis, and interpretation will be compatible and comparable across monitoring sites. Improved data documentation and data sharing will facilitate the use of water data for answering the case-study policy question. Tools such as the Consortium of Universities for the Advancement of Hydrologic Sciences, Inc. HydroDesktop (Consortium of Universities for the Advancement of Hydrologic Sciences, Inc., 2014), Water Quality Exchange (WQX) (U.S. Environmental Protection Agency, 2014) and the Water Quality Portal provide the infrastructure for organizations to format and share their data, but greater participation is needed. Consistent, thorough data documentation and wide availability of data sources through these services will increase the value of water-quality data from all monitoring agencies and reduce the amount of time needed to access and prepare data for new applications. A continued commitment to water-quality data-sharing systems is essential for maximizing use of existing water-quality data.

13.4 Conclusion

Targeted, robust monitoring networks for both surface water and groundwater are critical for identifying whether the influx of shale gas development activity in the Susquehanna River Basin is generating adverse changes in water quality. The existing water-quality data in the Susquehanna River Basin are inadequate to serve this purpose. This report presents strategies for collecting the water data needed to detect whether shale gas development activities are contaminating surface water or groundwater in the Susquehanna River Basin in a policy-relevant time frame. Key steps to generating the needed information include increased monitoring in strategic locations, design and implementation of a systematic groundwater monitoring program, and a long-term commitment to water-quality monitoring in the Susquehanna River

Basin supported by a coordinating entity. Water-quality data collection and analysis, with participation from the key stakeholders, can answer this urgent policy question of critical importance to the Northeast-Midwest region and prepare for the questions that will emerge with further growth of the shale gas industry. The sooner the region gets started, the better.

Summary of information needs to answer "Do Shale Gas Development Activities Contaminate Surface Water or Groundwater in the Susquehanna River Basin?"

- Increase monitoring at a subset of targeted surface water monitoring sites.
- Maintain data collection and analysis at enhanced surface water monitoring sites for a minimum of 10 years and as long as shale gas development activities continue in the Susquehanna River Basin.
- Design and implement a systematic, long-term groundwater monitoring program, building on data collected by the shale gas industry, if appropriate.
- Establish a coordinating entity to develop and implement surface water and groundwater monitoring plans in the Susquehanna River Basin, with representation from water monitoring organizations, the shale gas industry, domestic well owners, and public citizens.

14. References

Abdalla, Charles, and Drohan, Joy, 2009, Shaping proposed changes to Pennsylvania's total dissolved solids standard—A guide to the proposal and the commenting process: Pennsylvania State University, College of Agricultural Sciences, Agricultural Research and Cooperative Extension, 7 p., *http://extension.psu.edu/natural-resources/water/conservation/consumption-and-usage/TDS-highres-updateDec09.pdf*.

Advisory Committee on Water Information, 2015, National Ground-Water Monitoring Network, accessed November 23, 2015, *http://cida.usgs.gov/ngwmn/*.

Argue, Denise, Deacon, J., and Gilliom, R., 2014, A national compilation of water-quality monitoring data to support local, regional, and national scale water quality assessments: National Water Quality Monitoring Council, 9th National Monitoring Conference, April 28-May 2, 2014, Cincinnati, Ohio, Abstracts, p.2, *http://acwi.gov/monitoring/conference/2014/.*

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

Barton, Andrew, 2014, Application of an aqueous methane sensor to evaluate spatial and temporal changes in dissolved methane concentrations in a groundwater monitoring well: National Water Quality Monitoring Conference, 9th National Monitoring Conference, April 28 – May 2, 2014, Cincinnati, Ohio.

Battelle, 2013, Bradford and Susquehanna County, Pennsylvania retrospective case study characterization report: submitted by Battelle to the American Petroleum Institute and America's Natural Gas Alliance, February 2013, Technical Memo 4: Battelle Contract No. CON00011206, *http://www.api.org/~/media/Files/Policy/Hydraulic_Fracturing/Battelle-Studies/BradSusqReport-13-Feb-22.pdf*.

Bowen, Z.H., Oelsner, G.P., Cade, B.S., Gallegos, T.J., Farag, A.M., Mott, D.N., Potter, C.J., Cinotto, P.J., Clark, M.L., Kappel, W.M., Kresse, T.M., Melcher, C.P., Paschke, S.S., Susong, D.D., and Varela, B.A., 2015, Assessment of surface water chloride and conductivity trends in areas of unconventional oil and gas development—Why existing national data sets cannot tell us what we would like to know, Water Resour. Res., 51, 704–715, *http://dx.doi.org/10.1002/2014WR016382*.

Boyer, E.W., Swistock, B.R., Clark, James, Madden, Mark, and Rizzo, D.E., 2012, The impact of Marcellus gas drilling on rural drinking water supplies: The Center for Rural Pennsylvania, revised March 2012, originally published October 2011, 28 p., *http://www.rural.palegislature.us/documents/reports/Marcellus_and_drinking_water_2012.pdf*.

Brabec, Elizabeth, Schulte, Stacey, and Richards, P.L., 2002, Impervious surfaces and water quality: A review of current literature and its implications for watershed planning: Journal of Planning Literature, v. 16, no. 4, p. 499-514.

Brantley, S.L., 2013, Using CUAHSI Hydrologic Information System (HIS) in the Shale Network to assess water quality and quantity data in regions of hydraulic fracturing: Consortium of Universities for the Advancement of Hydrologic Science, Inc., 2013 Spring HIS Cyberseminars, March 27, 2013, *https://www.cuahsi.org/Posts/Entry/8915.*

Carter, K.M., Kresic, Nevin, Muller, Peter, and Vittorio, L.F., 2013, Technical rebuttal to article claiming a link between hydraulic fracturing and groundwater contamination: Pennsylvania Council of Professional Geologists, 13 p., *https://pcpg.wildapricot.org/Resources/Documents/Shale%20Gas/PAGS%20PCPG%20Rebuttal%20to%20Frac%20Induced%20GW%20Contamination%20Article%201.pdf*.

Centre County Pennsylvania Senior Environmental Corps, 2014, accessed March 3, 2014, at *http://ccpasec.org/now/index.php?option=com_content&view=frontpage&Itemid=1*.

Chaplin, J., 2012, Baseline monitoring in selected areas of the Marcellus play: Presented at Susquehanna River Basin Commission Water Quality Advisory Committee Meeting, May 22, 2012, *http://www.srbc.net/programs/docs/wqac052212baselinemonitoring.PDF*.

Coleman, N.P. and McElreath, Debby, 2012, Short-term intra-well variability in methane concentrations from domestic well waters in Northeastern Pennsylvania: Ground Water Protection Council, Stray Gas Incidence & Response Forum, July 24-26, 2012, Cleveland, Ohio, *http://www.gwpc.org/sites/default/files/eventsessions/Coleman_Nancy.pdf*.

Coles, J.F., McMahon, Gerard, Bell, A.H., Brown, L.R., Fitzpatrick, F.A., Scudder Eikenberry, B.C., Woodside, M.D., Cuffney, T.F., Bryant, W.L., Cappiella, Karen, Fraley-McNeal, Lisa, and Stack, W.P., 2012, Effects of urban development on stream ecosystems in nine metropolitan study areas across the United States: U.S. Geological Survey Circular 1373, 138 p., *http://pubs.usgs.gov/circ/1373/*.

Community Science Institute, 2015, Baseline testing; documenting pre-drilling water quality, accessed December 9, 2015 at http://communityscience.org/baseline-testing/.

Considine, Timothy, Watson, Robert, Considine, Nicholas, and Martin, John, 2012, Environmental impacts during Marcellus Shale gas drilling: Causes, impacts, and remedies: Shale Resources and Society Institute Report 2012-1, State University of New York, Buffalo, 44 p., *http://cce.cornell.edu/EnergyClimateChange/NaturalGasDev/Documents/UBSRSI-Environmental%20Impact%20Report%202012.pdf*.

Consortium of Universities for the Advancement of Hydrologic Sciences, Inc., 2014, CUASHI Water Data Center, accessed April 18, 2014 at *http://wdc.cuahsi.org/*.

Cooley, Heather, and Donnelly, Kristina, 2012, Hydraulic fracturing and water resources: Separating the frack from the fiction: Pacific Institute, Oakland, California, 34 p., ISBN: 1-893790-40-1, *http://pacinst.org/wp-content/uploads/2013/02/full_report5.pdf.*

Council of Canadian Academies, 2014, Environmental Impacts of Shale Gas Extraction in Canada, Ottawa (ON): The Expert Panel on Harnessing Science and Technology to Understand the Environmental Impacts of Shale Gas Extraction, Council of Canadian Academies.

Davis, S.N., Whittemore, D.O., and Fabryka-Martin, 1998, Uses of chloride/bromide ratios in studies of potable water: Groundwater, vol. 36, n. 2., pp. 338-350.

Drohan, P.J., and Brittingham, M., 2012, Topographic and soil constraints to shale-gas development in the Northcentral Appalachians: Soil Science Society of America Journal, v. 76, no. 5, p. 1696-1706, doi:10.2136/sssaj2012.0087, *http://www.uvm.edu/~pss/documents/Drohan_publications.pdf*.

Drollette, B.D., Hoelzer, K., Warner, N.R., Darrah, T.H., Karatum, O., O'Connor, M.P., Nelson, R.K., Fernandez, L.A., Reddy, C.M., Vengosh, A., Jackson, R. B., Elsner, M., Plata, D.L., 2015, Elevated levels of diesel range organic compounds in groundwater near Marcellus gas operations are derived from surface activities: Proceedings of the National Academy of Sciences, v. 112, n. 43, p. 13184-13189, doi: 10.1073/pnas.1511474112.

Engelder, T., 2012, Capillary tension and imbibition sequester frack fluid in Marcellus gas shale: Proceedings of the National Academy of Sciences, USA, v. 109, no. 52, E3625 at *http://www.pnas.org/content/109/52/E3625.full?ijkey=54e9b424d60785de1631a6b1645d2129902db24d&keytype2=tf_ipsecsha*.

Engelder, T., Cathles, L.M., and Bryndzia, L.T., 2014, The fate of residual treatment water in gas shale: Journal of Unconventional Oil and Gas Resources, v. 7, p. 33-48, *http://dx.doi.org/10.1016/j.juogr.2014.03.002*.

Esri, Inc., 2012, ArcGIS for Desktop, V. 10: Redlands, California, Esri, Inc., *http://www.esri.com/*.

Falcone, J.A., Carlisle, D.M., Wolock, D.M, and Meador, M.R., 2010, GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States: Ecology, v. 91, no. 2, Ecological Archives E091-045, p. 621, *http://www.datadryad.org*.

Fisher, Madeline, 2012, Fracking's footprint: Scientists study impact of shale gas development on Pennsylvania's forests: Soil Horizons, v. 53, no. 4, p. 1-3, doi:10.2136/sh2012-53-4lf.

Flewelling, S.A., Tymchak, M.P., and Warpinski, Norm, 2013, Hydraulic fracture height limits and fault interactions in tight oil and gas formations: Geophysical Research Letters, v. 40, no. 14, p. 3602–3606, doi:*10.1002/grl.50707*.

Flewelling, S.A., and Sharma, Manu, 2014, Constraints on upward migration of hydraulic fracturing fluid and brine: Ground Water, v. 52, no. 1, p. 9-19, first published online July 29, 2013, doi: 10.1111/gwat.12095.

Focazio, M.J., Reilly, T.E., Rupert, M.G., and Helsel, D.R., 2002, Assessing ground-water vulnerability to contamination: Providing scientifically defensible information for decision makers: U.S. Geological Survey Circular 1224, 33 p., *http://pubs.usgs.gov/circ/2002/circ1224/*.

Fontenot, B.E., Hunt, L.R., Hildenbrand, Z.L., Carlton, D.D., Jr., Oka, Hyppolite, Walton, J.L., Hopkins, Dan, Orsio, Alexandra, Bjorndal, Bryan, Hu, Q.H., and Schug, K.A., 2013, An evaluation of water quality in private drinking water wells near natural gas extraction sites in the Barnett Shale Formation: Environmental Science and Technology, v. 47, no. 17, p. 10032-10040, doi: 10.1021/es4011724, publication date July 25, 2013.

FracFocus, 2013, FracFocus Chemical Disclosure Registry, data downloaded November 7, 2013 at *http://fracfocus.org/chemical-use/chemicals-public-disclosure*.

Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., and Wickham, J., 2011, Completion of the 2006 National Land Cover Database for the conterminous United States: PE&RS (Photogrammetric Engineering and Remote Sensing), v. 77, no. 9, p. 858-864, *http://digital.ipcprintservices.com/publication/?i=78634*.

Gallegos, T.J., Varela, B.A., Haines, S.S., Engle, M.A., 2015, Hydraulic fracturing water use variability in the United States and potential environmental implications, Water Resources Research, 51, 5839–5845, doi:10.1002/2015WR017278.

Gallegos, T.J., and Varela, B.A., 2015, Trends in hydraulic fracturing distributions and treatment fluids, additives, proppants, and water volumes applied to wells drilled in the United States from 1947 through 2010—Data analysis and comparison to the literature: U.S. Geological Survey Scientific Investigations Report 2014–5131, 15 p., *http://dx.doi.org/10.3133/sir20145131*.

Geng, X., Davatzes, N., Soeder, D., Torlapati, J., Rodriguez, R., and Boufadel, M., 2014, Migration of high-pressure air during gas well drilling in the Appalachian Basin: Journal of Environmental Engineering, v. 140, 46 p.

Gilbert, R.O., 1987, Statistical methods for environmental pollution monitoring: Van Nostrand Reinhold, New York, 320 p.

Gorody, A.W., 2012, Factors affecting the variability of stray gas concentration and composition in groundwater: Environmental Geosciences, v. 19, no. 1, p. 17-31, doi: 10.1306/eg.12081111013.

Groom, Nichola, 2013, Analysis: Fracking water's dirty little secret – recycling: Reuters, Los Angeles, CA, June 15, 2013, *http://www.reuters.com/article/2013/07/15/us-fracking-water-analysis-idUSBRE96E0ML20130715.*

Haluszczak, L.O., Rose, A.W., and Kump, L.R., 2013, Geochemical evaluation of flowback brine from Marcellus gas wells in Pennsylvania, USA: Applied Geochemistry, v. 28, p. 55-61, *http://dx.doi.org/10.1016/j.apgeochem.2012.10.002*.

Harper, J.A. 1999, *The Geology of Pennsylvania*, Chap. 7, "Devonian," ed. C.H. Shultz, 108–127. Harrisburg, Pennsylvania, and Pittsburgh, Pennsylvania: Pennsylvania Geological Survey and Pittsburgh Geological Society.

Harrison, S.S., 1985, Contamination of aquifers by overpressuring the annulus of oil and gas wells: Groundwater, v. 23, no. 3, p. 317-324, doi: 10.1111/j.1745-6584.tb00775.x.

Hayes, T. 2009. Sampling and analysis of water streams associated with the development of Marcellus shale gas: Gas Technology Institute Des Plaines, IL, for the Marcellus Shale Coalition, 44 p.

Health Effects Institute Special Scientific Committee on Unconventional Oil and Gas Development in the Appalachian Basin, 2015, Strategic research agenda on the potential impacts of 21st century oil and gas development in the Appalachian region and beyond, 240p., accessed 11/9/2015, *http://www.healtheffects.org/UOGD/UODG-Research-Agenda-Nov-4-2015.pdf*.

Heisig, P.M., and 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: U.S. Geological Survey Scientific Investigations Report 2013-5190, 44 p., *http://pubs.er.usgs.gov/publication/sir20135190*.

Heilweil, V. M., Stolp, B. J., Kimball, B. A., Susong, D. D., Marston, T. M., and Gardner, P. M., 2013, A stream-based methane monitoring approach for evaluating groundwater impacts associated with unconventional gas development: Ground Water, v. 51, issue 4, p. 511–524.

Hetcher-Aguila, K.K. and Eckhardt, D.A.V., 2006, Ground-water quality in the upper Susquehanna River Basin, New York, 2004-05: U.S. Geological Survey Open-File Report 2006-1161, 20 p., *http://ny.water.usgs.gov/pubs/of/of061161/*.

Jackson, R.B., Vengosh, A., Darrah, T.H., Warner, N.R., Down, A., Poreda, R.J., Osborn, S.G., Zhao, K., and Karr, J.D., 2013, Increased stray gas abundance in a subset of drinking water wells near Marcellus Shale gas extractions: Proceedings of the National Academy of Sciences, USA, v. 110, no. 28, p. 11250-11255, published ahead of print June 24, 2013, doi:10.1073/pnas.1221635110.

Johnson, J.D., and Graney, J.R., 2015, Fingerprinting Marcellus Shale waste products from Pb isotope and trace metal perspectives: Applied Geochemistry, v. 60, p. 104-115, http://dx.doi.org/10.1016/j.apgeochem.2015.04.021.

Johnson, J.D., Graney, J.R., Capo, R.C., and Stewart, B.W., 2015, Identification and quantification of regional and road salt sources in watersheds along the New York/Pennsylvania border, USA: Applied Geochemistry, v. 60, p. 37-50, *http://dx.doi.org/10.1016/j.apgeochem.2014.08.002*.

Johnson, Nels, 2010, Pennsylvania energy impacts assessment, Report 1: Marcellus Shale natural gas and wind, 46 p. The Nature Conservancy, Pennsylvania Chapter, *http://www.nature.org/media/pa/tnc_energy_analysis.pdf*.

Kassotis, C.D., Tillitt, D.E., Davis, J.W., Hormann, A.M., and Nagel, S.C., 2013, Estrogen and androgen receptor activities of hydraulic fracturing chemicals and surface and ground water in a drilling-dense region: Endocrinology, v. 155, no. 3, p. 897-907, doi: *http://dx.doi.org/10.1210/en.2013-1697*.

Katz, B.G., S. M. Eberts, and L. J. Kauffman, 2011, Using Cl/Br ratios and other indicators to assess potential impacts on groundwater quality from septic systems: A review and examples from principal aquifers in the United States, Journal of Hydrology, 397 (pp. 151–166).

Lautz, L.K., Hoke, G.D., Lu, Z, Siegel, D. I., 2013, Characterizing water quality in southern New York state before hydraulic fracturing: project SWIFT (Shale-Water Interaction Forensic Tools): Geological Society of America Annual Meeting, October 27-30, 2013, Denver, Colorado.

Lindsey, B.D., and Rupert, M.G., 2012, Methods for evaluating temporal groundwater quality data and results of decadal-scale changes in chloride, dissolved solids, and nitrate concentrations in groundwater in the United States, 1988–2010: U.S. Geological Survey Scientific Investigations Report 2012–5049, 46 p.

Lu, Z., Hummel, S.T., Lautz, L.K., Hoke, G.D., Zhou, X., Leone, J. Siegel, D.I., 2015, Iodine as a sensitive tracer for detecting influence of organic-rich shale in shallow groundwater: Applied Geochemistry, v. 60, p. 29-36, http://dx.doi.org/10.1016/j.apgeochem.2014.10.019.

Lustgarten, A., 2009, Officials in 3 states pin water woes on oil fields: Propublica, April 26, 2009, *http://www.propublica.org/article/officials-in-three-states-pin-water-woes-on-gas-drilling-426.*

Madrigal, A.C., 2013, Shooting the well: The petroleum torpedoes of the early oil fields: The Atlantic, August 21, 2013, *http://www.theatlantic.com/technology/archive/2013/08/shooting-the-well-the-petroleum-torpedoes-of-the-early-oil-fields/278901/.*

Molofsky, L.J., Connor, J.A., Farhat, S.K., Wylie, A.S., Jr., and Wagner, T., 2011, Methane in Pennsylvania water wells unrelated to Marcellus Shale fracturing: Oil and Gas Journal, v. 109, no. 49, p. 54-67, December 5, 2011, *http://www.ogj.com/1/vol-109/issue-49/exploration-development/methane-in-pennsylvania-water-full.html*.

Molofsky, L.J., Connor, J.A., Wylie, A.S., Wagner, T., and Farhat, S.K., 2013, Evaluation of methane sources in groundwater in Northeastern Pennsylvania: Groundwater, v. 51, no. 3, p. 333–349, first published online April 5, 2013.

Mullaney, J.R., Lorenz, D.L., Arntson, A.D., 2009, Chloride in groundwater and surface water in areas underlain by the glacial aquifer system, northern United States: U.S. Geological Survey Scientific Investigations Report 2009–5086, 41 p.

National Geographic, 2014, Breaking fuel from the rock: National Geographic Daily News: accessed January 23, 2014, at *http://news.nationalgeographic.com/news/2010/10/101022-breaking-fuel-from-the-rock/*.

