EPA Hydraulic Fracturing Study
Quality Assurance Project Plan

Modeling the Impact of Hydraulic Fracturing on Drinking Water Resources Based on Water Acquisition Scenarios: Phase 2

Prepared by:

Kate Sullivan, Ph.D.

Stephen Kraemer, Ph.D.

Christopher Knightes, Ph.D.

Thomas Purucker, Ph.D.

Mark Gabriel, Ph.D.

Katie Price, Ph.D.

Lourdes Prieto

U.S. Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory (NERL)
Ecosystems Research Division (ERD)
Athens, Georgia 30605 USA

Version 1.0

August 30, 2013

Disclaimer

EPA does not consider this internal planning document an official Agency dissemination of information under the Agency's Information Quality Guidelines, because it is not being used to formulate or support a regulation or guidance; or to represent a final Agency decision or position. This planning document describes the overall quality assurance approach that will be used during the research study. Mention of trade names or commercial products in this planning document does not constitute endorsement or recommendation for use.

The EPA Quality System and the HF Research Study

EPA requires that all data collected for the characterization of environmental processes and conditions are of the appropriate type and quality for their intended use. This is accomplished through an Agency-wide quality system for environmental data. Components of the EPA quality system can be found at http://www.epa.gov/quality/. EPA policy is based on the national consensus standard ANSI/ASQ E4-2004 *Quality Systems for Environmental Data and Technology Programs: Requirements with Guidance for Use*. This standard recommends a tiered approach that includes the development and use of Quality Management Plans (QMPs). The organizational units in EPA that generate and/or use environmental data are required to have Agency-approved QMPs. Programmatic QMPs are also written when program managers and their QA staff decide a program is of sufficient complexity to benefit from a QMP, as was done for the study of the potential impacts of hydraulic fracturing (HF) on drinking water resources. The HF QMP describes the program's organizational structure, defines and assigns quality assurance (QA) and quality control (QC) responsibilities, and describes the processes and procedures used to plan, implement and assess the effectiveness of the quality system. The HF QMP is then supported by project-specific QA project plans (QAPPs). The QAPPs provide the technical details and associated QA/QC procedures for the research projects that address questions posed by EPA about the HF water cycle and as described in the *Plan to Study the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources* (EPA/600/R-11/122/November 2011/www.epa.gov/hydraulic fracturing). The results of the research projects will provide the foundation for EPA's 2014 study report.

This QAPP provides information concerning the Water Acquisition Stage of the HF water cycle as found in Figure 1 of the HF QMP and as described in the HF Study Plan. Appendix A of the HF QMP includes the links between the HF Study Plan questions and those QAPPs available at the time the HF QMP was published. This project is the Water Availability Modeling Project and is referred to as **Project 5B Phase 2** throughout this QAPP.

This project is classified as ORD QA Category 1 research and will strictly adhere to all Quality Assurance requirements accordingly.

Acknowledgements

This document builds on QAPP Phase 1 (version 1.0) (revised September 4, 2012) submitted to EPA by the Cadmus Group, Inc.(Laura Blake, Project Manager, Jonathan Koplos, Andy Somor, Corey Godfrey) under EPA Contract EP-C-08-002, Task Order 32 (Stephen Kraemer, EPA WAM) in conjunction with subcontractors AQUA TERRA Consultants (Paul Duda, John Imhoff, Tony Donigian) and Texas A&M University (Debjani Deb, Raghavan Srinivasan). We also acknowledge the contributions in the area of data collection during Phase 1 by Shaw Environmental Consultants (Johnathan Shireman) under EPA Contract EP-C-08-034, Work Assignment HF-2-10 (Susan Mravik, EPA WAM).

Section A – Project Management

A1 Approval Page

Signatures indicate approval of this Quality Assurance Project Plan (QAPP) and commitment to following the procedures noted.

Signature on File **Aug 30, 2013**

Kathleen Sullivan, Ph.D Chief, Ecosystems Assessment Branch Date
EPA Office of Research and Development (ORD), National Exposure Research Lab (NERL)
Ecosystems Research Division (ERD)

Signature on File **Aug 30, 2013**

Roy Sidle, Ph.D. Director, Ecosystems Research Division Date
EPA ORD/NERL/ERD

Signature on File **Aug 30, 2013**

James Kitchens, Quality Assurance Manager, EPA ORD/NERL/ERD Date

This page is blank.

A2 Table of Contents

List of Figures

List of Tables

A3 Distribution List

An electronic copy of this QAPP will be provided to all staff involved in this project, including:

Kate Sullivan, EPA ORD NERL ERD

Stephen Kraemer, EPA ORD NERL ERD A

Chris Knightes, EPA ORD NERL ERD

Thomas Purucker, EPA ORD NERL ERD

Mark Gabriel, EPA ORD NERL ERD

Katie Price, EPA ORD NERL ERD

Lourdes Prieto, EPA ORD NERL ERD

Keewook Kim, ORISE post doc, Athens, GA

Mike Cyterski, EPA ORD NERL ERD

Yusuf Mohamoud, EPA ORD NERL ERD

Alex MacPherson, EPA OAR RTP, NC

James Kitchens, EPA ORD NERL ERD

An electronic copy of this QAPP will also be provided to the following individuals for informational purposes:

John Johnston, EPA ORD NERL ERD Athens, GA

Roy Sidle, EPA ORD NERL ERD

Andrew Gillespie, EPA ORD NERL, RTP, NC

Jeanne Briskin, EPA ORD OSP, Washington, DC

Michelle Henderson, EPA ORD NERL, Cincinnati, OH

A4 Project/Task Organization

Project organization:

ORD HF Study Team

Jeanne Briskin

(Susan Burden)

Jeff Frithsen

(Steve Leduc)

ORD Technical Support

Mike Cyterski (Statistics)

Yusuf Mohamoud (hydrology)

Alex MacPherson (NEMS)

Contract Support:

Keewook Kim, post doc

ORD Quality Team

Michelle Henderson

James Kitchens

NERL Project Team

Kate Sullivan

Steve Kraemer

Thomas Purucker

Chris Knightes

Mark Gabriel

Katie Price

Lourdes Prieto

ORD Communications Team

Dayna Gibbons

Katie Wagner

Emily Smith

ORD Management Team

Jennifer Orme-Zavaleta

(Andy Gillespie)

Roy Sidle

John M. Johnston

Kate Sullivan

ORD Peer Review and Clearance Team

Ramona Trovato

Virginia Houk

Breanne Benton, TIM

Contract Support: ERG

Figure 1. Project organization chart.

Jeanne Briskin (supported by Susan Burden) is the HF Study Coordinator and oversees the planning and budgeting aspects of the study and facilitates communications.

Jeff Frithsen (supported by Steve Leduc) oversees the HF synthesis report to Congress (draft to the Science Advisory Board, December 2014). The water acquisition chapter consumes the research findings coming from this project.

Jennifer Orme-Zavaleta (supported by Andy Gillespie) is the Director of the National Exposure Research Laboratory (NERL) and oversees the NERL management chain of command and the implementation of the HF Quality Management Plan (QMP). The management chain includes Roy Sidle as the Director of the Ecosystems Research Division, John Johnston as the Chief of the Regulatory Support Branch and immediate supervisor to Steve Kraemer, Chris Knightes, Katie Price and Mark Gabriel, and Kate Sullivan as the Chief of the Ecosystems Assessment Branch, and immediate supervisor to Yusuf Mohamoud, Lourdes Prieto, and Thomas Purucker.

Kate Sullivan is the Chief of the Ecosystems Assessment Branch, and Project Lead and is responsible for overall oversight and assists in the preparation, maintenance, and implementation of this QAPP, completes QA Review Forms, works with the QAM on Technical Systems Audits (TSA), manages records, manages review and clearance of all draft and final deliverables, and communicates progress reports with the NERL immediate office and the EPA HF Study leads. Sullivan is a Hydrologist.

Stephen Kraemer is an ERD Research Hydrologist, is leading the ground water modeling components of this study.

Mark Gabriel is an ERD Federal Post Doctoral appointment and Environmental Engineer.

Chris Knightes is an ERD Environmental Engineer and they are supporting the project leading SWAT modeling in the futures modeling of the refined catchments in Garfield County of the Upper Colorado River basin and the Susquehanna River Basin study area.

Katie Price is an ERD Federal Post Doctoral appointment and hydrologist and is supporting the project in the analytical impact analysis in both study areas.

Thomas Purucker is an ERD Research Ecologist with experience in time series analysis and flow model calibration.

Lourdes Prieto is an ERD Physical Scientist and the project's geographical information systems (GIS) analyst.

Yusuf Mohamoud is an ERD Research Hydrologist and is a consultant to the project as an HSPF model expert.

Keewook Kim is an Oak Ridge Institute for Science and Education (ORISE) post doc conducting HSPF modeling.

Mike Cyterski is an ERD Research Ecologist and is a statistics expert and consultant to the project.

Michelle Henderson is the HF Program Quality Assurance Manager (PQAM) and the NERL Director of Quality Assurance (NERL-DQA) and serves as the liaison between the ORD-DQA, the HF Study Team Coordinator, and the ERD Quality Assurance Manager (QAM).

James Kitchens is the Ecosystems Research Division Quality Assurance Manager (QAM) and is responsible for review and approval of this QAPP, conducting Technical Systems Audits (TSA), and reviews of the draft and final deliverables for quality assurance.

Alex MacPherson, Environmental Economist, OAR/OAQPS/HEID in RTP, NC is supporting the project in the extraction of the Energy Information Administration NEMS data.

The HF Peer Review and Clearance Team is lead by Ramona Trovato, ORD Associate Assistance Administrator. The HF Peer Review Coordinator is Virginia Houk and she manages the records of the EPA Science Inventory. The ERD Technical Information Manager (TIM) is Breanne Benton and she manages the records in the Science and Technical Information Clearance System (STICS). ERG will assist in the performance of technical reviews of the reports.

The HF Communications Team is lead by Dayna Gibbons and supported by Katie Wagner and Emily Smith.

A5 Problem Definition and Background

Introduction

Natural gas plays a key role in our nation's energy future. Recent advances in drilling technologies—including horizontal drilling and hydraulic fracturing—have made vast reserves of natural gas economically recoverable in many areas of the United States. First introduced in the 1940s, hydraulic fracturing is designed to increase the permeability of shale, coalbeds, and tight sands buried deep below the earth's surface, allowing extraction of oil and natural gas trapped within these unconventional reservoirs. The hydraulic fracturing process involves injecting water mixed with a proppant and a variety of chemicals under sufficient pressure to induce and maintain fractures through which oil and gas can flow to a producing wellbore (API, 2010). Some of the water/chemical mixture flows back to the surface where it must be treated and/or disposed of following the fracturing operation.

Hydraulic fracturing (HF) has expanded dramatically in recent years to allow commercial production in the unconventional oil/gas "plays" distributed widely throughout North America (Figure 2). As use of HF has accelerated and spread to new areas, so have concerns about its potential impact on human health and the environment, especially the possible effects on the quality and quantity of drinking water resources.

EPA has initiated a multifaceted research program to elucidate potential impacts of hydraulic fracturing on drinking water resources and to identify factors that affect their severity and frequency (www.epa.gov/hfstudy).

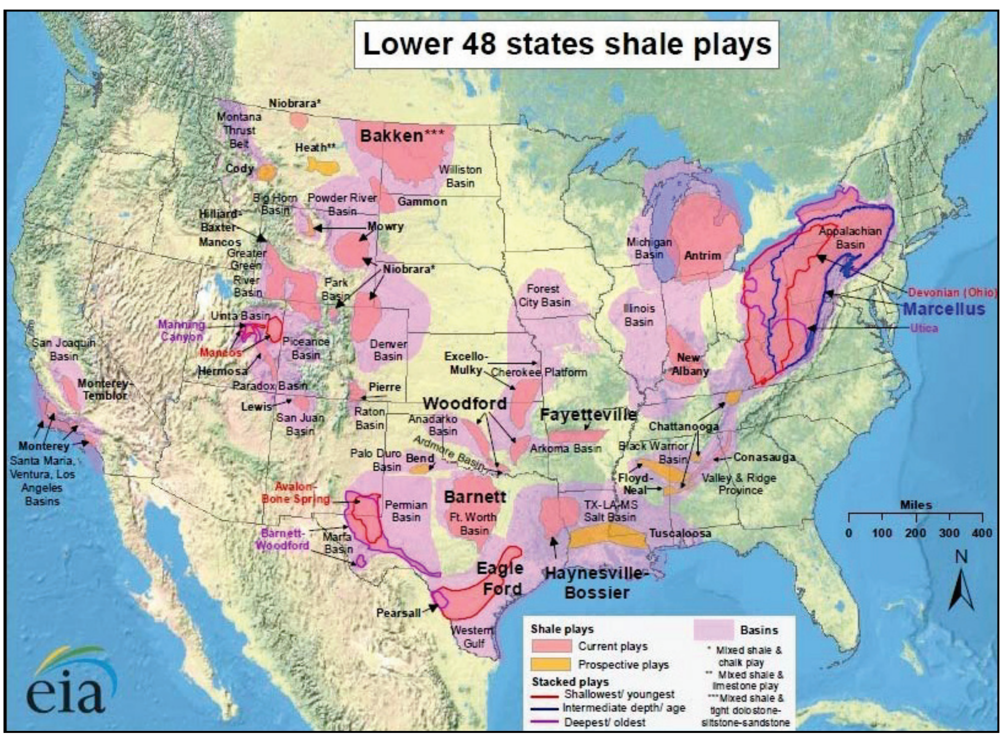


Figure 2. Location of unconventional shale plays in the United States

EPA's research is organized around the five major stages of the water cycle associated with hydraulic fracturing operations: (1) water acquisition; (2) chemical mixing; (3) well injection; (4) flow back and produced water; and (5) wastewater treatment and waste disposal (Figure 3).

This Quality Assurance Project Plan (QAPP) is one of a group of projects contributing to primary and secondary Water Acquisition questions:

What are the possible impacts of large volume water withdrawals from ground and surface water on drinking water resources?