National Water Quality Monitoring Council, 2014a, Water Quality Portal: accessed April 18, 2014 at *http://www.waterqualitydata.us/index.jsp*.

National Water Quality Monitoring Council, 2014b, National Environmental Methods Index: accessed May 16, 2014 at *https://www.nemi.gov/home/.*

New York State Department of Environmental Conservation, 2011, Revised draft Supplemental Generic Environmental Impact Statement on the oil, gas and solution mining regulatory program (September 2011) - Well permit issuance for horizontal drilling and high-volume hydraulic fracturing to develop the Marcellus Shale and other low-permeability gas reservoirs, Chapter V and VI, potential environmental impacts: 288 p., *http://www.dec.ny.gov/docs/materials_minerals_pdf/rdsgeisch6a0911.pdf*.

New York State Department of Environmental Conservation, 2015, Data on oil, gas and other wells in New York State, well data files in csv format: Division of Mineral Resources, accessed August 27, 2015, at *http://www.dec.ny.gov/energy/1603.html/.*

New York State Department of Health, 2014, A public health review of high volume hydraulic fracturing for shale gas development, available at *https://www.scribd.com/doc/250405540/High-Volume-Hydraulic-Fracturing*.

Olmstead, S.M., Muehlenbachs, L.A., Shih, Jhih-Shyang, Chu, Ziyan, and Krupnick, A.J., 2013, Shale gas development impacts on surface water quality in Pennsylvania: Proceedings of the National Academy of Sciences, USA, v. 110, no. 13, p. 4962-4967, published ahead of print March 11, 2013, doi: 10.1073/pnas.1213871110.

Osborn, S.G., Vengosh, A., Warner, N.R., and Jackson, R.B., 2011, Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing: Proceedings of the National Academy of Sciences, USA, v. 108, no. 20, p. 8172–8176, published ahead of print May 9, 2011, doi: 10.1073/pnas.1100682108.

Pennsylvania State Public Broadcasting, 2011, Explore shale - An exploration of natural gas drilling and development in the Marcellus Shale: Pennsylvania State University, accessed February 24, 2014, at *http://exploreshale.org/*.

Pennsylvania Department of Environmental Protection, 2005, Elements of Pennsylvania's monitoring and assessment program: Division of Water Quality Standards, 43 p., *http://www.epa.gov/reg3wapd/monitor/pdf/pa_strategy.pdf*.

Pennsylvania Department of Environmental Protection, 2010, Pennsylvania hydraulic fracturing state review: 40 p., *http://www.strongerinc.org/sites/all/themes/stronger02/downloads/PA%20HF%20Review%20Print%20Version.pdf*.

Pennsylvania Department of Environmental Protection, 2011, Form 26R - Chemical analysis of residual waste annual report by the generator: Bureau of Waste Management 2540-PM-BWM0347, revised 1/2011, accessed February 20, 2014, at *http://www.elibrary.dep.state.pa.us/dsweb/Get/Document-80512/01%20Instructions%202540-PM-BWM0347.pdf.*

Pennsylvania Department of Environmental Protection, 2012a, Policy for erosion and sediment control and stormwater management for earth disturbance associated with oil and gas exploration, production, processing, or treatment operations or transmission facilities, document number 800-2100-008: December 29, 2012, at *http://www.elibrary.dep.state.pa.us/dsweb/Get/Document-92195/800-2100-008.pdf*

Pennsylvania Department of Environmental Protection, 2012b, Pennsylvania's Surface Water Quality Monitoring Network (WQN), last revised April 2012, at *http://www.elibrary.dep.state.pa.us/dsweb/Get/Document-91443/3800-BK-DEP0636.pdf*

Pennsylvania Department of Environmental Protection, 2013, Abandoned and orphaned well program, available at *http://files.dep.state.pa.us/OilGas/BOGM/BOGMPortalFiles/AbandonedOrphanWells/WellPluggingProgram.pdf*

Pennsylvania Department of Environmental Protection, 2014, Pennsylvania quality assurance work plan Chesapeake Bay non-tidal network water quality monitoring for water year 2014 at *http://www.chesapeakebay.net/documents/PADEP_Nontidal_QAPP_FINAL_May14.pdf*

Pennsylvania Department of Environmental Protection, 2015a, Oil and gas reports: Office of Oil and Gas Management, accessed August 25, 2015 at *http://www.portal.state.pa.us/portal/server.pt/community/oil_and_gas_reports/20297.*

Pennsylvania Department of Environmental Protection, 2015b, PADEP oil & gas reporting website-statewide data downloads by reporting period: Office of Oil and Gas Management, accessed September 23, 2015, at: *https://www.paoilandgasreporting.state.pa.us/publicreports/Modules/DataExports/DataExports.aspx.*

Pennsylvania Geological Survey, 2014, Pennsylvania groundwater information system (PaGWIS): Pennsylvania Geological Survey, 4th ser., SQL database, accessed June 10, 2014, 05:48 p.m. UTC, at *http://dcnr.state.pa.us/topogeo/groundwater/pagwis/records/index.htm.*

Pennsylvania State University, 2014, FAQs—Marcellus center for outreach and research, accessed January 30, 2014 and February 20, 2014, at *http://www.marcellus.psu.edu/resources/faq.php#*.

Perry, A.E., Smith, Bert, Bothun, Rikka, and Hollingsworth, Mark, 2012, The occurrence of methane in shallow groundwater from extensive pre-drill sampling: Ground Water Protection Council, Stray Gas Incidence & Response Forum, July 24-26, 2012, Cleveland, Ohio, *http://www.gwpc.org/sites/default/files/event-sessions/Perry_Elizabeth.pdf*.

Rantz, S.E., and others, 1982, Measurement and computation of streamflow: Volume 1. Measurement of stage and discharge, U.S. Geological Survey Water-Supply Paper 2175, *http://pubs.usgs.gov/wsp/wsp2175/pdf/WSP2175_vol1a.pdf.*

Reardon, K., 2014, Continuous monitoring of methanol concentration in water treatment processes associated with hydrofracturing: National Water Quality Monitoring Conference, 9th National Monitoring Conference, April 28 – May 2, 2014, Cincinnati, Ohio.

Reddy, J.E., and Risen, A.J., 2012, Groundwater quality in the Upper Susquehanna River Basin, New York, 2009: U.S. Geological Survey Open-File Report 2012–1045, 29 p., *http://pubs.usgs.gov/of/2012/1045/.*

Reese, S. O., Neboga, V. V., Pelepko, Seth, and others, 2014, Groundwater and petroleum resources of Sullivan County, Pennsylvania: Pennsylvania Geological Survey, 4th ser., Water Resource Report 71, 99 p., 6 pls., 27 p. appendix.

Saiers, J.E., and Barth, E., 2012, Potential contaminant pathways from hydraulically fractured shale aquifers: Ground Water, v. 50, no. 6, p. 826-828, doi: 10.1111/j.1745-6584.2012.00990.x.

Sanders, T.G., Ward, R.C., Loftis, J.C., Steele, T.D., Adrian, D.D., and Yevjevich, V., 1983, Design of networks for monitoring water quality: Water Resources Publications, Littleton, Colorado, 328 p.

Schueler, T.R., 1994, The importance of imperviousness: Watershed Protection Techniques, *v.* 1, no. 3, pp. 100-111.

Shale Network, 2014, Shale network - A network of shale water researchers: accessed January 16, 2014, at *http://www.shalenetwork.org/*.

Siegel, D.I., Smith, B, Hollingsworth, M., Perry, E., Bothun, R., Whisman, C., Wardrop, R.T., Good, D., 2013, The prevalence of methane, salinity and trace metals in shallow groundwater consumed in northeastern Pennsylvania and southwestern Pennsylvania, Northern West Virginia, and Eastern Ohio: implications for regulatory assessment of background water quality: Geological Society of America Annual Meeting, October 27-30, 2013, Denver, Colorado.

Siegel, D.I., Smith, B, Perry, E., Bothun, R., Hollingsworth, M., 2015a, Pre-drilling water-quality data of groundwater prior to shale gas drilling in the Appalachian Basin: Analysis of the Chesapeake Energy Corporation dataset: Applied Geochemistry v. 63, pp. 37-57.

Siegel, D.I., Smith, B, Perry, E., Bothun, R., Hollingsworth, M., 2015b, Pre-drilling water-quality data of groundwater prior to shale gas drilling in the Appalachian Basin: Analysis of the Chesapeake Energy Corporation dataset, Supplemental information: Applied Geochemistry, *http://www.sciencedirect.com/science/article/pii/S0883292715300056#MMCvFirst*.

Slonecker, E.T., Milheim, L.E., Roig-Silva, C.M., and Winters, S.G., 2014, Landscape consequences of natural gas extraction in Bedford, Blair, Cambria, Centre, Clearfield, Clinton, Columbia, Huntingdon, and Luzerne counties, Pennsylvania, 2004–2010: U.S. Geological Survey Open-File Report 2014–1089, 49 p., *http://dx.doi.org/10.3133/ofr20141089.*

Sloto, R.A., 2013, Baseline groundwater quality from 20 domestic wells in Sullivan County, Pennsylvania, 2012: U.S. Geological Survey Scientific Investigations Report 2013–5085, 27 p., *http://pubs.usgs.gov/sir/2013/5085/.*

Soeder, D. J., 1988, Porosity and permeability of eastern Devonian gas shale: Society of Petroleum Engineers Formation Evaluation, Vol. 3, No. 2, pp. 116-124.

Soeder, D.J., and Kappel, W.M., 2009, Water resources and natural gas production from the Marcellus Shale: U.S. Geological Survey Fact Sheet 2009–3032, 6 p.

Soeder, D.J., 2011, Preliminary analysis of the Weathering potential of Marcellus Shale drill cuttings: Geological Society of America Section Meeting, March 20, 2011, Pittsburgh, Pennsylvania, *https://gsa.confex.com/gsa/2011NE/finalprogram/abstract_186221.htm.*

Soeder, D.J., Sharma, S., Pekney N., , Hopkinson, L, Dilmore, R., Kutchko, B., Stewart, B., Carter, K., Hakala, A., Capo, R., 2014, An approach for assessing engineering risk from shale gas wells in the United States: International Journal of Coal Geology, v. 126, pp. 4–19.

Soller, D.R., Packard, P.H., and Garrity, C.P., 2012, Database for United States Geological Survey Map I-1970 -- Map showing the thickness and character of quaternary sediments in the glaciated United States east of the Rocky Mountains: U.S. Geological Survey Data Series 656, available at *http://pubs.usgs.gov/ds/656/*.

Somers, Keith M., 1997, Power analysis: A statistical tool for assessing the utility of a study: Ontario Ministry of the Environment Technical Report, 44p.

Spooner, J., Dressing, S.A., and Meals, D.W., 2011, Minimum detectable change analysis. Tech Notes 7, December 2011: Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., *http://www.bae.ncsu.edu/programs/extension/wqg/319monitoring/tech_notes.htm.*

Spotts, Pete, 2013, Fracking waste water contaminated Pennsylvania streambeds, study finds: The Christian Science Monitor, October 2, 2013, *http://www.csmonitor.com/Environment/2013/1002/Fracking-waste-water-contaminated-Pennsylvania-streambeds-study-finds-video*.

STAC (Chesapeake Bay Program Scientific and Technical Advisory Committee), 2013, Exploring the environmental effects of shale gas development in the Chesapeake Bay watershed: STAC Publication 13-01, Edgewater, MD, 30 p., *http://www.chesapeake.org/pubs/297_Gottschalk2013.pdf*.

Streets, D.A., 2012, Assessing the surface water quality impacts of Marcellus shale development in Whiteday Creek watershed, West Virginia: Morgantown, West Virginia University, M.S. thesis, no. 1522485, 109 p.

Susquehanna River Basin Commission, 2012a, Remote water quality monitoring network - Data report of baseline conditions for 2010/2011; a Summary: Susquehanna River Basin Commission Publication no. 280, April 5, 2012, *http://mdw.srbc.net/remotewaterquality/assets/downloads/pdf/RWQMN_datareport_2010-2011.pdf*.

Susquehanna River Basin Commission, 2012b, Quality assurance/work plan - Susquehanna River Basin Remote Water Quality Monitoring Network: Susquehanna River Basin Commission Document Number QA044, available at: *http://mdw.srbc.net/remotewaterquality/assets/downloads/pdf/QualityAssuranceWorkPlan.pdf*.

Susquehanna River Basin Commission, 2013, Natural gas well development in the Susquehanna River Basin: Susquehanna River Basin Commission, Information Sheet, Revised January, 2013, 4 p., *http://www.srbc.net/programs/docs/naturalgasinfosheetjan2013.pdf*.

Susquehanna River Basin Commission, 2014a, General SRBC and watershed facts: Susquehanna River Basin Commission, accessed January 30, 2014, *http://www.srbc.net/pubinfo/factsheets.htm*.

Susquehanna River Basin Commission, 2014b, Remote Water Quality Monitoring Network: Watershed Profiles, accessed November 25 2014, *http://mdw.srbc.net/remotewaterquality/*

Susquehanna River Basin Commission, 2015, Remote water quality monitoring network - Data report of baseline conditions for 2010-2013: Susquehanna River Basin Commission Publication no. 297, June 2015, *http://mdw.srbc.net/remotewaterquality/assets/downloads/pdf/RWQMN_datareport_2010-2013.PDF*.

Swistock, Bryan, and Rizzo, Dana, 2012, Water facts #29—Testing drinking water supplies near gas drilling activity: College of Agricultural Sciences, Pennsylvania State University, Penn State Extension, 6 p., accessed December 27, 2013, at *http://extension.psu.edu/natural-resources/water/marcellus-shale/drinking-water/testing-drinking-water-supplies-near-gas-drilling-activity/extension_publication_file*.

U.S. Department of Agriculture, 2012, Cropland data layer: National Agricultural Statistics Service, published crop-specific data layer, accessed December 12, 2013, at *http://nassgeodata.gmu.edu/CropScape/*.

U.S. Department of Energy, 2013, Modern shale gas development in the United States: An update: National Energy Technology Laboratory, Strategic Center for Natural Gas and Oil, September, 2013, 79 p., *http://www.netl.doe.gov/File%20Library/Research/Oil-Gas/shale-gas-primer-update-2013.pdf*.

U.S. Energy Information Administration, 2015, Natural Gas Growth Withdrawals and Production, accessed October 29, 2015 at *http://www.eia.gov/dnav/ng/ng_prod_sum_a_EPG0_FPD_mmcf_a.htm*.

U.S. Department of Labor, 2015, Bureau of Labor Statistics, Consumer Price Index Calculator, *http://data.bls.gov/cgi-bin/cpicalc.pl?cost1=4105&year1=2012&year2=2015*.

U.S. Energy Information Administration, 2011, US_ShalePlays_EIA_May2011, accessed on January 15, 2013 at *http://www.eia.gov/pub/oil_gas/natural_gas/analysis_publications/maps/maps.htm#geodata* and modified using details from Pennsylvania Geological Survey, 2011, Extent of Pennsylvania underlain by Marcellus Shale, accessed October 25, 2013 at *http://www.dcnr.state.pa.us/topogeo/econresource/oilandgas/marcellus/index.htm.*

U.S. Environmental Protection Agency, 2002, Guidance on choosing a sampling design for environmental data collection (EPA QA/G-5S). Environmental Protection Agency Office of Environmental Information, Report EPA/240/R-02/005, 166 p.

U.S. Environmental Protection Agency, 2006, Guidance on systematic planning using the data quality objectives process (EPA QA/G4): Environmental Protection Agency Office of Environmental Information, Report EPA/240/B-06/001, 111 p.

U.S. Environmental Protection Agency, 2010, Level III ecoregions of the continental United States (revision of Omernik, 1987): Corvallis, Oregon, U.S. Environmental Protection Agency National Health and Environmental Effects Research Laboratory, Map M-1, various scales.

U.S. Environmental Protection Agency, 2014, STOrage and RETrieval (STORET) data warehouse: STORET/WQX homepage, *http://www.epa.gov/storet/*, website last updated January 15, 2014.

U.S. Environmental Protection Agency, 2015a, Assessment of the Potential Impacts of Hydraulic Fracturing for Oil and Gas on Drinking Water Resources (External Review Draft), U.S. Environmental Protection Agency, Washington, DC EPA/600/R-15/047.

U.S. Environmental Protection Agency, 2015b. Review of State and Industry Spill Data: Characterization of Hydraulic Fracturing-Related Spills, Office of Research and Development, Washington, DC. EPA/601/R-14/001.

U.S. Environmental Protection Agency, 2015c, Retrospective Case Study in Northeastern Pennsylvania Study of the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources, Office of Research and Development, Washington, DC, May, 2015 EPA/600/R-14/088.

U.S. Fish & Wildlife Service, 2014, 2011 National survey of fishing, hunting, and wildlife- associated recreation: Pennsylvania, revised January 2014, 82 p., *www.census.gov/prod/2013pubs/fhw11-pa.pdf.*

U.S. Geological Survey, 1998a, National Water Information System (NWIS): U.S. Geological Survey Fact Sheet 027-98, 2 p., *http://pubs.usgs.gov/fs/FS-027-98/fs-027-98.pdf*.

U.S. Geological Survey, 1998b, Data elements groundwater sites: Office of Groundwater Technical Memorandum 98.02.

U.S. Geological Survey, 2002a, NWISWeb: New site for the nation's water data: U.S. Geological Survey Fact Sheet 128-02, 2 p., *http://pubs.usgs.gov/fs/fs-128-02/pdf/fs-128-02.pdf*.

U.S. Geological Survey, 2002b, Aquifers of alluvial and glacial origin, in the National Atlas of the United States of America, accessed November 8, 2013, at *http://nationalatlas.gov/atlasftp.html*.

U.S. Geological Survey, 2003, Principal aquifers of the 48 conterminous United States, Hawaii, Puerto Rico, and the U.S. Virgin Islands, *http://nationalatlas.gov/mld/aquifrp.html*.

U.S. Geological Survey, 2005, State and county boundaries of the United States: National Atlas of the United States: Reston, VA, Digital data from 1:2,000,000-scale State boundary Digital Line Graph (DLG) files.

U.S. Geological Survey, 2006, Assessment of Appalachian Basin oil and gas resources: Devonian Shale–Middle and Upper Paleozoic total petroleum system: U.S. Geological Survey Open-File Report 2006-1237, 70 p., *http://pubs.usgs.gov/of/2006/1237/*.

U.S. Geological Survey, 2012, 1:1,000,000-Scale Streams of the United States, in the *National Atlas of the United States of America*, accessed on October 17, 2012, at *http://nationalatlas.gov/atlasftp-1m.html*.

U.S. Geological Survey, 2014a, Water-quality assessments of principal aquifers, accessed May 23, 2014, *http://water.usgs.gov/nawqa/studies/praq/.*

U.S. Geological Survey, 2014b, U.S. National streamflow information program, accessed September 15, 2014, *http://water.usgs.gov/nsip/definition9.html.*

U.S. Geological Survey, 2014c, National Water Information System: Web Interface, accessed April 3, 2014, *http://waterdata.usgs.gov/nwis/nwisman/?site_no=01550000&agency_cd=USGS.*

U.S. Geological Survey, 2015, Lycoming County Groundwater Study, accessed November 3, 2015, *http://pa.water.usgs.gov/projects/groundwater/lycomingCty/*.

U.S. House of Representatives, 2011, Chemicals used in hydraulic fracturing: United States House of Representatives Committee on Energy and Commerce, Minority staff, April 2011, prepared by committee staff for Henry A. Waxman, Edward J. Markey, and Diana DeGette, 30 p., *http://democrats.energycommerce.house.gov/sites/default/files/documents/Hydraulic-Fracturing-Chemicals-2011-4-18.pdf*.

Vaidyanathan, Gayathri, 2013, Hydraulic fracturing: When 2 wells meet, spills can often follow: EnergyWire, Monday, August 5, 2013, Environment and Energy Publishing, LLC, *http://www.eenews.net/stories/1059985587*.

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, v. 340, no. 6134, May 17, 2013: v. 340 no. 6134 DOI: 10.1126/science.1235009 *http://www.marcellus.psu.edu/news/PDFs/scienceArticleMay2013.pdf*.

Warner, N.R., Jackson, R.B., Darrah, T.H., Osborn, S.G., Down, A., Zhao, K., White, A., and Vengosh, A., 2012, Geochemical evidence for possible natural migration of Marcellus Formation brine to shallow aquifers in Pennsylvania: Proceedings of the National Academy of Sciences, USA, v. 109, no. 30, p. 11961-11966, published ahead of print July 9, 2012, doi: 10.1073/pnas.1121181109.