> *How much water is used in HF operations and what are its sources?*
>
> *How might water withdrawals affect short- and long-term water availability in an area with hydraulic fracturing activity?*
>
> *What are possible impacts of water withdrawals for HF operations on local water quality?*

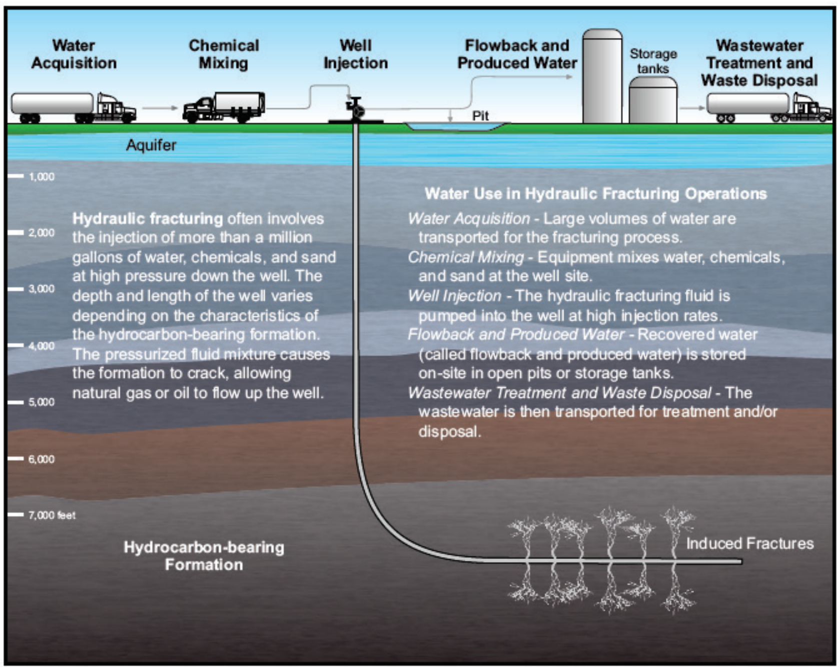


Figure 3. Illustration of the five stages of the hydraulic fracturing water cycle. The cycle includes acquisition of water for the hydraulic fracturing fluid, onsite mixing of chemicals and water to create the hydraulic fracturing fluid, injection of the fluid under high pressures to fracture the oil- or gas-containing formation, recovery of flow back and produced water (hydraulic fracturing wastewater) after the injection is complete, and treatment and/or disposal of the wastewater (from EPA, 2012).

Oil and gas development within a region is one of many consumptive uses, including domestic water supply, competing for available water resources. Water use must be balanced against water availability over time to sustain human, ecosystem and economic values, and is expressed as:

Water Available at Source – Ecosystem Needs = Water Available for Consumption for all uses

The potential impacts of HF on water supply in an area depend on:

- water availability reflecting local geology, hydrology and climate,
- scale of hydraulic fracturing operations including rate of development and well needs, and

- competing demands for water including drinking water, agricultural and industrial uses and ecosystem requirements.

All of these factors vary regionally and locally within the United States.

Overview of Hydraulic Fracturing and Water Use

To evaluate potential impacts, it is essential to understand the hydraulic fracturing process and how water is utilized and sourced. The following brief description is based on information available in the scientific literature provided during EPA and National Academy of Sciences workshops since 2012 and from industry and organization web sites.

The Hydraulic Fracturing Process. Oil and gas wells are drilled vertically to great depths (2,000 to 10,000 ft) and then horizontally (1/3 to 1 mile or more) to allow extraction from a greater area within the rock formation. Once drilled, the well intervals are cased with steel and cement and producing intervals are then "stimulated" to release the gas or oil tightly held within fine-grained matrices of the rock. Hydraulic fracturing involves injecting a water/proppant/chemical mixture into the well under high pressure. The pumping process expands existing fissures or cracks or creates new ones that allow the oil or gas to seep slowly from rocks into the well. The proppant materials comprised of sand, ceramics or other inert particulates keep the fissures and cracks from closing. A variety of chemicals may be added to the water/proppant mix such as friction reducers, scale inhibitors and biocides (Vidic *et al*., 2013). The Proppant generally makes up about 9% and chemicals 1% or less of the total injected volume.

Well operations are centered at "well pads" that are generally about two acres in size (Figures 4 and 5). Activities conducted at the pad include water storage, chemical mixing, and water collection and, possibly, treatment facilities. Several individual wells can be drilled from one well pad which minimizes the construction footprint on the landscape, including access roads.

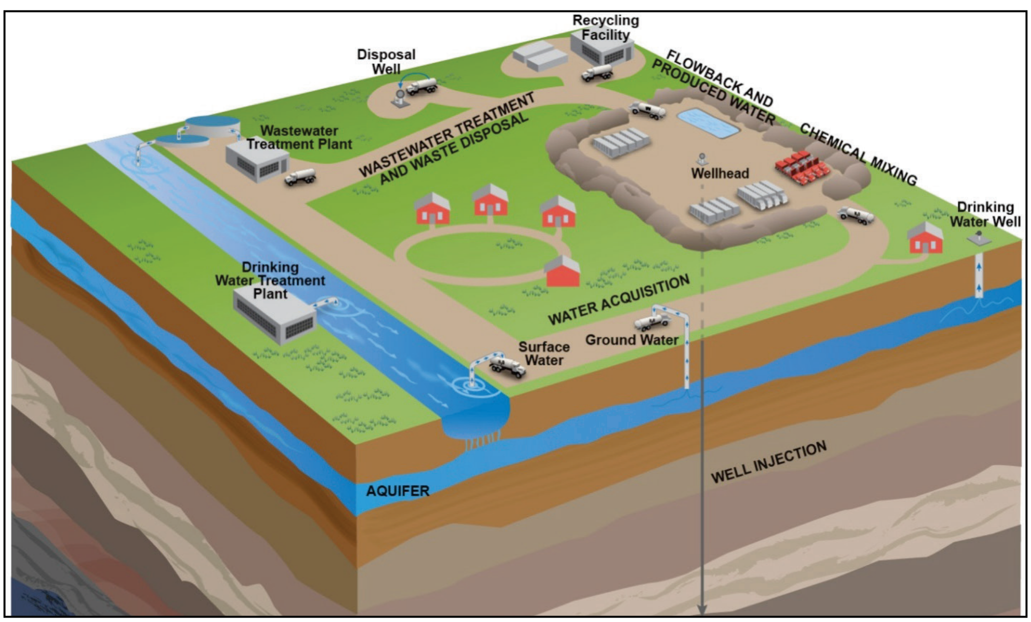


Figure 4. Sketch of typical hydraulic fracturing well pad.

Drilling and constructing each well may take as little as several months or up to a year to complete, although the actual fracturing process that stimulates the well is usually completed in five to seven days. Large volumes of water are needed onsite to fracture each well during this roughly week-long period due to the vertical and lateral extents of the wellbores and characteristics of the rock formations.

Figure 5. Photograph of well pad with active hydraulic fracturing operations in the Marcellus shale. The upper reservoir is probably produced water.

Water is usually trucked to the site where it must be stored in temporary reservoirs or tanks, as shown in the photograph. Trucking water to and from the well site is a major cost in well development. Non-potable, low quality water can be used for hydraulic fracturing. Once the well is in service, there is very little need for additional water. Some water is used during drilling.

The HF fluid is injected into the well under high pressure. After the pressure is released, the wellhead valve is opened and "flow back water" is collected. After the well begins production (generally about 30 days after fracturing), the flow back water is termed "produced" water. Most of the water is recovered as flow back water, but some produced water will be recovered throughout the life of the well, which could be decades. The fraction of flow back water recovered from each well varies from region to region, and may be as small as 10% in the Marcellus shale in the northeastern US (Vidic *et al.*, 2013) or as large as 80% in the tight sands of the Piceance play in Colorado (Cadmus, 2012b).

Flow back and produced water must be disposed of and potential effects on water quality from spills or treatment of the water/chemical mixture is a concern (Vidic *et al.*, 2013).

Water Use for Hydraulic Fracturing. EPA's water acquisition study focuses only on acquiring water for the hydraulic fracturing process. The amount of water needed depends both on the volume used per well and the amount of water required, in aggregate, for broader, long-term, area-wide development programs anticipated by individual companies in each play (API, 2010). The amount of water injected into wells varies significantly among areas, depending on characteristics of the formation being fractured (e.g., coalbed, shale or tight sands) and design of the production well and fracturing operation (e.g., depth, length, vertical or directional drilling) (GWPC and ALL Consulting, 2009). Estimates of water needs per well have been reported as high as 13 million gallons for shale gas production. For perspective, five million gallons of water are the equivalent amount of water used by approximately 50,000 people in one day. The water usage for hydraulic fracturing in shale gas plays are two orders of magnitude greater than more conventional energy sources such as coalbed methane reservoirs which use about 65,000 gallons per well.