Warner N.R., Kresse, T.M., Hays, P.D., Down, A., Karr, J.D., Jackson, R.B., Vengosh, A., 2013a, Geochemical and isotopic variations in shallow groundwater in areas of Fayetteville Shale development, north-central Arkansas: Applied Geochemistry, v. 35, p. 207-220, *http://dx.doi.org/10.1016/j.apgeochem.2013.04.013.*

Warner, N.R., Christie, C.A., Jackson, R.B., and Vengosh, A., 2013b, Impacts of shale gas wastewater disposal on water quality in western Pennsylvania: Environmental Science & Technology, v. 47, no. 20, p. 11849-11857, October 2, 2013, DOI: 10.1021/es402165b.

Weisberg, Deborah, 2011, Unassessed waters initiative, Pennsylvania Angler & Boater, January/February 2011, *http://fishandboat.com/anglerboater/2011ab/vol80num1_janfeb/03unassessed.pdf*

Whisman, C., Wardrop, R., Good, D., McElreath, D., Olmsted, C., and Smith, B., 2012, Real-time monitoring system for evaluating long-term variability of methane in domestic water wells in Pennsylvania: Ground Water Protection Council, Stray Gas Incidence & Response Forum, July 24-26, 2012, Cleveland, Ohio, *http://www.gwpc.org/sites/default/files/event-sessions/Whisman_Chuck.pdf*.

Whittemore, D.O., 1988, Bromide as a tracer in ground-water studies—Geochemistry and analytical determination, *in* Ground Water Geochemistry Conference, Denver, Colo., February 16–18, 1988, Proceedings: National Water Well Association, p. 339–359.

Williams, J.H., Taylor, L. E. and Low, D.J., 1998, Hydrogeology and groundwater quality of the glaciated valleys of Bradford, Tioga, and Potter Counties, Pennsylvania: Pennsylvania Geological Survey, 4th ser., Water Resource Report 68, 89 p.

Williamson, H. F., Andreano, R.L., and Menezes, C., 1966, The American petroleum industry, chap. 6 of Brady, D.S., ed, Output, employment, and productivity in the United States after 1800: National Bureau of Economic Research, ISBN: 0-870-14186-4, p. 349-404, *http://www.nber.org/books/brad66-1*.

Winter, T.C., Harvey, J.W., Franke, O.L., and Alley, W.M., 1998, Ground water and surface water - A single resource: U.S. Geological Survey Circular 1139, 79 p. *http://pubs.water.usgs.gov/circ1139*.

15. Appendix A: Surface and Groundwater Characterization

List of Figures

List of Tables

15.1 Surface-water data

Water chemistry in the Susquehanna River Basin can be classified by ecoregion (Figure A-1), with additional classifications for land use or geology within each ecoregion (Susquehanna River Basin Commission, 2012). Boxplots were prepared to illustrate the quantity of data in the Susquehanna data set and to compare the distribution of concentrations of a selected constituent when organized by grouping variables such as ecoregion. Figure A-2 is an explanation of the information shown in boxplots and serves as the explanation for all the boxplots used in this appendix. Statistical tests were used to determine whether data grouped by ecoregion are statistically different. Sample populations that share the same letter symbol are not significantly different at the 0.05 level of significance; for example, an ecoregion designated by the letter "A" is significantly different from an ecoregion designated with the letter "B" but is not significantly different from an ecoregion designated by the letters "AB."

Nonparametric tests were used because they do not require the assumption that the data come from normal distributions. A multiple comparison (pair-wise) test on ranks was used to determine significant differences for all combinations of groups with uncensored data. All statistical summaries and tests were calculated with the TIBCO Spotfire S+ software package (TIBCO, 2008).

Figure A-3 and Figure A-4 show how specific conductance, dissolved barium, and total barium vary by the main ecoregions in the shale gas-producing portion of the Susquehanna River Basin. Only pre-2007 data are included in the boxplots to characterize water quality prior to shale gas development in the region. Since surface water is influenced not only by land use, but by season and streamflow, these boxplots include only sites with at least 6 samples, to show data for different seasons and for a range of streamflow. The number of records is the number of samples that were used to make the box and whisker plot, and the number of sites is the number of monitoring sites that contributed records.

The maximum reporting level for barium was 10 µg/L. The maximum reporting level is the highest reporting level in the data set used to censor non-detections. Elevated non-detections (above 10 µg/L) were removed from the data set so they would not affect the summary statistics and boxplots. (Examples of elevated non-detections are < 20 µg/L, < 100 µg/L, and < 500 µg/L.) A few values reported as zero were removed, as the minimum reporting level could not be determined. Preparing the barium data resulted in so few (1 percent for dissolved and 0.2 percent for total) remaining censored data that non-detections were treated as detections to simplify the analysis. There were not enough barium data to conduct valid statistical tests to determine if barium concentrations differed significantly by ecoregion.

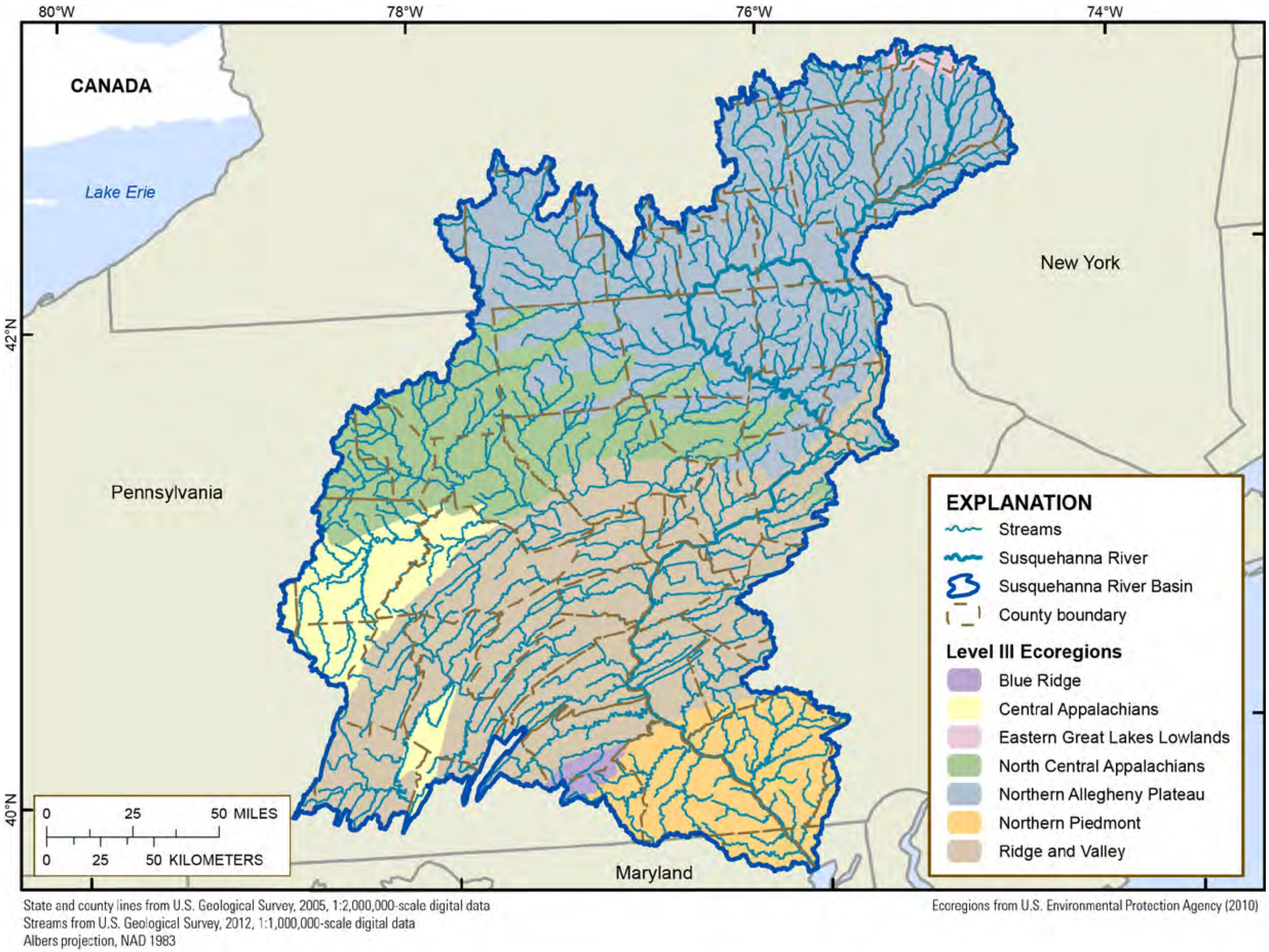


Figure A-1. Ecoregions in the Susquehanna River Basin.

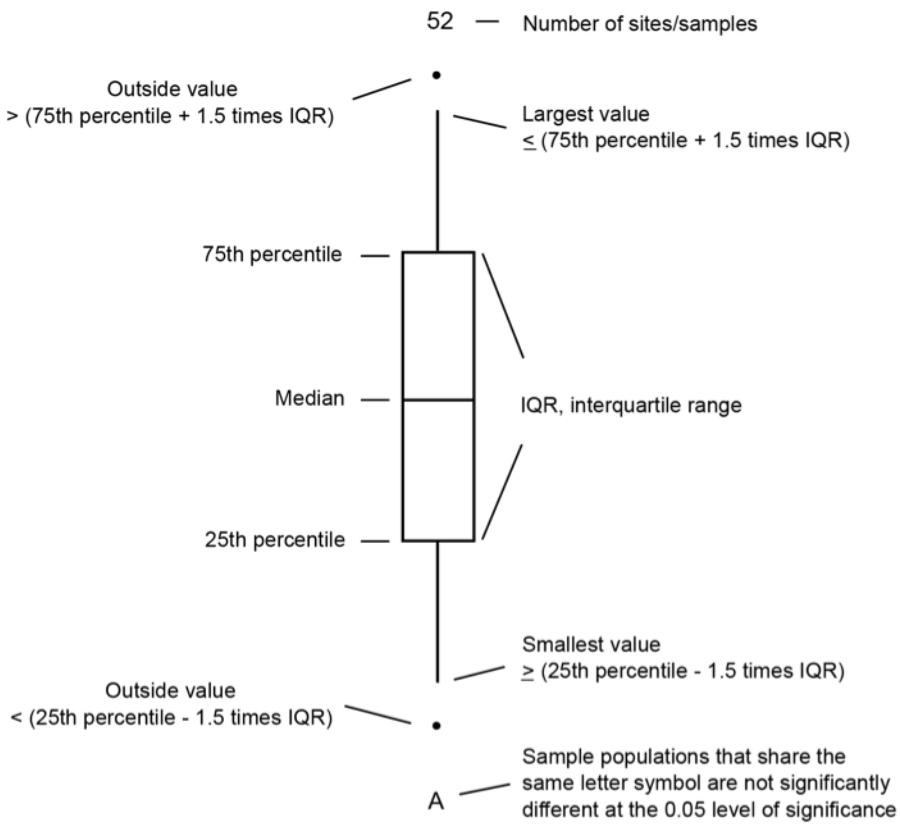


Figure A-2. Boxplot explanation.

15.1.1 Northern Allegheny Plateau Ecoregion

The Northern Allegheny Plateau ecoregion spans the northern portion of the Susquehanna River Basin (Figure A-1). Open valleys and low mountains support woodlands and agriculture in this region, which has experienced substantial shale gas development in northeastern Pennsylvania. This ecoregion has differences in glacial geology. Streams draining unglaciated portions have somewhat lower pH and conductance values when compared to the streams in glaciated areas (Susquehanna River Basin Commission, 2012). Overall, the SRBC found that streams in the ecoregion have higher and more variable specific conductance values than do streams in the North Central Appalachians ecoregion (Susquehanna River Basin Commission, 2012).

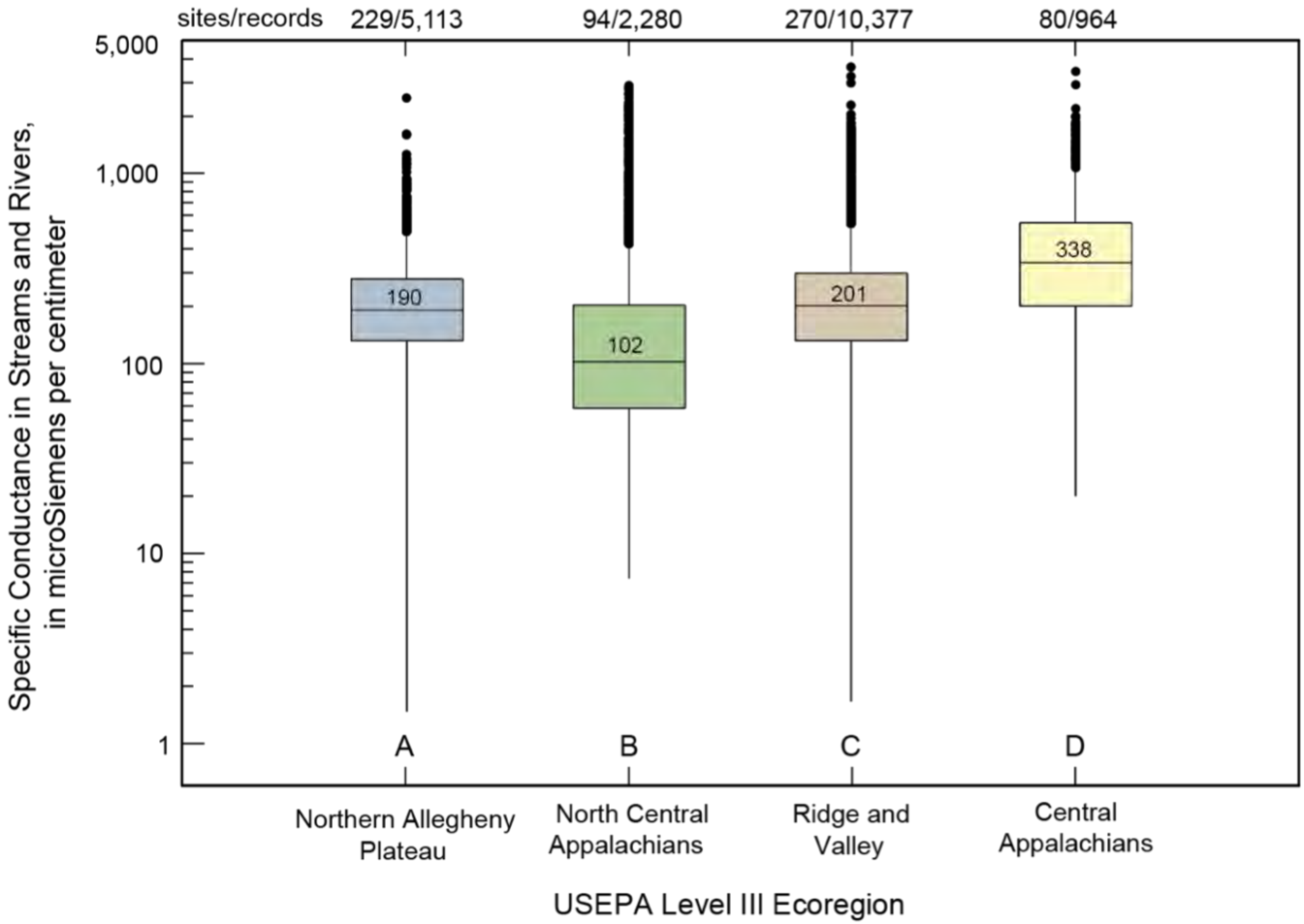


Figure A-3. Distribution of specific conductance records in surface water in the Susquehanna data set by USEPA Level III Ecoregion in the Marcellus and Utica Shale area of the Susquehanna River Basin.

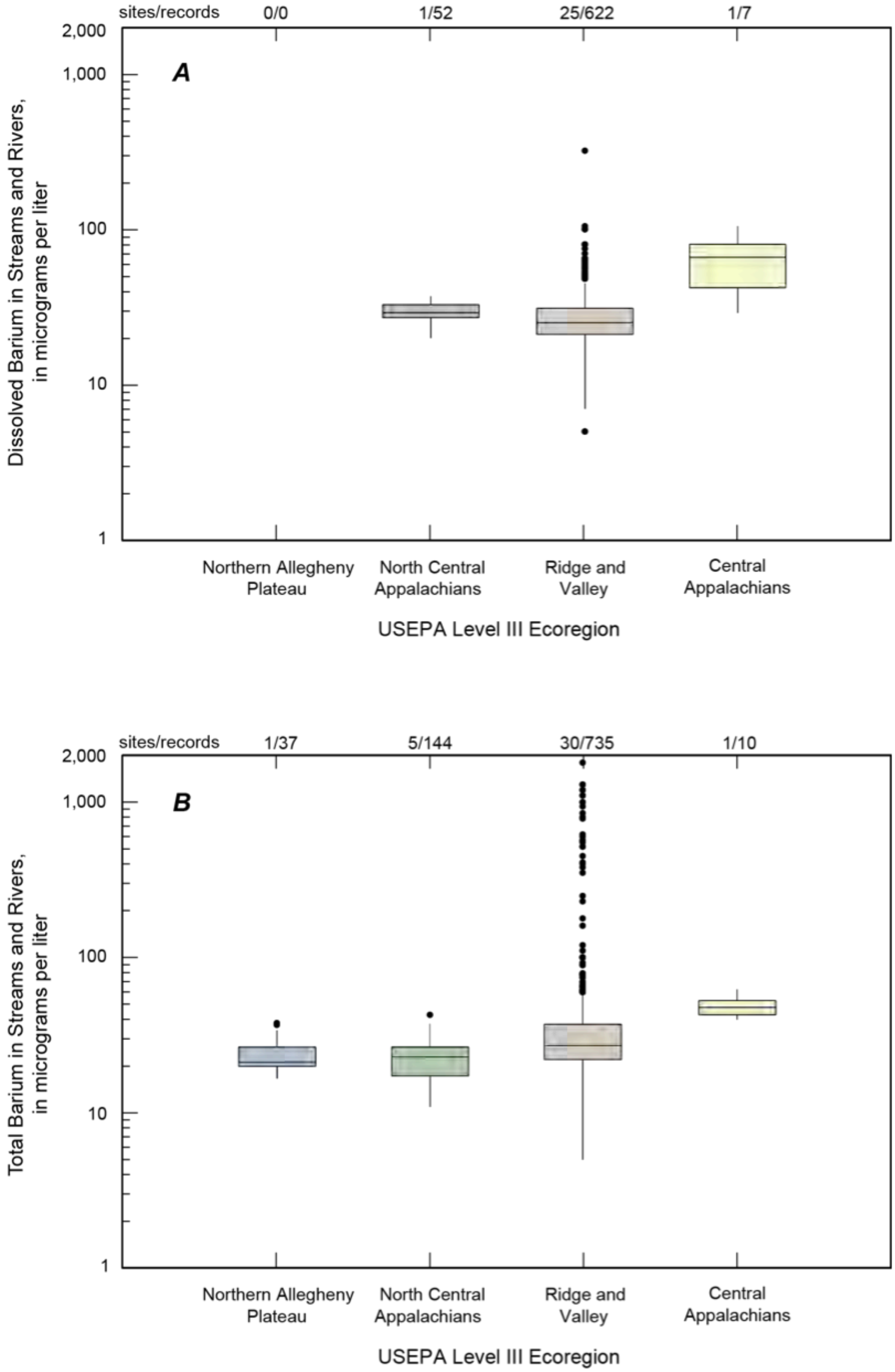


Figure A-4. Distribution of (A) dissolved barium, and (B) total barium records in surface water in the Susquehanna data set by USEPA Level III Ecoregion in the Marcellus and Utica Shale area of the Susquehanna River Basin.

15.1.2 North Central Appalachians Ecoregion

The North Central Appalachian ecoregion is a sedimentary upland that has high hills and low mountains. A substantial portion of the shale gas development in the Susquehanna River Basin is in this ecoregion. This ecoregion is located in the center of the Susquehanna River Basin and spans from the eastern to the western basin border. The ecoregion is primarily forested with a small fraction of agricultural land use. Unlike the Northern Allegheny Plateau ecoregion, differences in glacial geology do not seem to play a major role in differentiating water chemistry in streams in the North Central Appalachian ecoregion (Susquehanna River Basin Commission, 2012). The SRBC instead classifies the streams draining this ecoregion based on degree of agricultural land use (Susquehanna River Basin Commission, 2012). The North Central Appalachian ecoregion has the lowest mean specific conductance of the ecoregions shown in this analysis (Figure A-3). The SRBC found that the North Central Appalachian ecoregion has streams with the lowest specific conductance, pH, and turbidity values, and exhibit the lowest variability in specific conductance of the ecoregions they studied (Susquehanna River Basin Commission, 2012), despite the presence of a number of streams affected by acid mine drainage in portions of this ecoregion.

15.1.3 Central Appalachians Ecoregion

The Central Appalachian ecoregion covers a portion of the Marcellus and Utica Shale area of the Susquehanna River Basin, and is primarily forested with some agricultural land use. This area does not have substantial shale gas development, but may experience more in the future. The Central Appalachians region has the highest median specific conductance (Figure A-3), and is located in an area of prevalent acid mine drainage.

15.1.4 Ridge and Valley Ecoregion

The Ridge and Valley ecoregion contains an area of parallel ridges and valleys, with a mix of forest and agricultural land use. The Marcellus and Utica Shale areas underlie the ecoregion but this ecoregion may only be developed along the northern edge as most of the shale is thermally overmature. Streams impaired by acid mine drainage are prevalent in the eastern portion of this ecoregion. This ecoregion has the second highest median specific conductance in this analysis (Figure A-3); the SRBC also found high specific conductance levels in this ecoregion (Susquehanna River Basin Commission, 2012).

15.2 Groundwater

15.2.1 Groundwater data characterized by aquifer

Groundwater chemistry in the Susquehanna River Basin can be characterized by aquifer or principal aquifer. Boxplots are used here to illustrate the quantity and distribution of available data in the Susquehanna data set, and to compare the concentrations of selected constituents when grouped by aquifer or principal aquifer.

Data were processed for groundwater boxplots as follows. Sampling sites were included in the plots if they are more than 1 mile from the nearest oil or gas well (HVHF or conventional). A median concentration was used at sites that had more than one sample. A multiple comparison (pair-wise) test on ranks was used to determine significant differences for all combinations of aquifer groups with uncensored data. Sample populations that share the same letter symbol are not different at the 0.05 significance level; for example, an aquifer designated by the letters "AB" is not significantly different from an aquifer designated with either the letter "A" or "B", whereas an aquifer with the letter "A" designation is significantly different than an aquifer with the letter "B" designation.

When censored data were involved, the generalized Wilcoxon test was used to test for differences in left-censored, two-sample data. The generalized Wilcoxon test was applied to one pair of aquifer groups at a time. When multiple groups are being compared an overall error rate of 5 percent (alpha = 0.05) was used so that there was no more than a 5 percent chance of making one error; however, each group comparison must be made at an individual error rate smaller than 5 percent. The individual error rate is determined by Bonferroni's formula, where alpha is divided by the total number of comparisons (Helsel, 2005, p. 175). For example, if there are 4 aquifers, there will be 6 possible comparisons between groups, and the individual error rate will be 0.0083 (0.05 ÷ 6).

Figure A-5 through Figure A-12 illustrate boxplots for methane, specific conductance, bromide, and barium grouped by aquifer and principal aquifer. Principal aquifers were not included in a boxplot if they did not have the minimum required number of sites (25) for specific conductance or bromide. Most aquifers and principal aquifers did not have 25 monitoring sites for methane, but boxplots are shown in Figure A-5 and Figure A-6 to illustrate available data by appropriate grouping variables. The data presented here may not agree with the (larger) data set collected by industry, indicating the limitations of a limited (smaller) data set. Several aquifers and one principal aquifer had fewer than 25 monitoring sites for total barium, but boxplots are presented showing available data in Figure A-11B and Figure A-12B.

Dashed portions of a boxplot, or portions that fall below a shown reporting level, should be considered estimates due to censored data (non-detections). Percentiles and whisker plotting positions of boxplots were adjusted with results from ROS (regression on order statistics of log-transformed data) when enough censored data were present to affect summary statistics. The ROS method allows for multiple detection limits in the data set and yields good results even if the data are not log-normally distributed (Helsel and Cohn, 1988; Helsel, 2005).

For barium, elevated non-detections greater than 10 µg/L were removed from the data set so they would not affect the summary statistics and boxplots. For bromide, elevated non-detections greater than 0.05 mg/L were removed. There were no censored data for dissolved methane, and there were no elevated non-detections greater than 0.001 mg/L for total methane.

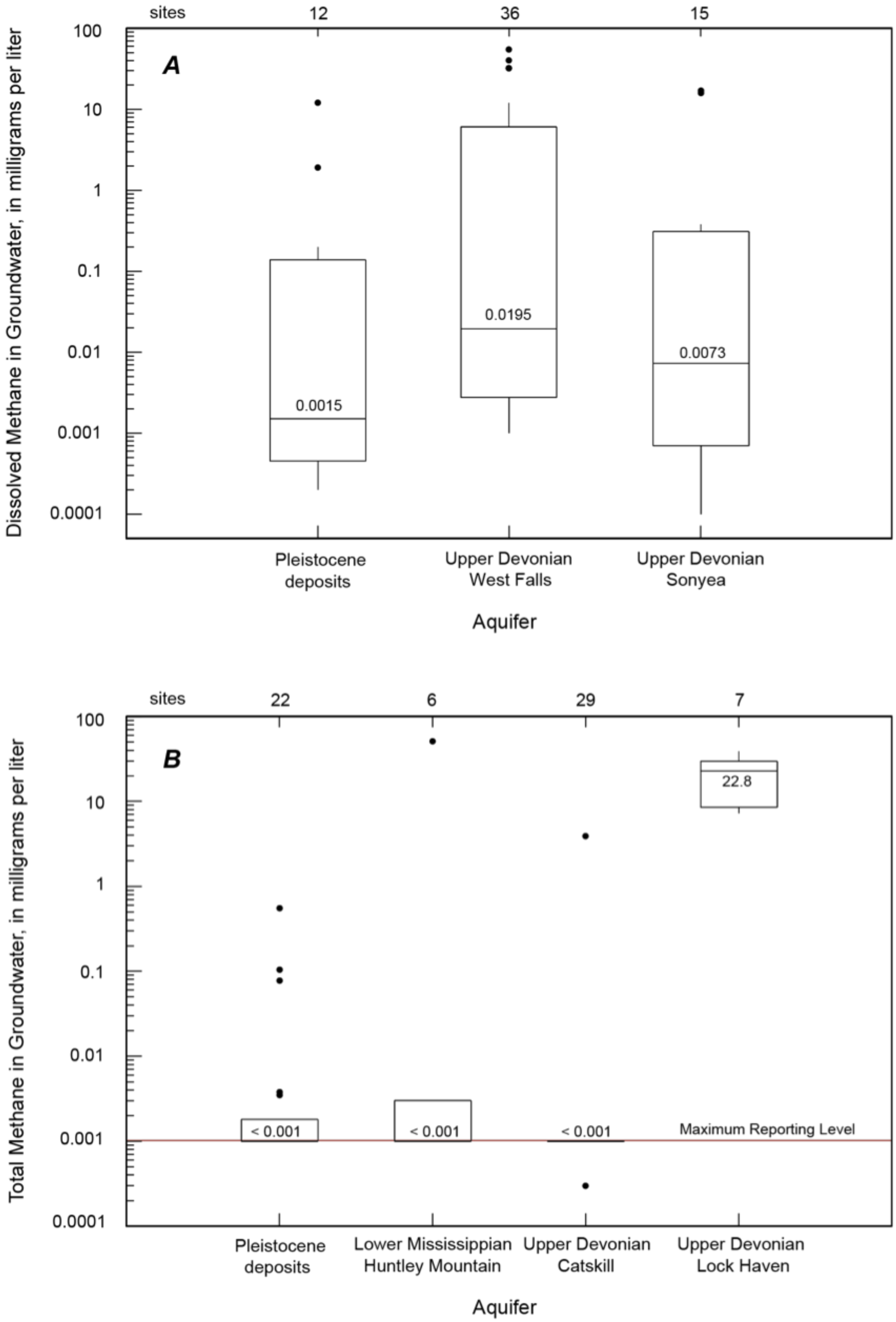


Figure A-5. Distribution of (A) dissolved methane, and (B) total methane concentrations by aquifer in the Marcellus and Utica Shale area of the Susquehanna River Basin for monitoring sites in the Susquehanna data set greater than 1 mile from any known oil or gas well. The maximum reporting level is the highest reporting level in the data set used to censor non-detections.

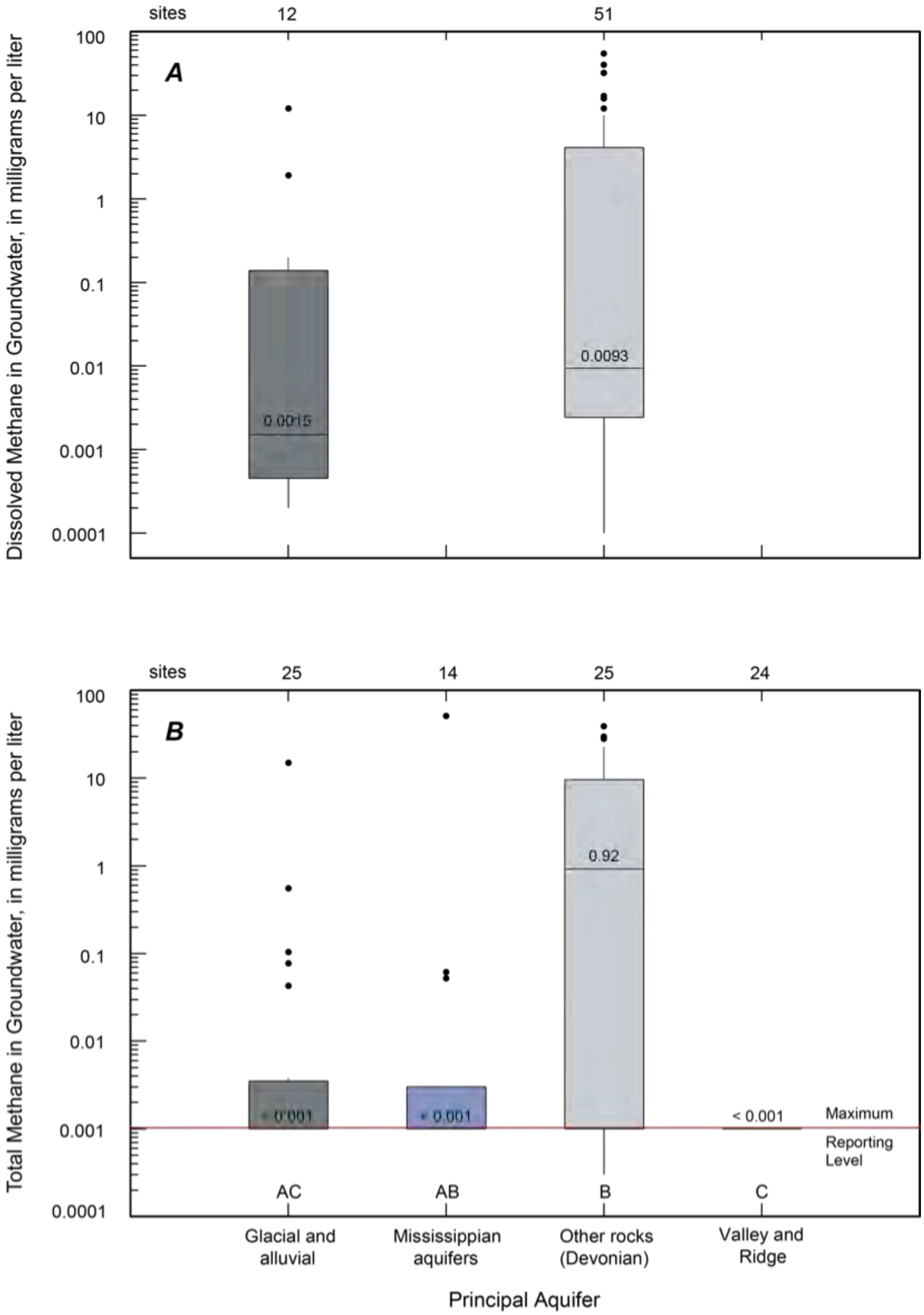
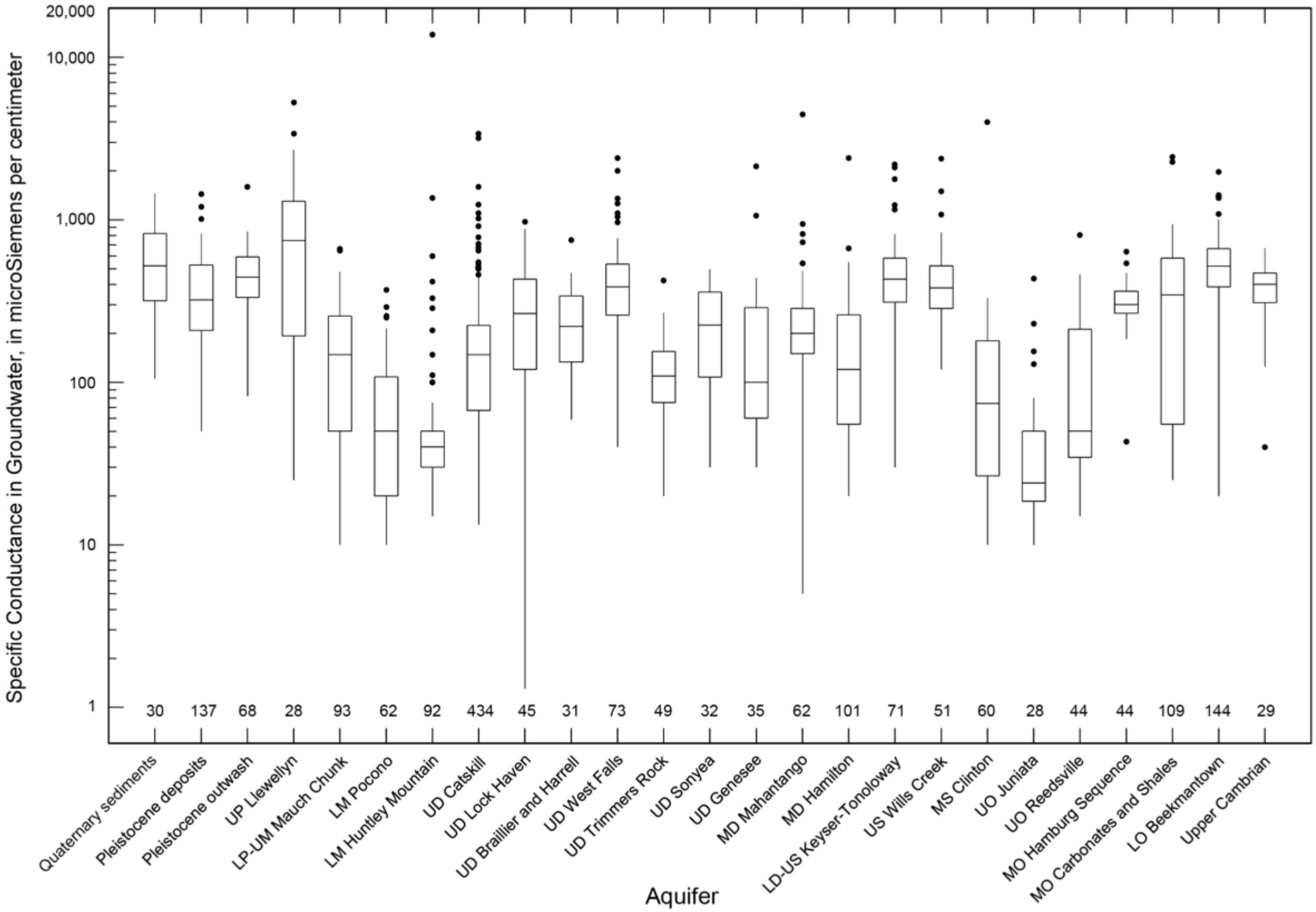


Figure A-6. Distribution of (A) dissolved methane, and (B) total methane concentrations by principal aquifer in the Marcellus and Utica Shale area of the Susquehanna River Basin for monitoring sites in the Susquehanna data set greater than 1 mile from any known oil or gas well. The maximum reporting level is the highest reporting level in the data set used to censor non-detections.



EXPLANATION

UP Upper Pennsylvanian	LM Lower Mississippian	LD Lower Devonian	UO Upper Ordovician	45 Number of sites
LP Lower Pennsylvanian	UD Upper Devonian	US Upper Silurian	MO Middle Ordovician	
UM Upper Mississippian	MD Middle Devonian	MS Middle Silurian	LO Lower Ordovician	

Figure A-7. Distribution of specific conductance by aquifer for groundwater samples in the Susquehanna data set collected from monitoring sites greater than 1 mile from any known oil or gas well in the Marcellus and Utica Shale area of the Susquehanna River Basin.

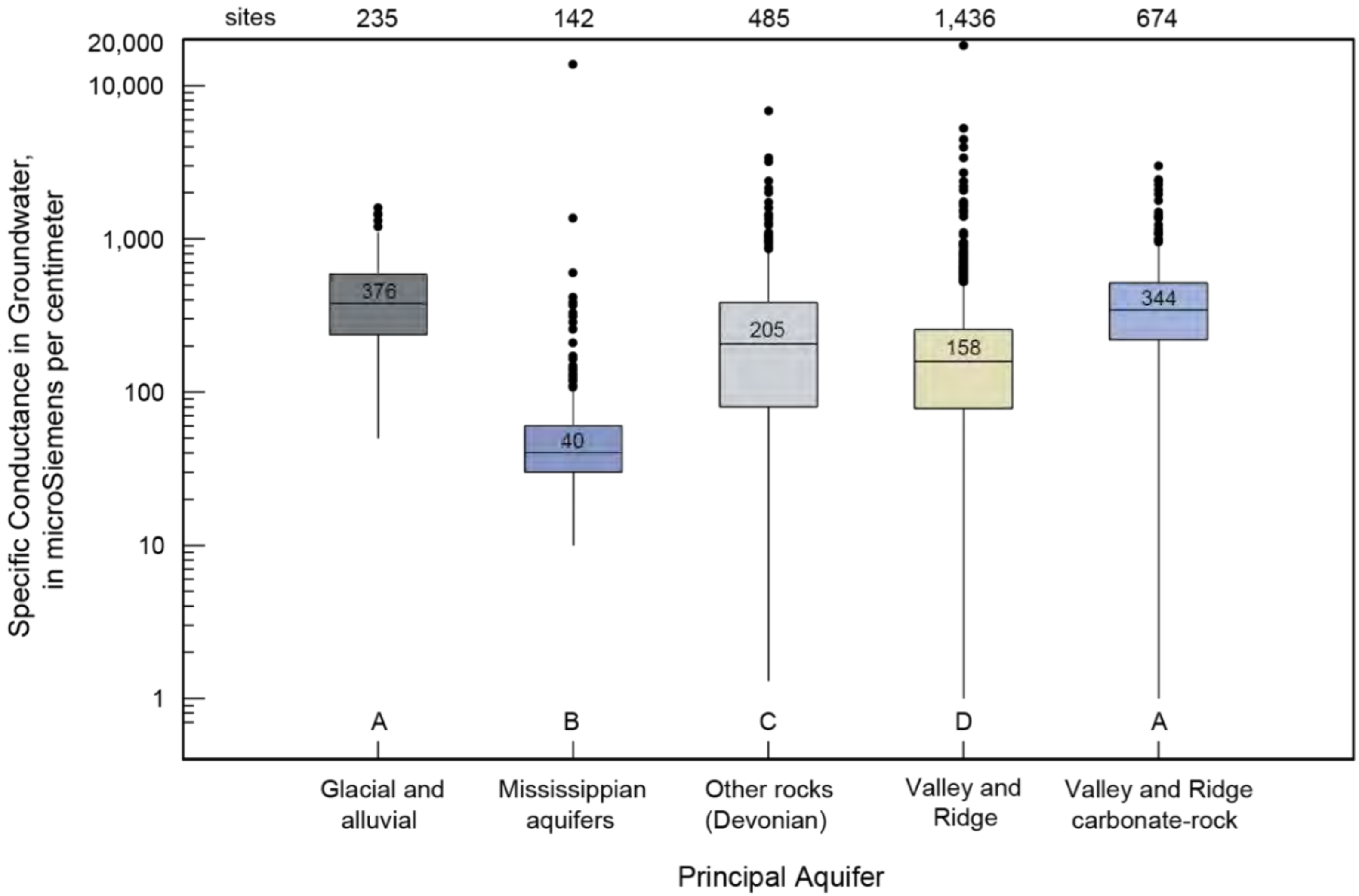


Figure A-8. Distribution of specific conductance by principal aquifer for groundwater samples in the Susquehanna data set collected from monitoring sites greater than 1 mile from any known oil or gas well in the Marcellus and Utica Shale area of the Susquehanna River Basin.