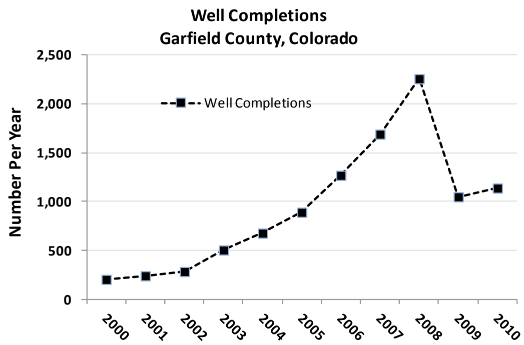
Responding to an EPA request for information, nine oil and gas operators provided data on water management from 330 wells fractured between 2009 and 2010 in many of the most intensively developed hydraulic fracturing areas. Data from Garfield and Mesa Counties in Colorado shows water use per well ranges between one to nine million gallons, with a median of 1.3 million gallons (GWPC, 2012). Individual well water usage in the Marcellus shale in Pennsylvania ranges from two to four million gallons (API, 2010; GWPC and ALL Consulting, 2009; and Satterfield *et al*,. 2008). These estimates are consistent with industry values reported in the FracFocus database (GWPCC and IOGCC, 2013).

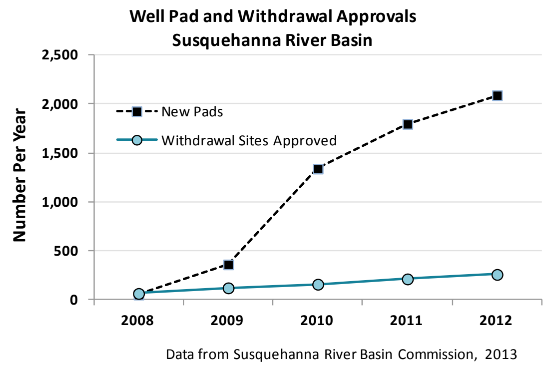
Area-Wide Water Use. Well pads are spaced systematically within the landscape to ensure maximum subsurface extraction of gas from the formation, with the pattern depending on well characteristics and topography. Final density of wells within a play reflects the maximum reach of horizontal drilling and well productivity. In the close-up area shown in the example photograph of a mature well field in the Uinta Basin in Utah (Figure 6), wells are spaced approximately 700 ft apart, yielding a local density of about one well per 11 acres. Oil and gas developer Range Resources projects the final well density in the Marcellus shale in Pennsylvania at one per 80 acres (www.rangeresources.com).

Figure 6. Mature well field in the Uinta Basin, Utah. Density of wells within area in left photo is 1 per 11 acres. Right photo is view from 6000 ft elevation. Images from GoogleEarth.

While hydraulic fracturing has been utilized for decades and some plays are well-developed, the recent pace of well drilling has increased in many areas and will likely continue to accelerate over the next several decades. With increased drilling comes increased demand for water to support drilling, along with additional demands from population growth or industrialization. Figure 7 shows annual well completion in recent years for the Marcellus shale of Pennsylvania (Richenderfer, 2013), and the Piceance play in western Colorado. Phase 1 of the water acquisition project estimated a maximum future development rates based on NEMS and USGS energy development projections that indicate drilling will continue to increase (Cadmus Group, 2012a,b). Drilling rate has increased dramatically to about 2000 per year in the Marcellus shale in the last several years. Water sources required to supply the drilling must increase commensurately where wells use significant amounts of “new” water with low rates of recycling.

Figure 7. Annual rate of well completions in Garfield County, CO and the SRB in the Marcellus shale.





Water Sources. Water is obtained from surface water (rivers, streams, lakes, reservoirs) or from groundwater (wells). Consumers either purchase water from a "public" supply or are "self-supplied". The U.S. Geological Survey defines a source as "public" if water is delivered by a public or private entity to at least 25 customers for domestic use, public services or commercial, industrial, or agricultural purposes. "Self-supplied" sources include user-direct withdrawals from surface water or private wells.

Table 1. Potential hydraulic fracturing water sources.

	Self-Supplied	**Public**
Groundwater	Yes	Yes
Surface Water	Yes	Yes

Water used for hydraulic fracturing operations is typically obtained from a mix of sources in the area including surface or groundwater, and public or self-supplied (Table 1). Hydraulic fracturing operations can use low quality surface water or more brackish or saline water than is required for domestic water consumption, and thus are not always in direct competition for domestic supplies that require high quality water. The quality of water needed for HF depends on other chemicals in the fracturing fluid formulations, availability of recycled HF fluids, and chemical and physical properties of the formations.

Every five years, the U.S. Geological Survey conducts a national survey of water use in the United States. Data are reported by county, state and nationally with the most recent report updating water use to 2005 (Kenny *et al.,* 2009). The following facts on water use at a national level are taken from this report.

- An estimated 258 million people, 86% of the population, rely on public water supplies for household use. The proportion of the population drawing water from public sources has increased over time.
- At a national level, two-thirds of water withdrawn for public supply in 2005 was from surface water such as lakes and streams. Just 15 states obtained more than half their public water supplies from groundwater.

- An estimated 43 million people in the United States, 14 percent of the population, supplied their own water for domestic use. Most of them live in rural areas and obtain their supplies from wells.
- The weighted national average per capita water use for domestic deliveries from public suppliers was 98 gal/day. Generally, per capita use is greater in arid regions compared to humid regions.

The total volume of water used and relative amount drawn from surface versus groundwater and public versus private sources varies by state and locally. Figure 8 shows the daily consumption of ground versus surface water for all uses in selected states within unconventional gas and oil regions.