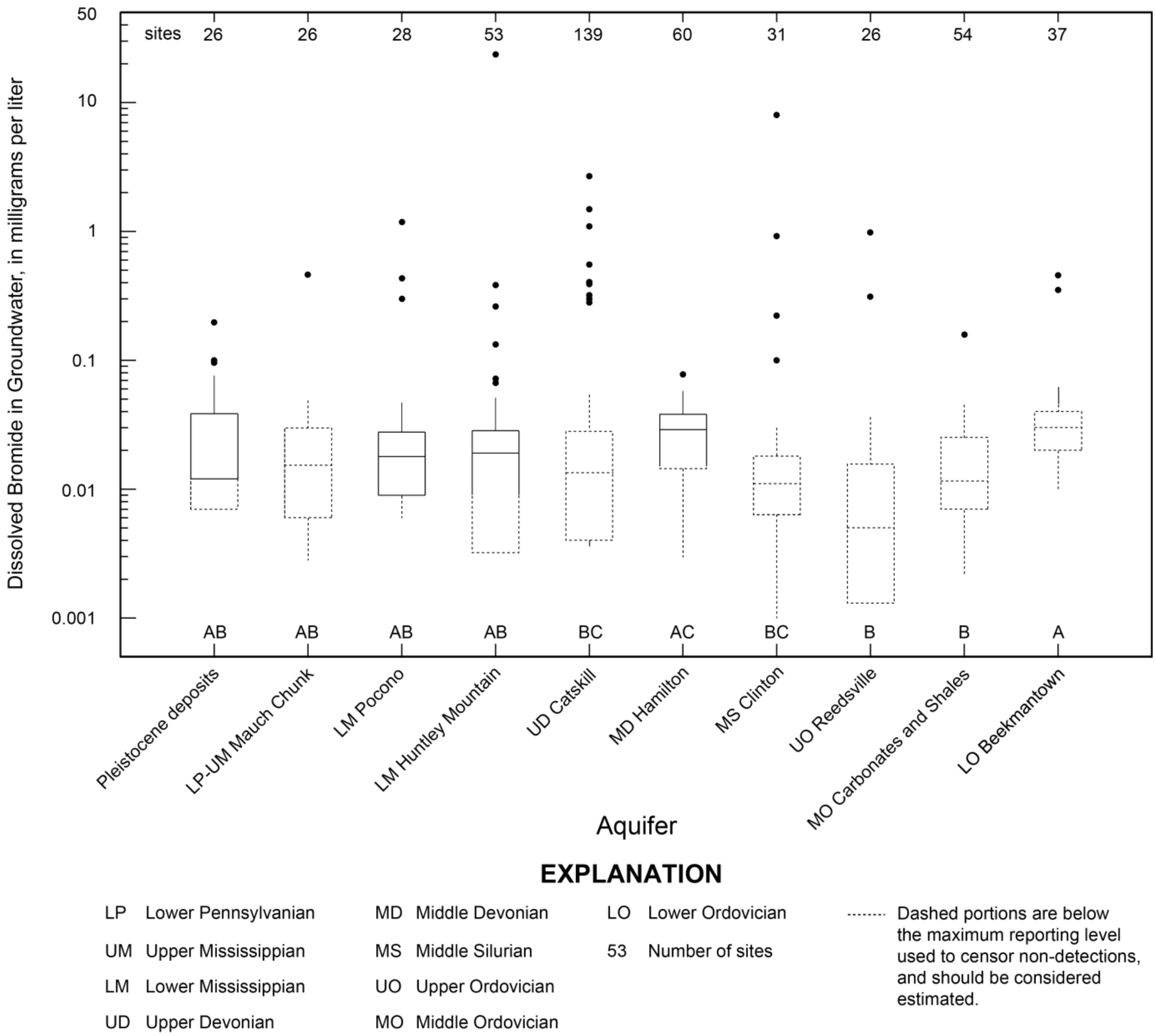


Figure A-9. Distribution of bromide by aquifer for groundwater samples in the Susquehanna data set collected from monitoring sites greater than 1 mile from any known oil or gas well in the Marcellus and Utica Shale area of the Susquehanna River Basin.

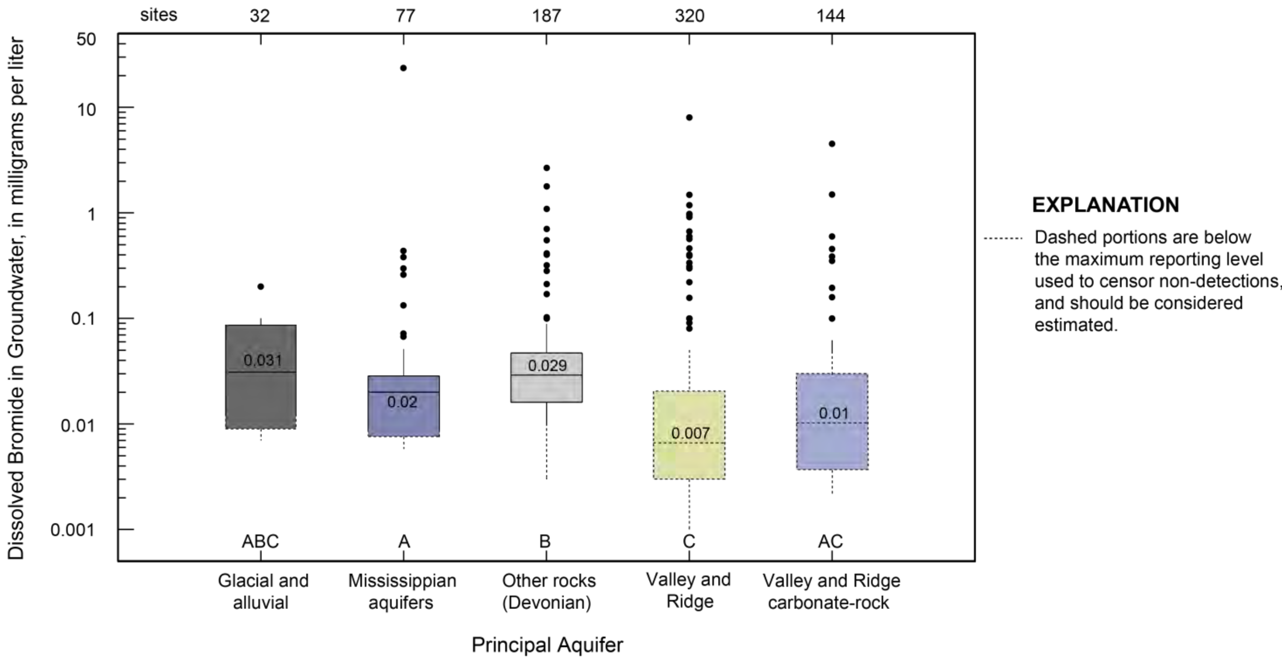


Figure A-10. Distribution of bromide by principal aquifer for groundwater samples in the Susquehanna data set collected from monitoring sites greater than 1 mile from any known oil or gas well in the Marcellus and Utica Shale area of the Susquehanna River Basin.

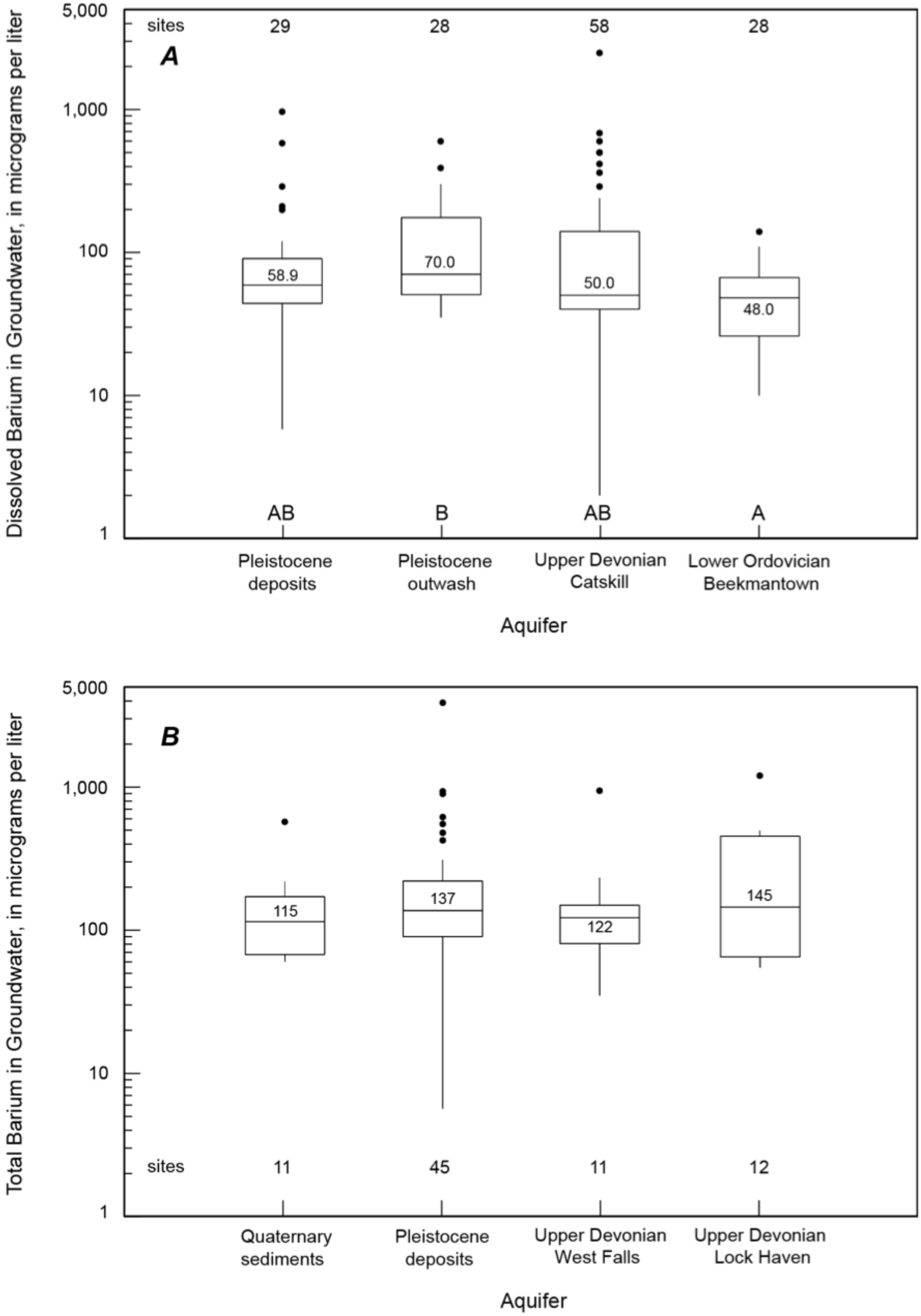


Figure A-11. Distribution of (A) dissolved barium, and (B) total barium by aquifer for groundwater samples in the Susquehanna data set collected from monitoring sites greater than 1 mile from any known oil or gas well in the Marcellus and Utica Shale area of the Susquehanna River Basin.

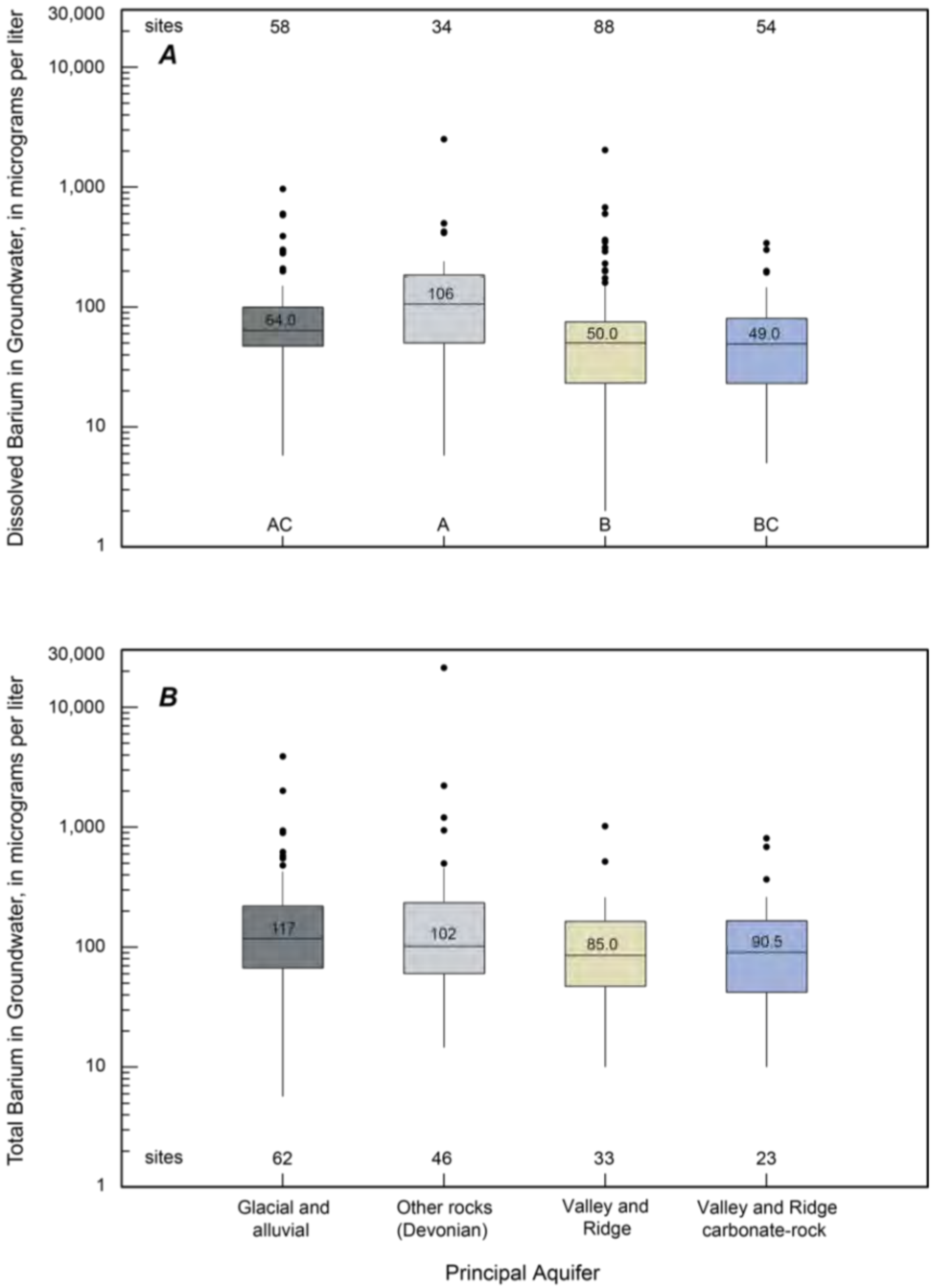


Figure A-12. Distribution of (A) dissolved barium, and (B) total barium by principal aquifer for groundwater samples in the Susquehanna data set collected from monitoring sites greater than 1 mile from any known oil or gas well in the Marcellus and Utica Shale area of the Susquehanna River Basin.

Some of the differences in aquifer water quality may be due to variability in physical settings that can have an important effect on residence times and groundwater chemistry. For example, the thick glacial sand and gravel “valley fill” aquifers of New York and northern Pennsylvania are typically at a lower altitude than thinner upland glacial aquifers in surrounding upland areas, and may be covered by depositional silt and clay that serve as a confining layer over the sand and gravel. These glacial sand and gravel aquifers in valleys have the potential of holding much older water (Heisig, 2012). Older water is often associated with reduced oxygen levels, or suboxic to anoxic “redox” conditions. According to McMahon et al. (2009):

> *“Reduction/oxidation (redox) processes affect the quality of groundwater in all aquifer systems. Redox processes can alternately mobilize or immobilize potentially toxic metals associated with naturally occurring aquifer materials, contribute to the degradation or preservation of anthropogenic contaminants, and generate undesirable byproducts, such as dissolved manganese (Mn^{2+}), ferrous iron (Fe^{2+}), hydrogen sulfide (H_2S), and methane (CH_4). Determining the kinds of redox processes that occur in an aquifer system, documenting their spatial distribution, and understanding how they affect concentrations of natural or anthropogenic contaminants are central to assessing and predicting the chemical quality of groundwater.”*

15.2.2 Groundwater data availability at different spatial scales

Criteria for groundwater data needed to answer the case-study policy question were presented in Chapter 9 of the case study report. Combinations of different aquifers and topographic settings (upland versus valley settings) were selected as the best network unit scale for detecting change in groundwater quality related to shale gas development. The availability of groundwater data at aquifer, principal aquifer, and ecoregion scales was examined for this case study and those results are presented here.

Table A-1 presents the number of existing focus parameter sampling sites by selected aquifers. There are potentially more than 25 “before” data records in the three primary drinking water aquifers that could be used as the first sampling event of a network analysis. The Upper Devonian Catskill aquifer has the most data, with 31 bromide and 156 specific conductance sites within 1 mile of a HVHF well. The Pleistocene deposits have 43 sites with barium measurements, and the Upper Devonian Lock Haven aquifer has 72 sites with specific conductance measurements.

Table A-2 shows counts of groundwater monitoring sites and records for focus parameters by principal aquifer. There are many records and sites for all parameters, but fewer sites fall within 1 mile of a HVHF well. There are potentially enough “before” data (greater than 25 data sites) in three principal aquifers that could be used as the first sampling event of a network analysis (see section 9.3). “Aquifers of alluvial and glacial origin” have 45 barium sites within 1 mile of a HVHF well, and “Mississipian aquifers” have 36 specific conductance sites and 25 barium sites. “Other rocks” is the name assigned to units that are minimally permeable, but may contain locally productive aquifers (U.S. Geological Survey, 2003). In the Susquehanna River Basin, “other rocks” is composed of Devonian-age bedrock aquifers and represent most of the area in the Susquehanna River Basin where intensive shale gas development is occurring. “Other rocks” have 47 bromide and 228 specific conductance sites within 1 mile of a HVHF well. Not all of these samples will necessarily have been collected in the right decade or have the appropriate minimum detection limits for network analysis.

Table A-3 shows counts of groundwater monitoring sites and records for focus parameters by ecoregion. There are many records and sites for all parameters, but again few of these fall within 1 mile of a HVHF well. There are potentially enough “before” data (greater than 25 monitoring sites) in two ecoregions for selected parameters that could be used as the first sampling event of a network analysis. Again, not all of these samples will necessarily have been collected in the right decade or have the appropriate minimum detection limits for network analysis. Ecoregions would be the least informative grouping variable for groundwater quality because they cover such large geographic areas. Ecoregions will include sites with different aquifers and even different principal aquifers.

15.2.3 Methane in groundwater in the Marcellus and Utica Shale area of the Susquehanna River Basin

In an investigation of background concentrations of methane in the glaciated 1,810-square-mile area of south-central New York along the Pennsylvania border, Heisig and Scott (2013) found that:

> *“Wells completed in bedrock within valleys and under confined groundwater conditions were most closely associated with the highest methane concentrations. Fifty-seven percent of valley wells had greater than or equal to 0.1 mg/L of methane, whereas only 10 percent of upland wells equaled or exceeded that concentration. Isotopic signatures differed between these groups as well. Methane in valley wells was predominantly thermogenic in origin, likely as a result of close vertical proximity to underlying methane-bearing saline groundwater and brine and possibly as a result of enhanced bedrock fracture permeability beneath valleys that provides an avenue for upward gas migration. Isotopic signatures of methane from four upland well samples indicated a microbial origin (carbon-dioxide reduction) with one sample possibly altered by microbial methane oxidation.”*

The work by Heisig and Scott (2013) demonstrates the importance of subsurface information (e.g., hydrogeology and well construction) in understanding methane occurrence and provides an initial conceptual framework that can be used in investigation of stray gas in south-central New York and possibly in the glaciated part of northern Pennsylvania.

A comprehensive study of over 2,300 baseline gas samples from Marcellus wells and groundwater in northeastern Pennsylvania shows that “microbial, mixed microbial and thermogenic, and thermogenic gases of different thermal maturities occur in some shallow aquifer systems and throughout the stratigraphy above the Marcellus Formation” (Baldassare et al., 2014). The origin of thermogenic gas may be associated with coal that is widely found in the Upper Devonian Catskill and Lock Haven aquifers in northeastern Pennsylvania (Wilson, 2014).

Table A-1. Groundwater sampling sites in the Susquehanna data set that are in or near the Marcellus and Utica Shale area with data for the focus parameters, sorted by selected aquifers.
[Abbreviations: SC, specific conductance; --, not applicable; HVHF, high-volume hydraulic fracturing]

Aquifer	Counts of groundwater sites and records for the focus parameters											
	Number of records				Number of sites				Number of sites within 1 mile of a HVHF well[1]			
	Barium	Bromide	Methane	SC	Barium	Bromide	Methane	SC	Barium	Bromide	Methane	SC
Lower Mississippian Huntley Mountain	14	83	11	138	10	80	8	133	1	13	1	20
Lower Mississippian Pocono	6	46	4	131	6	46	4	90	0	6	0	12
Lower Ordovician Beekmantown Group	28	39	--	344	28	37	--	146	0	0	--	0
Lower Pennsylvanian-Upper Mississippian Mauch Chunk	10	32	4	106	10	32	4	106	0	4	0	5
Lower Silurian Tuscarora Sandstone	--	14	--	18	--	14	--	18	--	0	--	0
Lower Devonian	3	--	--	5	3	--	--	5	0	--	--	0
Lower Mississippian	12	5	--	16	1	1	--	5	0	0	--	0
Middle Devonian Hamilton Group	19	70	--	118	7	65	--	106	0	0	--	0
Middle Pennsylvanian Allegheny	234	3	--	247	18	3	--	31	0	0	--	1
Middle Pennsylvanian Pottsville	32	19	1	57	10	19	1	35	0	2	0	4
Middle Devonian	1	1	1	5	1	1	1	5	0	0	0	0
Pleistocene deposits	146	26	35	206	137	26	35	175	**27**	0	0	11
Pleistocene outwash	129	38	19	248	69	10	15	100	16	0	0	4
Quaternary sediment	22	--	2	62	20	--	2	51	0	--	0	0

Aquifer	Counts of groundwater sites and records for the focus parameters											
	Number of records				Number of sites				Number of sites within 1 mile of a HVHF well[1]			
	Barium	Bromide	Methane	SC	Barium	Bromide	Methane	SC	Barium	Bromide	Methane	SC
Unknown (bedrock)	502	59	--	277	83	59	--	98	0	0	--	0
Unknown (no information)	443	747	124	1,100	261	724	106	907	1	1	0	1
Upper Cambrian	13	3	--	57	13	3	--	30	0	0	--	0
Upper Devonian	--	5	--	23	--	5	--	23	--	0	--	0
Upper Devonian Canadaway Group	1	11	--	21	1	11	--	21	0	0	--	0
Upper Devonian Catskill aquifer	121	221	38	727	90	207	30	657	2	**31**	0	**156**
Upper Devonian Foredknobs and Scherr	--	--	--	20	--	--	--	20	--	--	--	0
Upper Devonian Genesee Group	16	28	6	51	13	28	6	44	0	0	0	0
Upper Devonian Lock Haven aquifer	36	44	28	201	32	41	28	173	12	17	3	**72**
Upper Devonian Sonyea Group	--	--	--	33	--	--	--	33	--	--	--	0
Upper Devonian Trimmers Rock	1	17	1	51	1	17	1	51	0	0	0	1
Upper Devonian West Falls Group	23	31	41	114	21	31	41	111	0	0	0	0
Upper Pennsylvanian Conemaugh Group	17	--	--	47	17	--	--	47	0	--	--	3
Upper Silurian Wills Creek	1	8	--	53	1	8	--	52	0	0	--	0

[1]Aquifers with 25 or more groundwater sampling sites are shown in bold and may have enough data for a network analysis.

Table A-2. Groundwater monitoring sites in the Susquehanna data set that are in or near the Marcellus and Utica Shale area with data for the focus parameters, sorted by principal aquifer.

[Abbreviations: SC, specific conductance; --, not applicable; HVHF, high-volume hydraulic fracturing]

Principal aquifer	Counts of groundwater sites and records for the focus parameters											
	Number of records				Number of sites				Number of sites within 1 mile of a HVHF well[1]			
	Barium	Bromide	Methane	SC	Barium	Bromide	Methane	SC	Barium	Bromide	Methane	SC
Aquifers of Alluvial and Glacial Origin	297	64	56	516	226	36	52	326	**45**	0	0	16
Mississippian aquifers	54	139	19	277	25	132	16	222	1	**25**	1	**36**
New York and New England carbonate-rock aquifers	1	2	1	4	1	2	1	4	0	0	0	0
Other rocks	152	306	102	965	126	297	102	887	12	**47**	3	**228**
Pennsylvanian aquifers	468	17	1	491	48	17	1	100	1	2	0	8
Unknown (unconsolidated or bedrock)	234	60	95	214	234	60	94	204	1	1	0	1
Valley and Ridge aquifers	377	846	61	1,670	149	815	36	1,476	3	6	0	11
Valley and Ridge carbonate-rock aquifers	309	338	--	1,002	93	327	--	680	0	0	--	0

[1]Aquifers with 25 or more groundwater sampling sites are shown in bold and may have enough data for a network analysis.

Table A-3. Groundwater monitoring sites in the Susquehanna data set that are in or near the Marcellus and Utica Shale area with data for the focus parameters, sorted by U.S. Environmental Protection Agency Level III ecoregion.
[**Abbreviations**: SC, specific conductance; --, not applicable; HVHF, high-volume hydraulic fracturing]

Ecoregion	Counts of groundwater sites and records for the focus parameters											
	Number of records				Number of sites				Number of sites within 1 mile of a HVHF well[1]			
	Barium	Bromide	Methane	SC	Barium	Bromide	Methane	SC	Barium	Bromide	Methane	SC
Central Appalachians	503	28	--	628	52	24	--	140	1	2	--	8
Eastern Great Lakes Lowlands	5	3	2	11	5	3	2	11	0	0	0	0
North Central Appalachians	189	242	36	486	140	220	29	390	8	**32**	1	**58**
Northern Allegheny Plateau	514	283	236	1,394	463	265	235	1,178	**51**	**42**	3	**226**
Ridge and Valley	675	1,216	61	2,614	238	1,174	36	2,176	3	5	0	8
Western Allegheny Plateau	6	--	--	6	4	--	--	4	0	--	--	0

[1]Aquifers with 25 or more groundwater sampling sites are shown in bold and may have enough data for a network analysis.

In an analysis of water data prior to shale gas development, methane was found in 78 percent of pre-drilling water samples in Susquehanna County (Molofsky et al., 2011). Methane was found in 24.2 percent of the wells in pre-drilling samples in Pennsylvania by Chesapeake Energy (Perry et al., 2012; Whisman et al., 2012; Siegel et al., 2015). The USEPA reports that methane is naturally present in much of the groundwater of northeastern Pennsylvania (U.S. Environmental Protection Agency, 2012). Sloto (2013) reports water samples from two wells that contained methane concentrations greater than 1 milligram per liter (mg/L) out of 20 wells sampled prior to shale gas development in Sullivan County, Pennsylvania. A study by Heisig and Scott (2013), showed that methane can be found in the southern tier counties of New York; measurements of over 60 groundwater wells showed that while the majority of wells (64.6 percent) had methane concentrations of less than 0.1 mg/L, 15.6 percent of wells had methane concentrations equal to or greater than 10 mg/L. A review of groundwater studies across all of New York State—including the Heisig and Scott data (2013)—showed that for 437 wells sampled, 73.0 percent had methane concentrations less than 0.1 mg/L and 8.7 percent of wells had concentrations equal to or greater than 10 mg/L (Kappel et al., 2014).

A handful of researchers have published their analyses of industry methane data sets (Heisig and Scott, 2013; Siegel et al., 2015; Molofsky et al., 2013; Whisman et al., 2012). According to Molofsky et al. (2013), "methane concentrations are best correlated with topography and groundwater geochemistry rather than shale gas extraction activities" as shown in Figure A-13. Methane concentrations appear to occur more frequently in sodium-chloride (Na-Cl) and sodium-bicarbonate (Na-HCO_3) water types (Heisig and Scott, 2013).

Temporal variability of dissolved methane concentrations in domestic wells is not unusual and will require more frequent monitoring to establish baseline conditions. This finding has major implications for the minimum set of monitoring data needed to characterize pre-drilling water quality. A single pre-drill methane sample is often taken to represent the baseline condition at a domestic well, but this single sample may not represent the full range of natural variability (Whisman et al., 2012; Figure A-14). Short-term variations of dissolved methane in groundwater can occur on a day-to-day basis (Coleman and McElreath, 2012.)

Isotope ratios, along with isotopic and mole ratios of heavier hydrocarbon gases such as ethane and propane, can be used to help identify or "fingerprint" the source of thermogenic methane, but these analyses require complex and expensive laboratory work and do not indicate the mechanism of migration (Molofsky et al., 2013; Mulder, 2012; Osborn et al., 2011). It is therefore critical to monitor baseline methane conditions before drilling and to use multiple lines of evidence to better identify the origins of gas migration (Baldassare, 2014; Vidic et al., 2013).

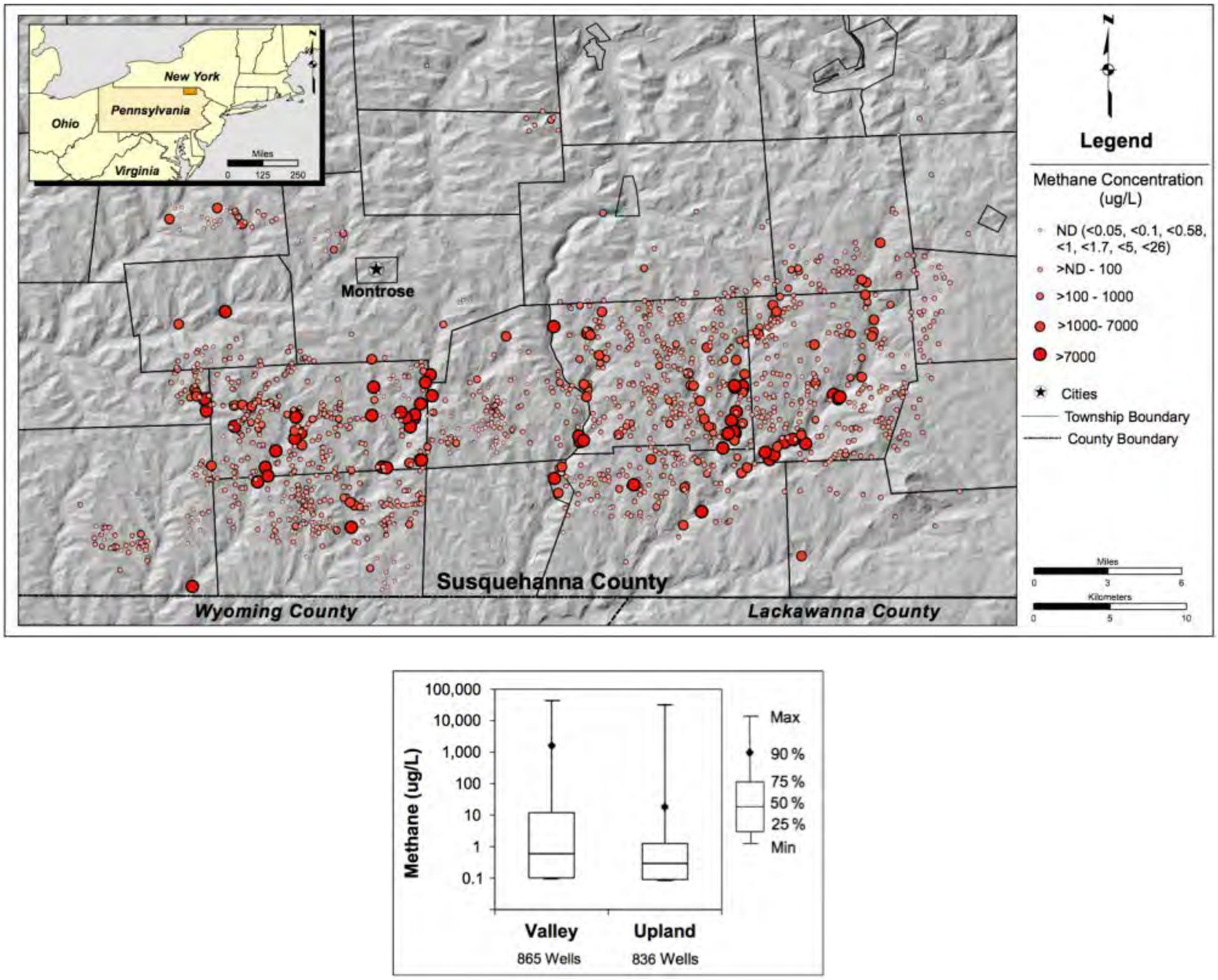


Figure A-13. Comparison between methane concentrations in groundwater samples from valley wells and upland wells in Susquehanna County, PA (from Molofsky et al., 2013; reprinted from Groundwater with permission of the National Groundwater Association, Copyright 2013).

Several studies have examined the incidence of methane near HVHF wells. One study of over 11,300 dissolved methane analyses collected by Chesapeake Energy found that methane concentrations in sampled wells were unrelated to their proximity to HVHF wells in northeastern Pennsylvania (Siegel et al., 2015). All of the non-industry publications examined for this study have small data sets. One study shows a high proportion of domestic wells with methane contamination near shale gas development in the Marcellus Shale area (Osborn et al., 2011), but this study does not demonstrate whether the methane existed in those locations prior to shale gas development activity in the area (Vidic et al., 2013). The averages reported in the Osborn study (2011) for sites both near and far from HVHF wells are not dissimilar from values for methane in groundwater from areas of Pennsylvania and West Virginia sampled by the U.S. Geological Survey before shale gas development, or samples in New York State where HVHF shale gas development is banned (Vidic et al., 2013). A similar study, conducted in Pennsylvania in areas of shale gas development looked at 48 wells both before and after shale gas development and found no significant relationship of methane concentrations to distance from a HVHF gas well (Boyer et al., 2012). Methane contamination incident data were not compiled or analyzed for this study.

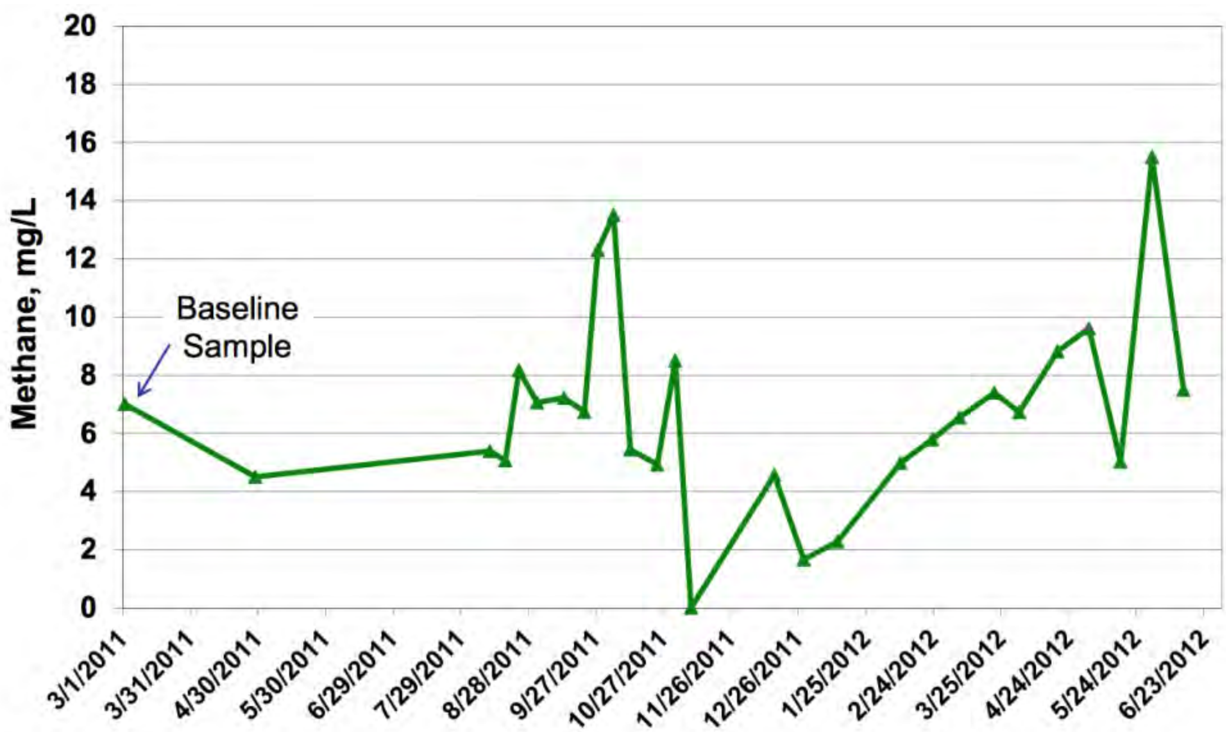


Figure A-14. Variability of methane at a well site in Sullivan County, Pennsylvania (from Whisman et al., 2012; reprinted with permission from Chesapeake Energy).

15.3 References

Baldassare, F.J., McCaffrey, M.A., and Harper, J.A., 2014, A geochemical context for stray gas investigations in the northern Appalachian Basin: Implications of analyses of natural gases from Neogene-through Devonian-age strata: AAPG Bulletin, v. 98, no. 2, pp. 341– 372.

Boyer, E.W.; Swistock, B.R.; Clark, James; Madden, Mark; and Rizzo, D.E., 2012, The impact of Marcellus gas drilling on rural drinking water supplies: The Center for Rural Pennsylvania, revised March 2012, originally published October 2011, 28 p., http://www.rural.palegislature.us/documents/reports/Marcellus_and_drinking_water_2012.pdf.

Coleman, N.P. and McElreath, Debby, 2012, Short-term intra-well variability in methane concentrations from domestic well waters in Northeastern Pennsylvania: Ground Water Protection Council, Stray Gas Incidence & Response Forum, July 24-26, 2012, Cleveland, Ohio, http://www.gwpc.org/sites/default/files/event-sessions/Coleman_Nancy.pdf.

Helsel, D.R., and Cohn, T.A., 1988, Estimation of descriptive statistics for multiply censored water quality data: Water Resources Research, v. 24, no. 12, p. 1997–2004.

Helsel, D.R., 2005, Nondetects and data analysis: Hoboken, N.J., John Wiley & Sons, 250 p.

Heisig, P.M., 2012, Hydrogeology of the Susquehanna River valley-fill aquifer system and adjacent areas in eastern Broome and southeastern Chenango Counties, New York: U.S. Geological Survey Scientific Investigations Report 2012–5282, 19 p., http://pubs.usgs.gov/sir/2012/5282.

Heisig, P.M., and 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: U.S. Geological Survey Scientific Investigations Report 2013-5190., 44 p., http://pubs.er.usgs.gov/publication/sir20135190.

Kappel, W.M., Nystrom, E.A., Reddy, J., Heisig, P., and Scott, T.M., 2014, Dissolved methane in groundwater and specific conductance of streams in New York – a retrospective and current evaluation of water-quality conditions: The Geological Society of America, Northeastern Section - 49th Annual Meeting, March 23-25th, 2014.

McMahon, P.B., Cowdery, T.K., Chapelle, F.H., and Jurgens, B.C., 2009, Redox conditions in selected principal aquifers of the United States: U.S. Geological Survey Fact Sheet 2009-3041, 6 p.

Molofsky, L.J., Connor, J.A., Farhat, S.K., Wylie, A.S., Jr., and Wagner, T., 2011, Methane in Pennsylvania water wells unrelated to Marcellus Shale fracturing: Oil and Gas Journal, v. 109, no. 49, p. 54-67, December 5, 2011, http://www.ogj.com/1/vol-109/issue-49/exploration-development/methane-in-pennsylvania-water-full.html.

Molofsky, L.J., Connor, J.A., Wylie, A.S., Wagner, T., and Farhat, S.K., 2013, Evaluation of methane sources in groundwater in Northeastern Pennsylvania: Groundwater, v. 51, no. 3, p. 333–349, first published online April 5, 2013.

Mulder, M.L., 2012, Ambient geochemical and isotopic variations in groundwaters across an area of accelerating shale gas development: West Virginia University, M.S. thesis, no. 1520742, 101 p.

Osborn, S.G., Vengosh, A., Warner, N.R., and Jackson, R.B., 2011, Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing: Proceedings of the National Academy of Sciences, USA, v. 108, no. 20, p. 8172–8176, published ahead of print May 9, 2011, doi: 10.1073/pnas.1100682108.

Perry, A.E.; Smith, Bert; Bothun, Rikka; and Hollingsworth, Mark, 2012, The occurrence of methane in shallow groundwater from extensive pre-drill sampling: Ground Water Protection Council, Stray Gas Incidence & Response Forum, July 24-26, 2012, Cleveland, Ohio, http://www.gwpc.org/sites/default/files/event-sessions/Perry_Elizabeth.pdf.