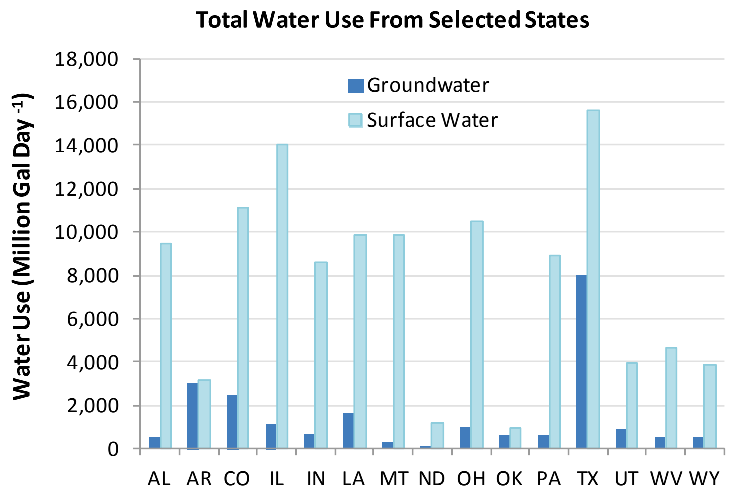


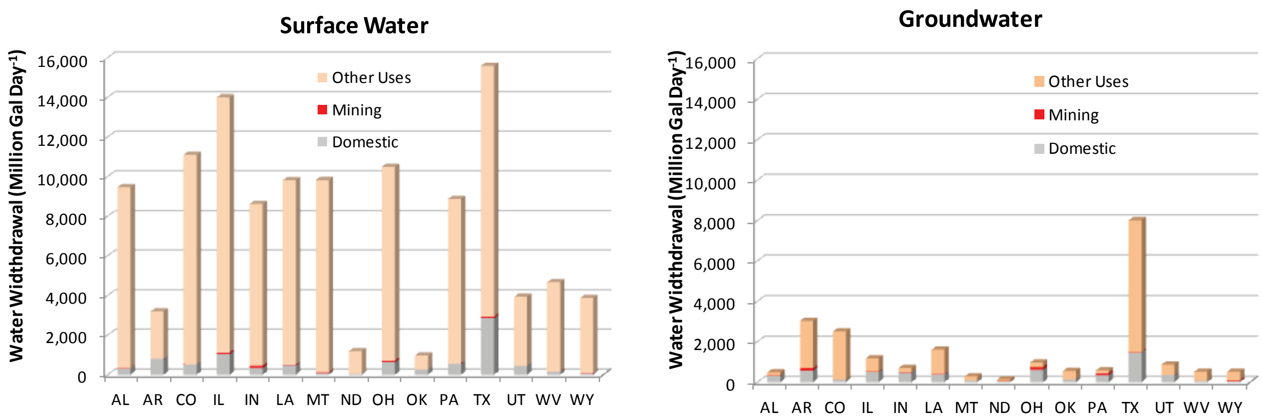
Figure 8. Surface and groundwater use rates in selected states with unconventional oil and gas reserves (USGS data reported in Kenny *et al*., 2009).

The oil and gas industry increasingly is treating and recycling flow back and produced water and reusing it in new wells. The extent of water reused varies regionally and by operator. Up to 10% of the original water pumped into a well may be recycled from other wells in the Marcellus shale region and up to 80% is reused in the Colorado Piceance. Water recyling is increasing throughout the industry and constitutes a significant portion of HF water for some companies, almost eliminating their need for “new” water. Recycling can have significant advantages in reducing acquisition and trucking costs associated with water management. Increased recyling is also expected to reduce potential impacts of HF on water acquisition demand as the pace of well development increases.

Local conditions and allocations determine availability of ground and surface water resources in proximity to planned operations (API, 2010). The options for acquiring water will depend upon volume and water quality requirements for HF in a given play, physical availability, competing uses, and regulatory constraints. Not all options may be available in all situations.

Water Users. There are many water users that draw continuously or episodically from public and self-supplied sources. Large volumes of water are used for farming, including irrigation and animal husbandry (livestock, aquaculture), and for industry, including manufacturing, mining, and thermoelectric power generation (Kenny *et al*., 2009). HF operations may also use secondary waters such as cooling water from power plants or treated wastewater from municipal and industrial treatment facilities (API, 2010). Since ecosystems also rely on water, maintaining integrity of ecosystem services in surface waters requires a minimum maintenance flow, or “passby” flow. Ecosystem services are considered a water user in this project.

Figure 9. Water consumption from surface and groundwater sources showing domestic, mining, and all other uses combined. U.S.G.S. data from Kenny *et al*. (2009).



Daily water consumption by combined use categories for many states with hydraulic fracturing activity is shown in Figure 9, based on USGS data (Kenny *et al*., 2009). Water volume withdrawals, sources, and uses vary widely among states. Domestic supply uses a relatively small amount of water compared to other users, while irrigation (included in "Other") is generally the largest user of both ground and surface water in most states. As of 2005, the relative amount of water consumed by mining, including hydraulic fracturing, is less than 1% of the total aggregated at the state level (barely visible on the bar charts.)

Water Stress Analysis. The basic approach to analyze potential effects of hydraulic fracturing on water availability is to quantify the balance between water supply and demand:

Water Available at Source – Ecosystem Needs = Water Available for Consumption for All Uses

Equation 1

Any area where water use approaches or exceeds what is available is under "water stress". Various analyses of this balance between supply and demand have been conducted for all water uses (e.g., Tidwell *et al.*, 2012; Tidwell, 2013), with specific focus on the role of hydraulic fracturing (Nicot, 2013). When the ratio of consumption to availability is compared at large spatial scales (e.g., states) and summed over long time frames (e.g., one year), hydraulic fracturing water use is a small fraction of water consumption, consistent with the low volume relative to other uses shown in Figure 9.

At the EPA Water Acquisition Workshop held in June 2013, Tidwell presented a national-scale map that showed the balance between consumption for all users and available water, based on data available at the state, county and watershed scales (Figure 10). The map expresses a water stress index computed as:

$$\boldsymbol{Water\ Stress = \ Volume\ Consumed/Volume\ Available}$$

Equation 2

Values approaching 1 indicate complete consumption of available water. It is clear in the Tidwell (2013) analysis (Figure 10) that consumption/availability balance does not adhere strictly to state boundaries. Not surprisingly, surface water stress is chronic in the arid western United States and in areas with significant agricultural activity and large population centers. Many of these also coincide with major unconventional oil and gas plays (Figure 2). Groundwater stress is particularly significant in the southwestern U.S and the Great Plains where groundwater from major regional aquifers are an important source of water.

Climatic fluctuations have a prominent effect on water withdrawals, particularly for irrigation, thermoelectric power generation, and public supply. Consumption imbalance can be more consequential locally and during times of year when water demands are higher, as when irrigation is active and surface water flow naturally is lower. Periodic droughts have also drawn attention to the limits of local and regional water supplies (Kenny *et al.*, 2009).

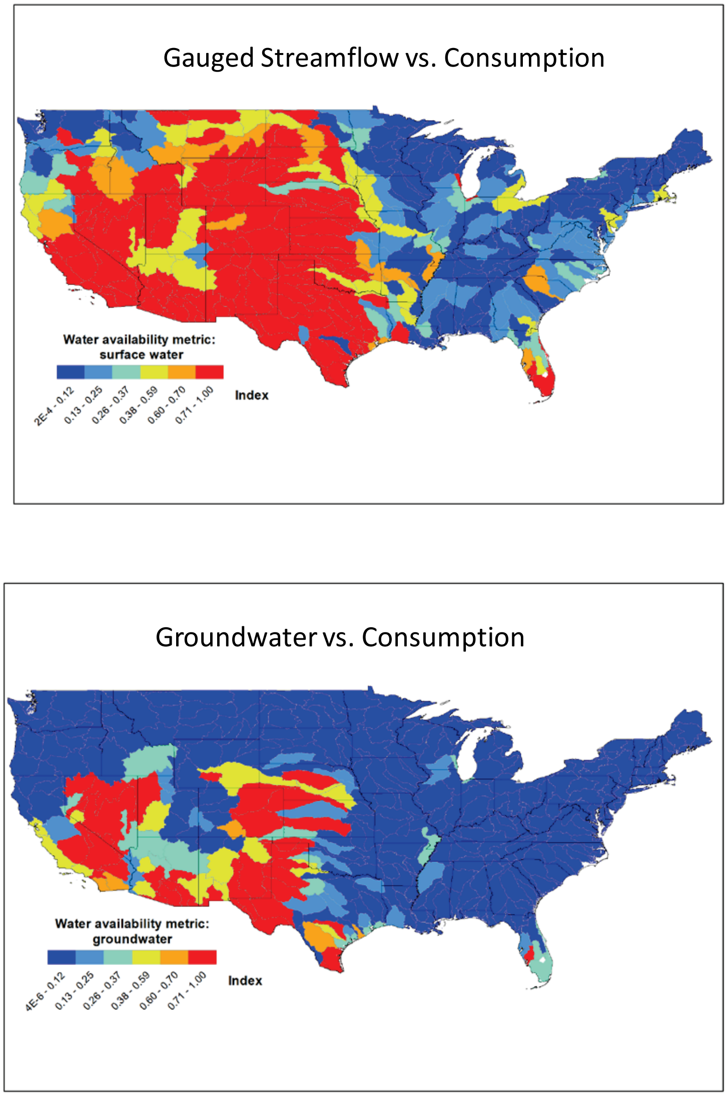


Figure 10. Water stress maps for surface water and groundwater illustrating consumption from all uses relative to water availability (from Tidwell *et al.* EPA Workshop, June 4, 2013).