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: Environmental Science and Technology, v. 49, issue 7, pp. 4106-4112.

Sloto, R.A., 2013, Baseline groundwater quality from 20 domestic wells in Sullivan County, Pennsylvania, 2012: U.S. Geological Survey Scientific Investigations Report 2013–5085, 27 p., http://pubs.usgs.gov/sir/2013/5085/.

Susquehanna River Basin Commission, 2012, Remote water quality monitoring network - Data report of baseline conditions for 2010/2011; a Summary: Susquehanna River Basin Commission Publication no. 280, April 5, 2012. http://mdw.srbc.net/remotewaterquality/assets/downloads/pdf/RWQMN_datareport_2010-2011.pdf.

TIBCO, 2008, TIBCO Spotfire S+ 8.1 for Windows User's Guide: TIBCO Software Inc., 572 p.

U.S. Geological Survey, 2003, Principal aquifers of the 48 conterminous United States, Hawaii, Puerto Rico, and the U.S. Virgin Islands, http://nationalatlas.gov/mld/aquifrp.html.

U.S. Environmental Protection Agency, 2012, Study of the potential impacts of hydraulic fracturing on drinking water resources: Progress report: Office of Research and Development, EPA 601/R-12/011, December, 2012, 262 p., http://www2.epa.gov/sites/production/files/documents/hf-report20121214.pdf.

Wilson, B., 2014, Geologic and baseline groundwater evidence for naturally occurring, shallowly sourced, thermogenic gas in northeastern Pennsylvania: American Association of Petroleum Geologists Bulletin, v. 98, no. 2, pp. 373– 394.

Whisman, C., Wardrop, R., Good, D., McElreath, D., Olmsted, C., and Smith, B., 2012, Real-time monitoring system for evaluating long-term variability of methane in domestic water wells in Pennsylvania: Ground Water Protection Council, Stray Gas Incidence & Response Forum, July 24-26, 2012, Cleveland, Ohio, http://www.gwpc.org/sites/default/files/event-sessions/Whisman_Chuck.pdf.

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, v. 340, no. 6134, May 17, 2013: v. 340 no. 6134 DOI: 10.1126/science.1235009 http://www.marcellus.psu.edu/news/PDFs/scienceArticleMay2013.pdf.

16. Appendix B: Water Quality Sampling Requirements to Assess Change: Statistical Power Analysis

By K.H. Reckhow

List of Figures

List of Tables

16.1 Introduction

The Northeast-Midwest Institute (NEMWI), in cooperation with the U.S. Geological Survey (USGS) National Water-Quality Assessment (NAWQA) Program, investigated the ability of the Northeast-Midwest (NEMW) region's water-monitoring programs to inform policy decisions through two case studies, one exploring the effectiveness of management practices for reducing nonpoint sources of nutrients in the Lake Erie Basin and one exploring shale gas development in the Susquehanna River Basin. Both case studies used a statistical power analysis to help quantify the water-quality data that are needed to answer each case-study policy question.

The goal of the statistical power analysis was to determine how many water-quality samples are needed to detect a "signal" (change) in the midst of the noise term, at specified statistical significance and power levels. To do this, deterministic features were modeled in the historic data; these features were expected to be trends, seasonal patterns, and/or a flow effect for each sampling station. These modeled features were then subtracted from the raw data, resulting in the residual standard error, which is the background noise term used in the power analysis. To conduct this analysis, water-quality and flow data were acquired from a number of monitoring stations on rivers flowing into western Lake Erie and in the Susquehanna River Basin. The power analysis method described here was used to estimate the number of water-quality samples needed to detect a change in water quality for both the nutrient case study and the shale gas development case study.

16.2 Method

A time series of historic data for each water-quality parameter and monitoring station of interest is needed for the statistical analysis. It is best if the sampling frequency is at least monthly, and several years of data are needed. These data are necessary to estimate the historic natural variability, which serves as an estimate of the natural variability that may be found in the future data to be collected for assessment of a water-quality change. In essence, an estimate of the background "noise" (natural variability) is needed to estimate how many samples (at a specified sampling frequency) are necessary to detect a "signal" (the change) in the midst of that noise. Once a suitable data set has been acquired, the following factors may need to be assessed in the analysis of the existing data:

- even spacing of observations/missing data
- censored data (i.e., data below detection limits)
- relationship between concentration and flow or stage
- seasonality
- trend
- autocorrelation
- normality

The purpose of the analysis of historic data that precedes the monitoring design (sample size) calculations is the elimination of all components of variability so that the remaining residual data series is stationary (i.e., the remaining variability is background white noise with possible autocorrelation).

Approximately 3 years (or more) of monthly (or more frequent) water-quality data are needed to begin the historic data analysis. The choice of 3 years of data is based on a general approximation of how much data are needed and on expert judgment. The 3-year minimum can be considered a rough approximation and an initial starting point for monitoring. If there are fewer data, or if the sampling is less frequent than monthly, it is more likely that seasonal patterns, historic trends, and background variability will be mis-estimated. This analysis emphasizes monthly data, defined as having a frequency of 25-35 days; anything outside of this frequency might be considered either too frequent (therefore possibly a replicate), or a missing value.

To begin the analysis, one comprehensive regression analysis was conducted on the monthly historic data for each monitoring site to model three deterministic features affecting the water-quality parameter. If the water-quality variable had an asymmetric distribution, the data were first logarithmically transformed to diminish the impact of data points having leverage and influence and to achieve approximate normality of regression residuals. All of the nutrient and streamflow data for the nutrient case study were logarithmically transformed, but most of the specific conductance and barium data for the shale case study were not transformed because these data generally appeared normally-distributed without a transformation. Once the appropriate data were logarithmically transformed, most of the power analysis method was the same for both case studies. The three deterministic features are:

1. **Relationship between concentration and flow** For rivers and streams, it is common to observe that flow and concentration are correlated, often due to either a high flow dilution effect of wastewater discharges or high rainfall-runoff from nonpoint sources. If there is a deterministic relationship, then a concentration-flow regression model term should be part of the analysis. For the current analysis, a concentration-flow model term was used.
2. **Trend in water quality over time** With a relatively short period of record for the historic data, a trend may not be likely. Any trend should be modeled, perhaps using either a linear or polynomial regression model (with time) term, or perhaps using differencing (differencing is based on the changes between time periods). For the current analysis, a linear regression model with time was used.
3. **Seasonal pattern** Modeling can be done using either a deseasonalization technique or more structured sinusoidal regression terms. For this analysis, a structured sinusoidal regression was used per equation 1.

$$Asin\left(\frac{2\pi}{12}\right)t + Bcos\left(\frac{2\pi}{12}\right) \qquad (1)$$

where A and B are regression coefficients (to be estimated) and t is the monthly time index.

In the nutrient case study, censored (e.g., below detection limits) or missing data were found to be small in number, so this issue was ignored here. These sample dates were deleted for most analyses and the impact was ignored.

One comprehensive regression analysis was applied for the three deterministic features listed above, based on a linear model assumption (log-linear when the log transform was applied to variables) the residuals that result from this regression for the water-quality variable were used for the remaining analysis.

The residuals from the historic data regression analysis were determined to be white noise (i.e., no autocorrelation); therefore, the residuals provide an estimate of the background variance (independent of the sampling frequency). If this is not the case, autocorrelation must be estimated in the residuals. Then, using only positive lag autocorrelations that are significantly different from zero at α= 0.05 (or if the lag-2 autocorrelation, r^2, is significant, but lag-1 is not, then use both) estimate the background variance for monthly sampling as (Gilbert, 1987, pg. 43):

$$\sigma_\epsilon{}^2 = s^2/[1 - \left(\frac{2}{n(n-1)}\right)\sum(n-k)r_k] \qquad (2)$$

where s^2 is the residual variance, $\sigma_\epsilon{}^2$ is the autocorrelation-corrected variance, n is the number of samples, k is the autocorrelation lag number (Gilbert, 1987, p. 43), and r_k is the lag-k autocorrelation. In most cases, 0-2 lags will be sufficient. This slight correction to the sample background variance (the "noise" term) is necessary when autocorrelated samples are used to estimate the residual variance. For the analyses conducted for the two case studies, autocorrelation corrections were not necessary as the residuals did not appear to demonstrate much autocorrelation, so equation 2 was not applied.

Finally, sample size (number, n) was estimated using the power functions in Berryman et al. (1988) (equations 3 and 4) for the t-test where α is the significance level (α = type I error probability) and 1-β is the power (β = type II error probability):

$$(1-\beta) = F_g(N_t - Z_{(1-\alpha/2)}) \quad (3)$$

In equation 3, F_g is the cumulative distribution function for the standard t-distribution with n-2 degrees of freedom, $Z_{(1-\alpha/2)}$ is the quantile of the standard normal distribution at probability (1-α/2), and N_t is a dimensionless trend statistic given by, for a linear trend/change:

$$N_t = \frac{|\Delta\mu|\sqrt{n}}{\sigma_\epsilon\sqrt{12}} \quad (4)$$

In equation 4, $\Delta\mu$ is the total change in mean level over the length of the series, n is the number of observations, and σ_ϵ is the standard error of the noise term with mean zero. Solving for n yields the number of samples required.

Equation 5 can be applied for a step trend/change:

$$N_t = \frac{|\mu_1 - \mu_2|\sqrt{n}}{2\sigma_\epsilon} \quad (5)$$

where μ_1 and μ_2 in equation 5 are the means before and after the step change. Equation 5 was not applied for the current analysis.

The seasonal Kendall Tau test (Hirsch et al., 1982; Hirsch and Slack, 1984) for water-quality trend detection may be used in place of equations 4 or 5; tests based on other parametric and nonparametric statistics can also be used. Lettenmaier (1976) has evaluated the adequacy of equations 4 and 5 for design and analysis using the nonparametric Mann-Whitney and Spearman tests. Adequate information on the power function for the seasonal Kendall test is not available, but Berryman et al. (1988) note that equations 4 or 5 provide a reasonable approximation for the seasonal Kendall test (sample size may be slightly underestimated). Thus the sample size can be estimated using the background variance estimate and the power expressions in either equation 4 or 5. This yields the number of independent samples, or the "effective sample size" (n_e) of a larger number (n) of correlated samples. To calculate the actual number of samples when samples are autocorrelated, the equation below can be used (Yevjevich, 1972, p. 188):

$n \approx n_e(1 + 2\sum r_k)$ (6)

where the approximation is good when n>k. In this case, only positive autocorrelations that are significant at the 0.05 level were used. Since monthly samples were used, it was assumed that autocorrelation was not a factor and equation (3) was applied in the analysis. Testing of the sample residuals confirmed that autocorrelation was not a significant factor.

A two-tailed test was applied at this point, with type I and type II errors set equal to each other for simplicity. Contours for 0.10, 0.20, and 0.30 error levels were plotted on a trend magnitude versus sample number plot, with the magnitude of change expressed as percent change for log-transformed data or absolute concentration values for untransformed data.

No suspected outliers were removed in either case study. Robust estimators for deterministic patterns and trends, for autocorrelation and for background variance, could be applied in other applications of this analysis approach.

Monthly samples were required for the power analysis; more-frequent-than-monthly samples were removed from high sampling-frequency data sets to ensure that there would be no autocorrelation in the residuals. Autocorrelation would violate the assumptions used to conduct the power analysis. The sample that fell closest to the 15th of the month was selected, and, if there were multiple samples on that day, the sample that was closest to noon was selected to generate the monthly samples data set.

16.3 An Example: Phosphorus in the Maumee River

Heidelberg University has collected nearly 40 years of daily phosphorus water-quality data on the Maumee River, Ohio. Due to expected autocorrelation in water-quality data collected more frequently than monthly, a monthly subsample of the Maumee data was used, collected on the 15th of each month. The method described in this Appendix was performed for both total phosphorus (TP) and dissolved reactive phosphorus (DRP).

Table B-1 contains the results of a regression analysis for the monthly TP data for the Maumee River. Note that the regression parameters for each of the deterministic terms are significant at the 0.05 level. The residual standard error is 0.161385; this is the background noise term that is used in the power analysis. Figure B-1 shows the relationship between sample size and linear (continuous) change in TP concentration in the Maumee River, as a function of the type I and type II error levels (significance and power) set equal to each other for simplicity of presentation. In this analysis, the samples were assumed to be independent; if autocorrelation is expected, the required sample size will be larger. It is clear from this analysis that a large number of samples is necessary to detect the signal (change in TP) from the noise. Note that the "fractional decrease" in TP concentration has a different scale on the x-axis than does the "fractional increase;" this occurs due to the asymmetry of the TP residuals in the concentration metric. (The concentration metric is the original metric that has not been log transformed.) The remaining analyses for the nutrient case study addressed only fractional decrease, since decrease is expected with management practice implementation.

A similar analysis was undertaken for dissolved reactive phosphorus (DRP) for the Maumee River; Figure B-2 shows the results of the power analysis for DRP in the Maumee. In comparing Figure B-1 and Figure B-2, note that substantially more samples are required to detect a change in DRP than required for detecting the same change in TP. The number of samples required to detect a change in DRP is greater because of the higher variability detected in background noise for log-transformed DRP, in comparison to log-transformed TP. Figure B-3 shows that, under the log-transformation, DRP regression residuals are far more variable than are TP residuals.

In this study, censored (e.g., below detection limits) or missing data were deleted for most analyses and the impact was ignored. However, Heidelberg University maintains zeros and small negative numbers in their data, and they consider zeros to represent what was measured. They do not censor their data at the detection limit (Peter Richards, Heidelberg University, written commun., 2014). To test what effect these values have on the regression analysis results, one regression analysis was completed for the nutrient case study as a sensitivity test, where 0.001 mg/L was substituted for 44 censored dissolved reactive phosphorus samples in the Sandusky River. The results of this sensitivity analysis indicated that including these censored values increased the background variability and consequently increased the residual standard error for this monitoring site. The number of years of sampling required to detect a 10-percent change in median value is on the order of hundreds of years, and tens of years for detecting a 40-percent change for both analyses. Including the censored values for other monitoring sites would not significantly affect the conclusions regarding the amount of DRP data needed to test the case-study hypotheses.

Table B-1. Regression analysis for deterministic features: Maumee TP data.

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	9.098658	2.274665	87.3355	1.14E-52
Residual	383	9.975286	0.026045		
Total	387	19.07394			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-1.59609	0.056336	-28.3319	5.14E-96	-1.70686	-1.48533	-1.70686	-1.48533
time	-0.00042	7.33E-05	-5.66565	2.88E-08	-0.00056	-0.00027	-0.00056	-0.00027
sine(m)	-0.08682	0.013875	-6.25736	1.05E-09	-0.1141	-0.05954	-0.1141	-0.05954
cos(m)	-0.03375	0.011635	-2.90054	0.00394	-0.05662	-0.01087	-0.05662	-0.01087
logFlow	0.290123	0.016562	17.51703	7.02E-51	0.257558	0.322687	0.257558	0.322687

Regression Statistics	
Multiple R	0.690667
R Square	0.47702
Adjusted R Square	0.471558
Standard Error	0.161385
Observations	388

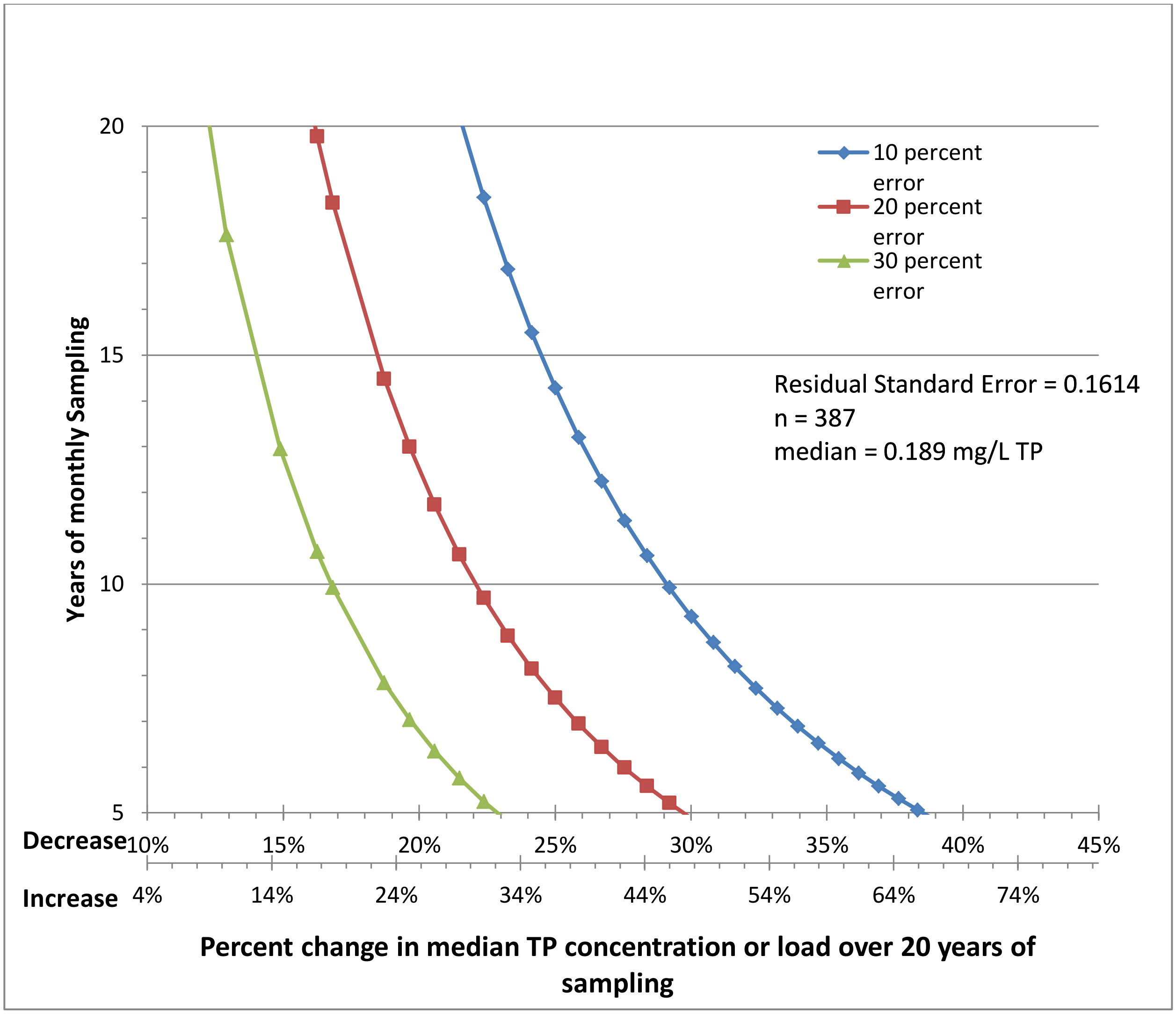


Figure B-1. Power analysis estimates of the number of years of monthly sampling needed to detect trends in median total phosphorus (TP) concentration or load, for different error levels, over 20 years of monitoring at Maumee River at Waterville, OH. The different scales for the percent decrease and increase on the horizontal axis occurs because the log transform introduces an asymmetric distribution in the original concentration metric.
[**Abbreviations:** mg/L, milligrams per liter]