Overview of the EPA Project on Water Acquisition Scenario Modeling

This project will study potential impacts of water acquisition on domestic water supplies and explore the balance of water consumption and availability, focusing on hydraulic fracturing as a growing consumptive use. The potential impacts of water acquisition will be studied in two watersheds where hydraulic fracturing activity has increased in recent years. The Susquehanna River Basin (SRB) overlying the Marcellus Shale gas reservoir is located primarily in Pennsylvania and New York and represents humid eastern climate (Figures 2 and 11). The Upper Colorado River Basin (UCRB) in semi-arid western Colorado overlies the Piceance structural basin and tight gas of the Williams Fork formation (Figures 2 and 11). These watersheds were selected because of the high current and projected rates of hydraulic fracturing activity anticipated over the next several decades, and because the EPA has previously calibrated and tested watershed models in these areas to investigate future climate change impacts on watershed hydrology (the "20 watersheds study"; Johnson *et al*., 2012).

This project utilizes the water stress approach by determining consumption relative to surface and ground water availability at the large basin (Phase 1) and small catchment (Phase 2) scales. Phase 1 assessed relative impacts of hydraulic fracturing water withdrawals in the two major river basins exceeding 18,000 mi^2 (Upper Colorado River and the Susquehanna River) at basically an annual scale of resolution. This project was conducted by extramural contractors who completed a draft report in 2012 currently under review (Cadmus Group, 2012a,b). The preliminary results appear to be consistent with other studies conducted at state and national scales (e.g., Nicot, 2013) that have found that hydraulic fracturing has little or no impact on water availability at these scales. Most elements of this Phase 2 project build on the Phase 1 study.

The Phase 2 study of HF water acquisition scenarios to be conducted by EPA/ORD will increase the spatial and temporal granularity of the consumption versus availability analysis by narrowing geographic scope to a continuum of watershed sizes less than 250 mi^2 located primarily in Bradford County within the Susquehanna basin and on Garfield County (approximately 2020 mi^2) within the Upper Colorado basin. The EPA project team will assess water consumption at local spatial scales within sub-watersheds and at temporal scales relevant to actual water use in an area (days to months).

A number of criteria were applied to narrow the analysis area within the major basins to ensure that results would be most relevant to HF scenario analysis and transferable to areas of high HF potential. The selection criteria and refined project areas are described in detail in Appendix A. Criteria included:

- Currently active long-term USGS streamflow gages within the watershed and at least one groundwater monitoring well in its vicinity
- Currently active long-term NWS hourly stations inside and around the watershed
- No major surface water reservoirs
- Contains drinking water supply intakes with detailed consumption data
- Located in a region of active and increasing pace of HF drilling activity

Components of the analysis defined in project tasks are briefly summarized in Table 2. The remainder of section A5 describes in some detail common steps of analysis to assess potential hydraulic fracturing impacts on water availability in both study areas. Modeling and calibration steps and methods applied in the project areas are further detailed in Appendix A.

Table 2. Brief description of project tasks for the Project 5B--Phase 2 refined water acquisition study.

Sub Task	Task	Objective/Work
1	**Quantify Available Water Resources**	Determine availability of surface water and groundwater over a range of flow levels at gaged and ungaged streams with watershed size from headwaters to HUC 12. Apply watershed and groundwater models and empirical statistical relationships.
2	**Quantify Consumptive Use**	Determine current cumulative surface water and groundwater use at withdrawal locations distributed within test watersheds. Available data sources such as USGS, States, SRBC and COGCC and spatial statistical modeling as needed.
3	**HF Scenario Analysis**	Gather detailed information from HF operators, regulatory agencies, the USGS, the literature, and EPA studies on specific activities used to guide water acquisition operations. Information will be synthesized to develop impact analysis scenarios. Compare four HF scenarios of water use on water availability within watersheds at a range of flow conditions. Identify noteworthy differences in water stress for: • baseline-pre HF • current HF development rate • "energy plus" estimating well development within an area during peak drilling • "recycling technology" reflecting recycling of produced and flow back water
4	**Report**	Presentation of results for each project area and synthesis of physiographic, demographic and HF management leading to higher vulnerability of water supplies. This will include a project report and at least one journal publication.

Figure 11. Water acquisition project areas.

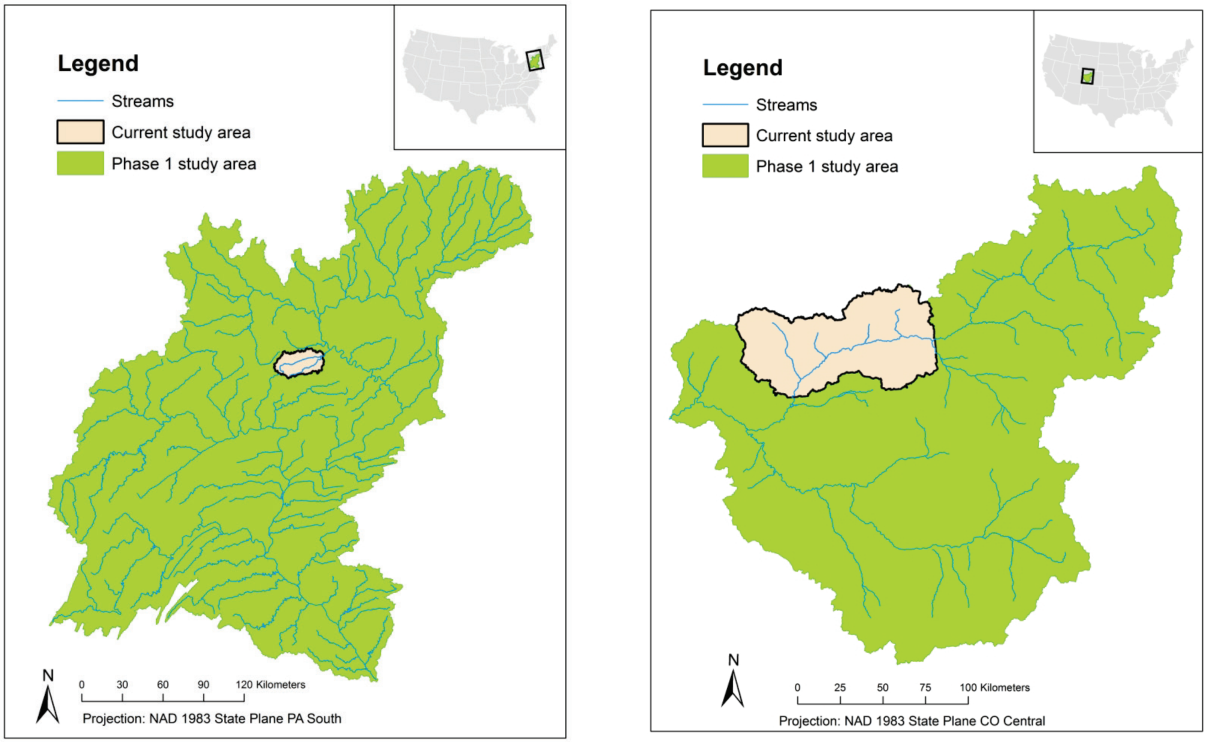


Table 3. Project implementation background.

Project 5B	Study Feature	Susquehanna River Basin (SRB)	Upper Colorado River Basin (UCRB)
Large Basin Study (Phase 1)	Total Basin Area	27,510 mi^2	17,800 mi^2
	Focal Area	Total basin	Total basin
	Project Team	Extramural (Cadmus + Aqua Terra)	Extramural (Cadmus + Texas A&M)
	Analytical Model	HSPF	SWAT
	Status	Draft Report in review Cadmus (2012a)	Draft Report in Review Cadmus (2012b)
Refined Analysis (Phase 2)	Focal Area	Bradford County—Towonda Creek	Garfield County
	Total Basin Area	215 mi^2	2021 mi^2
	Watershed size	Small headwaters to HUC 12	Small headwaters to HUC 12
	Temporal Scale	Short-term to annual	Short-term to annual
	Project Team	EPA-ORD	EPA-ORD
	Hydro Model	SWAT, HSPF	SWAT, HSPF
	Status	Start August 2013	Start August 2013

Water Availability Analysis (Task 1)

First we describe our conceptual approach to the analysis of water stress, then the details of how water availability analysis will be completed.

Conceptual Overview

The basis of the analysis strategy is the physical water balance (Figure 12) within a given area and time period. Water balance in its simplistic form is:

$$IN - OUT = CHANGE\ IN\ STORAGE$$

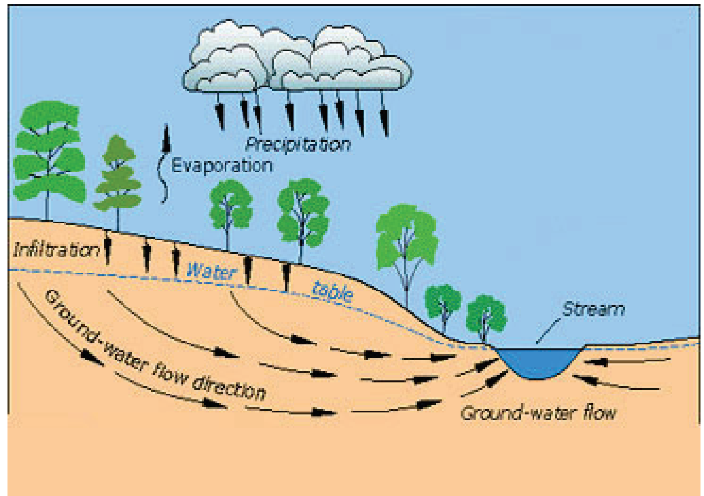


Figure 12. Schematic of water balance. Graphic from GroundwaterCommons.com.

The project objective is to characterize the volume of surface water in rivers and streams (and any temporary reservoirs constructed on them) and in relatively shallow saturated subsurface aquifers. Water found in these "buckets" is what is available for human consumption. Precipitation provides the input of water and initiates the water cycle. Some input water input is unavailable for consumption as it is lost to the atmosphere as evapotranspiration or to the deep earth through leakage, and some is held in unsaturated soils by tension. Surface water and subsurface water are intimately associated and are in a continuous process of exchange (Dunne and Leopold, 1978.) The total volume of available water above and below ground varies over time in response to episodic rainfall and through slow exchange from subsurface storage to surface flow during intervening dry periods.

Groundwater in the saturated zone below the water table contains 80% of the unfrozen fresh water in the United States and is an important source of water. Local flow systems in shallow unconfined aquifers that provide water to relatively shallow wells are the focus of groundwater assessment in this project (Figure 13). Very deep and/or regional aquifers tend to either have lower water quality and less value as a drinking water source or they are beyond the project's analytical scope.

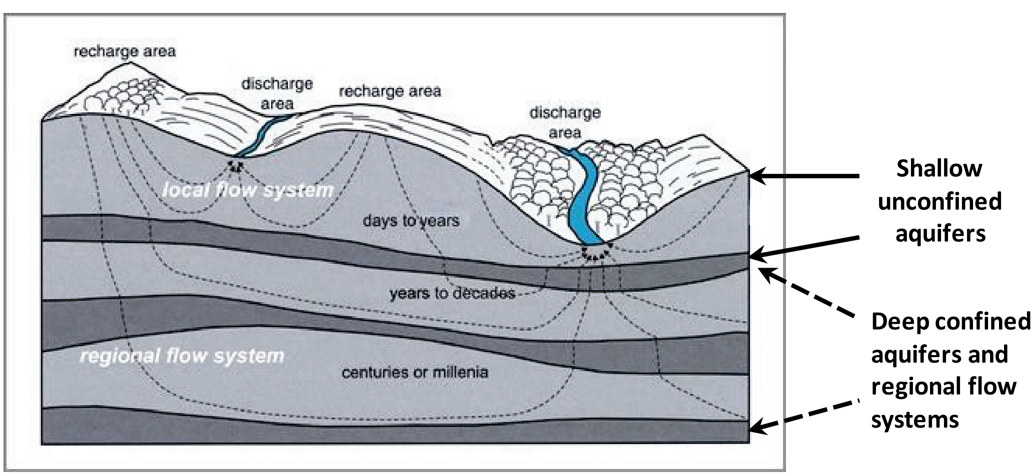


Figure 13. Sketch of groundwater flow systems. (Graphic from U.S. Geological Survey).

Water stress arises when the consumptive demand that drives the rate of water withdrawal exceeds what is available in the streams or groundwater aquifers within the time span of natural replenishment. The three factors that determine the vulnerability of the surface or subsurface water source are volume of available water and the timing and volume of withdrawal. The likelihood of water stress is lower when there is larger volume in the "bucket", or when the volume of water consumed is lower or withdrawn over a longer period. Storage within the water supply system alleviates short-term pressures during high use or lower flow periods.

Our conceptual framework for analysis of general vulnerability to water consumption in natural systems is illustrated in Figure 14. The potential vulnerability (Z-axis) is a water stress metric that expresses potential vulnerability to water withdrawal relative to availability, such as the ratio of consumption to available water used by Tidwell (2013) as shown in Figure 10. Spatial scale on the X-axis uses contributing watershed area as an index of available water in either the surface or subsurface storage reservoir based on the assumption that the larger the water- or groundshed, the greater the volume of water within it. Rivers with large drainage areas or large subsurface aquifers have enough available water that they can accommodate considerable consumptive use.

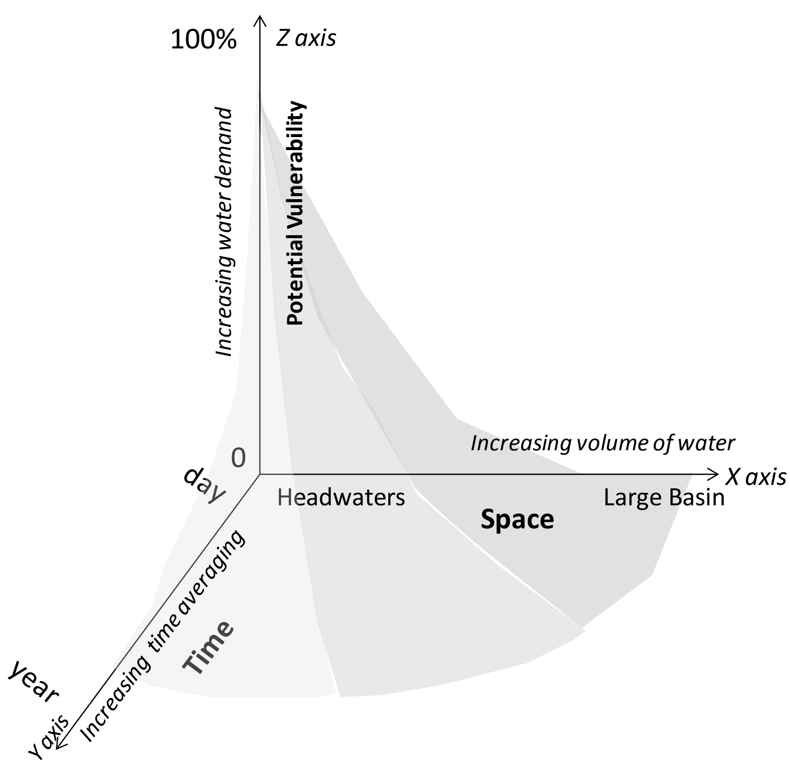


Figure 14. Conceptual framework for potential impact of water withdrawal in relation to temporal and spatial scales.