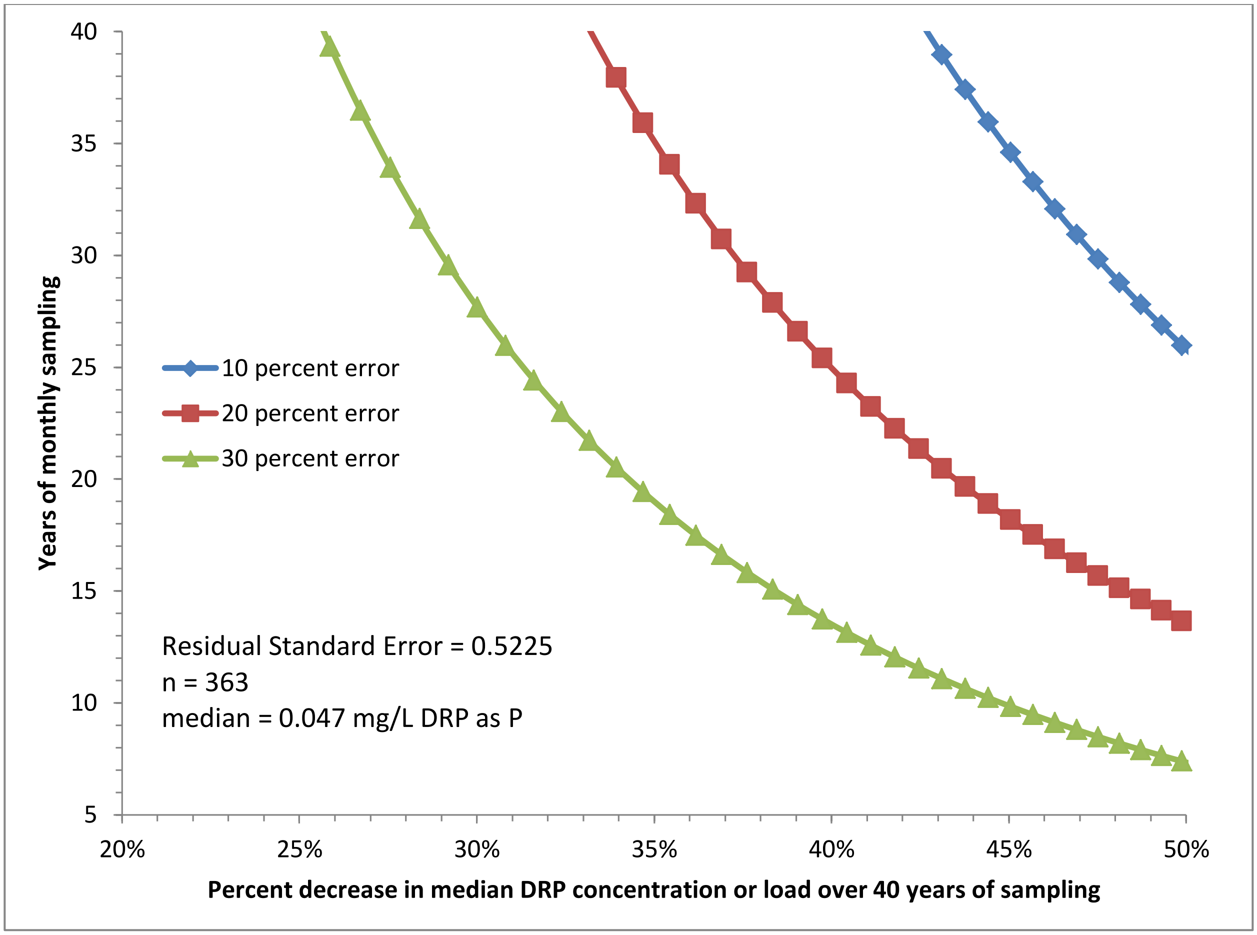


Figure B-2. Power analysis estimates of the number of years of monthly sampling needed to detect trends in median dissolved reactive phosphorus (DRP) concentration or load, for different error levels, over 40 years of monitoring at Maumee River at Waterville, OH.
[Abbreviations mg/L, milligrams per liter]

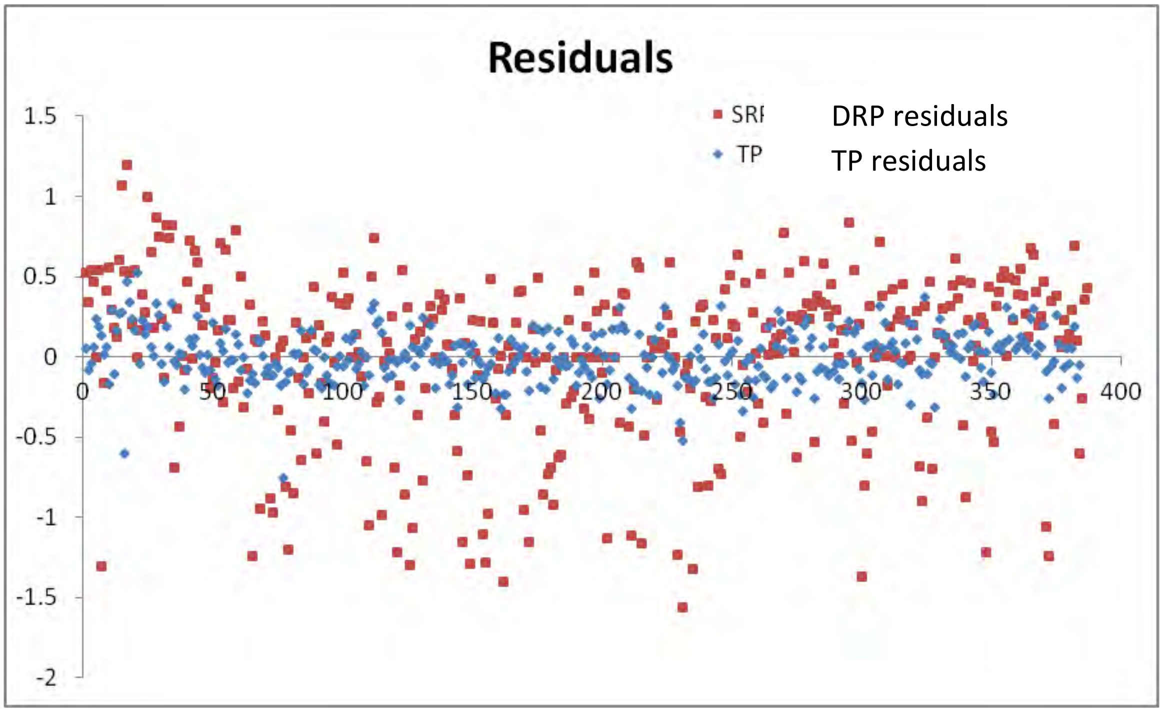


Figure B-3. Comparison of residuals from total phosphorus (TP) and dissolved reactive phosphorus (DRP) regression models.

Hirsch et al. (2010) stated that pollutant flux may be a better measure than pollutant concentration when looking for changes in nonpoint source pollutant loads; this conclusion is relevant when the daily concentration and flux estimates are not both based on a log-linear model for concentration and flow. Since the log-linear flow-concentration model was used here, concentration and flux changes were indistinguishable from each other.

The graphs of sample numbers versus change graphs all reflect a linear change. This is reasonable for assessing the impact of management practice implementation, as it may be expected that the water-quality effects of management practice implementation will be gradual. However, if water-quality monitoring to detect change is discontinued at the start of management practice implementation and is not re-established until sufficient time has elapsed to allow for a water-quality change to be observed, the change could be considered a step function from the perspective of the monitoring results. To illustrate, assume that management practice implementation is completed in 2014, and downstream water-quality monitoring is temporarily terminated at that time. Further, assume that 6 years are required (based on model results and a confidence interval assessment) for the water-quality impact to be measured at a downstream monitoring station. We can then consider the pre-2014 monitoring data to reflect “before” conditions and the post-2020 monitoring data to reflect “after” conditions; this is effectively a step-change assessment. This approach would require only 1/3 the number of independent samples (taken after 2020) as needed for the linear change assessment. In this case, equation 5 would be applied to determine the number of samples required to detect change at a given power and significance.

16.4 Results

Table B-2, Table B-3 and Table B-4 summarize the results of the power analyses completed for the Nutrient Case Study. Table B-5, Table B-6, and Table B-7 summarize the results of the power analyses completed for the Shale Gas Development Case Study.

Table B-2. Summary of total phosphorus and dissolved reactive phosphorus regression analysis results.
[**Abbreviations**: TP, total phosphorus; DRP, dissolved reactive phosphorus; mi^2, square miles; mg/L, milligrams per liter; P, phosphorus]

Monitoring site	Station Name	TP or DRP	Basin Area (mi^2)	Residual Standard Error[1]	Number of samples used in regression analysis[2]	Median Concentration (mg/L as P)
ohUSGS:04193500	Maumee River at Waterville, OH	TP	6013	0.1614	385	0.189
03231500	Scioto River at Chillicothe, OH	TP	3849	0.1607	137	0.279
ohUSGS:04198000	Sandusky River near Fremont, OH	TP	1252	0.2079	398	0.122
inLES060-0005 Flow site 04182000	Saint Mary's River St. Mary's River near Fort Wayne, IN	TP	716	0.1303	50	0.265
04178000	St. Joseph River near Newville, IN	TP	618	0.1908	66	0.130
ohUSGS:04189000	Blanchard River near Findlay, OH	TP	351	0.2344	62	0.215
04186500	Auglaize River near Fort Jennings, OH	TP	331	0.2278	120	0.174
ohUSGS:04199500	Vermilion R Near Vermillion, OH	TP	260	0.3431	118	0.069
ohUSGS:04197100	Honey Creek at Melmore, OH	TP	150	0.2483	409	0.121
ohUSGS:04197170	Rock Creek at Tiffin, OH	TP	37	0.3098	339	0.074
0402913084	Chickasaw River	TP	16	0.273	45	0.248
ohUSGS:04185440	Unnamed Tributary to Lost Creek near Farmer, OH	TP	4.3	0.2325	50	0.078
ohUSGS:04198000	Sandusky River near Fremont, OH	DRP[3]	1252	0.6137	398	0.030
ohUSGS:04193500	Maumee River at Waterville, OH	DRP	6336	0.5225	363	0.047
ohUSGS:04198000	Sandusky River near Fremont, OH	DRP	1252	0.5124	354	0.035
04178000	St. Joseph River near Newville, IN	DRP	618	0.2855	69	0.037
ohUSGS:4199500	Vermilion R Near Vermillion, OH	DRP	260	0.4063	72	0.019
ohBR	Old Woman Creek @ Berlin Road	DRP	56	0.2894	91	0.017
ohUSGS:04197170	Rock Creek at Tiffin, OH	DRP	37	0.4931	307	0.020
ohUSGS:04185440	Unnamed Tributary to Lost Creek near Farmer, OH	DRP	4.3	0.3744	51	0.030

[1] Log base 10-transformed concentration units
[2] The number of observations in the raw data minus observations deleted from regression due to "missingness" (missing data).
[3] Sensitivity analysis

Table B-3. Estimated years of monthly sampling required to detect a decrease in median concentration of total phosphorus (TP) of 10 percent and 40 percent for analyzed watersheds at the 20 percent error level.
[**Abbreviations**: mg/L, milligrams per liter; P, phosphorus; %, percent]

Station name	Residual standard error (log-transformed concentration units)	Median concentration (mg/L as P)	Years of monthly TP sampling required to detect 10% change from median value at the 20% error level	Years of monthly TP sampling required to detect 40% change from median value at the 20% error level
Watersheds larger than 1,000 square miles				
Maumee River at Waterville, OH	0.1614	0.189	55	2
Scioto River at Chillicothe OH[1]	0.1607	0.279	55	2
Sandusky River near Fremont, OH	0.2079	0.122	92	4
Watersheds between 50 and 1,000 square miles				
Saint Mary's River near Ft. Wayne, IN	0.1303	0.265	36	2
St. Joseph River near Newville, IN	0.1908	0.130	77	3
Blanchard River near Findlay, OH	0.2344	0.215	117	5
Auglaize River near Fort Jennings, OH	0.2278	0.174	110	5
Vermilion River Near Vermillion, OH	0.3431	0.069	250	11
Honey Creek at Melmore, OH	0.2483	0.121	131	6
Watersheds less than 50 square miles				
Rock Creek at Tiffin, OH	0.3098	0.074	204	9
Chickasaw River, OH[1]	0.273	0.248	159	7
Unnamed Tributary to Lost Creek near Farmer, OH	0.2325	0.078	115	5

[1]Not in the Lake Erie Basin

Table B-4. Estimated years of monthly sampling required to detect a decrease in concentrations of dissolved reactive phosphorus (DRP) of 10 percent and 40 percent for analyzed watersheds at the 20 percent error level.
[Abbreviations: mg/L, milligrams per liter; P, phosphorus; %, percent]

Station name	Residual standard error (log-transformed concentration units)	Median concentration (mg/L as P)	Years of monthly DRP sampling required to detect 10% change from median value at the 20% error level	Years of monthly DRP sampling required to detect 40% change from median value at the 20% error level
Watersheds larger than 600 square miles				
Maumee River at Waterville, OH	0.5225	0.047	581	25
Sandusky River near Fremont, OH	0.5124	0.035	559	24
St. Joseph River near Newville, IN	0.2855	0.037	173	8
Watersheds between 50 and 600 square miles				
Vermilion River Near Vermillion, OH	0.4063	0.019	351	15
Watersheds less than 50 square miles				
Old Woman Creek at Berlin near Huron, OH	0.2894	0.017	178	8
Rock Creek at Tiffin, OH	0.4931	0.020	517	23
Unnamed Tributary to Lost Creek near Farmer, OH	0.3744	0.030	298	13

Table B-5. Summary of barium and specific conductance regression analysis results.

[**Abbreviations**: mi^2, square miles; µg/L, micrograms per liter; µS/cm, microSiemens per centimeter]

Monitoring site	Station Name	Parameter	Basin Area (mi^2)	Residual Standard Error[1]	Number of samples used in regression analysis[2]	Median Concentration (Barium: µg/L; Specific Conductance: µS/cm)
1538709	West Branch Fishing Creek near Elk Grove, PA	Barium	20.2	0.7504	67	11.6
1545600	Young Womans Creek near Renovo, PA	Barium (diss.)	46.2	4.09	52	29
1548476	Cedar Run above Mine Hole Run near Cedar Run, PA	Barium	26.3	1.508	62	22.4
1533610	Unnamed Tributary to Tunkhannock Creek at Gelatt, PA	Barium	9.02	3.278	33	21.1
1557990	Sinking Run near Spruce Creek, PA	Barium	28.3	5.566	67	26.7
155979602	Bobs Creek below Wallacks Branch at Pavia, PA	Barium	22.1	3.608	74	45.6
1569195	Conodoguinet Creek above Reservoir near Roxbury, PA	Barium	27.2	2.998	68	28
1571820	Swatara Creek at Ravine, PA	Barium (diss.)	43.3	2.337	46	20
1571820	Swatara Creek at Ravine, PA	Barium	43.3	74.76	46	23
1538709	West Branch Fishing Creek near Elk Grove, PA	Spec. Cond.	20.2	2.919	64	27
1545600	Young Womans Creek near Renovo, PA	Spec. Cond.	46.2	4.601	252	39
1548423	Wilson Creek at Morris, PA	Spec. Cond.	22.8	87.29	44	412
1548476	Cedar Run above Mine Hole Run near Cedar Run, PA	Spec. Cond.	26.3	0.08681[3]	59	46
1508800	Factory Brook at Homer, NY	Spec. Cond.	15.8	44.03	59	321
1509150	Gridley Creek above East Virgil, NY	Spec. Cond.	10.4	24.91	71	195
1528000	Five mile Creek near Kanona, NY	Spec. Cond.	66.8	31.32	113	258
BNTY000.9-4276	Bentley Creek at Wellsburg, NY	Spec. Cond.	54.2	27.19	49	193
1557990	Sinking Run near Spruce Creek, PA	Spec. Cond.	28.3	32.77	64	186
155979602	Bobs Creek below Wallacks Branch at Pavia, PA	Spec. Cond.	22.1	9.464	72	76
1569195	Conodoguinet Creek above Reservoir near Roxbury, PA	Spec. Cond.	27.2	13.23	67	65
1571820	Swatara Creek at Ravine, PA	Spec. Cond.	43.3	27.93	101	203
CCPASEC_1216	Lick Run at Howard, PA	Spec. Cond.	11.4	98.12	37	438

[1] Barium: µg/L; Specific Conductance: µS/cm

[2] The number of observations in the raw data minus observations deleted from regression due to missingness

[3] Log transformed; Log base 10-transformed concentration units

Table B-6. Estimated number of samples needed to detect a 20-percent reduction from median values for barium by ecoregion for watersheds of less than 70 square miles.

[**Abbreviations**: mi^2, square miles; µg/L, micrograms per liter; %, percent; ecoregion, U.S. Environmental Protection Agency Level III Ecoregion]

Station number	Station name	Drainage area (mi^2)	Parameter fraction	Minimum concentration (µg/L)[1]	Median concentration (µg/L)[1]	Maximum concentration (µg/L)[1]	Number of barium samples required to detect a 20% change from median values	
							20% error	10% error
Northern Allegheny Plateau ecoregion								
01533610	Unnamed Tributary to Tunkhannock Creek at Gelatt, PA	9.02	Total	16.6	21.1	37.9	33	62
North Central Appalachians ecoregion—total barium								
01538709	West Branch Fishing Creek near Elk Grove, PA	20.2	Total	9.5	11.6	14.6	6	11
01548476	Cedar Run above Mine Hole Run near Cedar Run, PA	26.3	Total	17.2	22.4	31.8	6	12
North Central Appalachians ecoregion								
01545600	Young Womans Creek near Renovo, PA	46.2	Dissolved	20.0	29.0	37.0	27	51
Ridge and Valley ecoregion								
01557990	Sinking Run near Spruce Creek, PA	28.3	Total	17.9	27.0	55.6	59	112
0155979602	Bobs Creek below Wallacks Branch at Pavia, PA	22.1	Total	34.5	45.6	59.4	8	16
01569195	Conodoguinet Creek above Reservoir near Roxbury, PA	27.2	Total	22.8	28.0	40.6	15	29
01571820	Swatara Creek at Ravine, PA	43.3	Total	18.0	23.0	520	14,000[2]	27,000[2]
01571820	Swatara Creek at Ravine, PA	43.3	Dissolved	16.0	20.0	29.0	18	35

[1]Minimum, maximum and median values calculated based on the monthly data set.

[2]Monthly data set includes one potential storm sample resulting in increased background variability.

Table B-7. Estimated number of samples needed to detect a 20-percent reduction from median values for specific conductance levels by ecoregion for watersheds of less than 70 square miles.

[**Abbreviations**: mi^2, square miles; µS/cm, microSiemens per centimeter; %, percent; ecoregion, U.S. Environmental Protection Agency Level III Ecoregion]

Station number	Station name	Drainage area (mi^2)	Minimum specific conductance (µS/cm)	Median specific conductance (µS/cm)	Maximum specific conductance (µS/cm)	Number of specific conductance samples required to detect a 20% change from median values	
						20% error	10% error
Northern Allegheny Plateau ecoregion							
01508800	Factory Brook at Homer, NY	15.8	96	321	420	25	48
01509150	Gridley Creek above East Virgil, NY	10.4	75	195	335	22	42
01528000	Five mile Creek near Kanona, NY	66.8	78	258	450	20	38
BNTY000.9-4276	Bentley Creek at Wellsburg, NY	54.2	114	193	320	27	51
North Central Appalachians ecoregion							
01538709	West Branch Fishing Creek near Elk Grove, PA	20.2	21	27	36	16	30
01545600	Young Womans Creek near Renovo, PA	46.2	25	39	80	19	36
01548423	Wilson Creek at Morris, PA	22.8	130	412	975	61	115
01548476	Cedar Run above Mine Hole Run near Cedar Run, PA	26.3	21	46	176	45	86
Ridge and Valley ecoregion							
01557990	Sinking Run near Spruce Creek, PA	28.3	103	186	266	42	80
0155979602	Bobs Creek below Wallacks Branch at Pavia, PA	22.1	63	76	115	21	40
01569195	Conodoguinet Creek above Reservoir near Roxbury, PA	27.2	37	65	135	56	106
01571820	Swatara Creek at Ravine, PA	43.3	117	203	346	26	50
CCPASEC_1216	Lick Run at Howard, PA	11.4	139	438	635	68	129

[1]Minimum, maximum, and median values are calculated based on the monthly data set.

16.5 References

Berryman, D., Bobee, B., Cluis, D., and Haemmerli, J., 1988, Nonparametric tests for trend detection in water quality time series, Water Resources Bulletin, vol. 24, issue 3, p. 545-56.

Gilbert, R.O., 1987, Statistical Methods for Environmental Pollution Monitoring, Van Nostrand Reinhold, New York, pp. 336.

Hirsch, R.M., Slack, J.R., and Smith, R.A., 1982, Techniques of trend analysis for monthly water quality data, Water Resources Research, vol.18, no. 1, p. 107-21.

Hirsch, R.M., and Slack, J. R., 1984, A nonparametric trend test for seasonal data with serial dependence, Water Resources Research, vol. 20, issue 6, p. 727-32.

Hirsch, R.M., Moyer, D.L., and Archfield, S.A., 2010, Weighted Regressions on Time, Discharge, and Season (WRTDS), with an Application to Chesapeake Bay River Inputs: Journal of the American Water Resources Association, vol. 46, no. 5, p. 857-880.

Lettenmaier, D.P., 1976, Detection of trends in water quality data from records with dependent observations, Water Resources Research, vol. 12, issue 5, p. 1037-46.

Yevjevich, V., 1972, Probability and Statistics in Hydrology. Water Resources Publications, LLC, Fort Collins, Colorado, pp. 312.