The Y-axis represents temporal scales from daily to yearly.

Assuming some large volume of water is to be withdrawn to fracture a well, the smaller the stream or aquifer sourcing the water, and/or the shorter the time frame during which the water is taken, the greater the potential for withdrawing a significant portion of available water and the greater the vulnerability to over withdrawal.

The three-dimensional space/time continuum in Figure 14 conveys the principles of analysis. Next we describe the analytical steps that we will use to characterize vulnerability to water stress in meaningful spatial and temporal analyses.

Hydrological Modeling Characterizing Water Availability (Task 1)

The product of water availability analysis for the study areas will be a continuous time series of flow rates in the surface water and subsurface saturated zone storage beneath subwatersheds, where these catchments represent a range of contributing watershed areas. Hydrologic models will be used to extend the streamflow records from the limited number of gaged sites in each area to the ungaged study watersheds and to estimate subsurface water storage.

Modeling Approach.

Process-based, mechanistic hydrology models are available to quantify simplified forms of watershed processes governing the water balance. Such conceptualizations are generally depicted in a more detailed form of the water balance equation such as:

$$\boldsymbol{PPT + IMP - SF - DGL - EXP - CON - ET = \Delta S_{sw} + \Delta S_{snow} + \Delta S_{soil} + \Delta S_{gw}}$$

Equation 3

where P = precipitation; IMP = imported water in the catchment (e.g., inter-basin transfers and wastewater treatment discharges); SF = streamflow; DGL= deep groundwater losses (disconnected from the surface network); EXP = water exports out of the catchment; CON = consumptive water withdrawals including HF withdrawals; ET = evapotranspiration; and ΔS = change in storage which can be surface water reservoirs (ΔS_{sw}), snowpack (ΔS_{snow}), soil moisture (ΔS_{soil}), and/or groundwater storage (ΔS_{gw}).

Watershed (surface water) models have advanced capabilities to represent the surface water and soil water components of the hydrologic cycle. After accounting for water inputs and losses, these watershed models route water to streams as overland flow and interflow and recharge to the subsurface saturated zone that returns to the streams as baseflow (e.g. Figure 12). Widely used watershed models include HSPF (used extensively by EPA and USGS, among others) and SWAT (used extensively by USDA, among others). These watershed models have been shown to provide reliable simulations of rainfall-initiated stormflow and the transfers of water from subsurface to surface during rainless periods that produce very satisfactory estimations of streamflow records over time. The spatial and temporal data bases that inform these GIS-based models such as elevation, soils, landuse, rainfall, and streamflow are generally reliable and can be "ground-truthed" against observations within or nearby study areas.

There are also well tested and accepted groundwater models (e.g. GFLOW and MODFLOW) that have advanced capabilities to represent the saturated subsurface components of the hydrologic cycle. Groundwater models use a fundamentally different representation of subsurface characteristics to determine the storage volumes and saturated depths and to exfiltrate water to streams as baseflow than used by the watershed models (e.g., Figure 12). Due to the complexity of computations, watershed and groundwater hydrological models are generally not computationally linked, although there have been recent efforts to do so (e.g., INTBM, 2013; Guzman *et al.*, 2013). Groundwater models are more difficult to apply with known reliability given the lack of spatially distributed data on inherently heterogeneous subsurface aquifer and bedrock properties. Furthermore, well observation records required to calibrate and verify modeling simulations are usually very sparse.

One project objective is to represent time-dependent groundwater fluxes and storages in shallow aquifers beneath the study areas as an estimate of subsurface storage volume. We believe that the best estimates of groundwater storage would be derived from a groundwater model if one could be properly informed and validated, but we do not believe that is feasible for our project areas given the lack of necessary spatially distributed physical characterization and well observation data, as well as time and resource limitations. We therefore plan to use the recharge term from the watershed model to represent volume of water in subsurface storage.

To confirm that our surface-based watershed models represent baseflow-driving processes reasonably well, we will use a groundwater model to validate the watershed model in its predictions of groundwater storage. We will add a step to the model initialization and calibration process in each project area where we will utilize a cross-comparison between the watershed and groundwater models at steps where water flux is estimated by both models. The purpose of cross-referencing between the groundwater and watershed models is to develop confidence that the change in groundwater storage term from the watershed model can be used to characterize impact.

Application of Hydrologic Models in the Project Areas

Models. The project will apply two watershed models to assess water availability in the project areas--the Soil and Water Assessment Tool (SWAT) and/or the Hydrologic Simulation Program – Fortran (HSPF). Running two models will enable us to draw firmer conclusions from our scenario-modeling in these two watershed systems.

SWAT continues over 30 years of modeling efforts conducted by the USDA’s Agricultural Research Service and has been extensively peer-reviewed (Gassman *et al.*, 2007). SWAT hydrological response units can be parameterized based on publicly available GIS maps of land use, topography, and soils and is a good choice in less data-rich areas. SWAT model code has been tested under different environmental conditions nationally and internationally, as evidenced by more than 600 peer reviewed publications (Gassman *et al.*, 2007).

HSPF was developed nearly 20 years ago and is now jointly sponsored by the EPA and the USGS; it has extensive documentation and references (Donigian, 2000). HSPF has been applied to hundreds of watersheds throughout the United States and internationally. It has widespread acceptance by federal and state agencies and water districts. HSPF is included as a core watershed model in EPA’s BASINS modeling system (http://water.epa.gov/scitech/datait/models/basins/index.cfm) and the USACE’s Watershed Management System (http://chl.erdc.usace.army.mil/wms). It is listed as a “Nationally Accepted Hydrologic Model” by FEMA. (http://www.fema.gov/plan/prevent/fhm/en_hydro.shtm#2). As a subwatershed of the larger Chesapeake Bay system, our HSPF modeling in the SRB will allow benchmarking to the peer-reviewed and community-accepted Chesapeake Bay Program watershed model (Shenk *et al.,* 2012).

The project will select the most appropriate groundwater model to apply following assessment of available data to inform the model and evaluation of conditions in the project areas. Candidate groundwater models for benchmarking include GFLOW and MODFLOW. GFLOW (www.haitjema.com) is a regional groundwater models based on the analytic element solution method (www.analyticelements.org). This technique for regional groundwater modeling captures the accuracy of exact analytical solutions computationally (Strack, 1989; Haitjema, 1995a; Hunt, 2006; Kraemer, 2007). Solutions are based on the superposition of point sinks representing wells, line sinks representing rivers, and area elements representing heterogeneities in aquifer properties. In contrast, the US Geological Survey MODFLOW model (http://water.usgs.gov/nrp/gwsoftware/modflow.html) is a finite difference grid-based model, with particular strength in representing highly heterogeneous aquifer properties (assuming these are available) or tightly coupled watershed and groundwater systems.

GFLOW is the preferred model to represent groundwater storage because the vector basis of the analytic element method interfaces logically with that of SWAT and HSPF.

Model Initialization, Calibration and Cross-validation. Following is a general overview of how models will be parameterized, calibrated and cross-referenced. Application of this process in the project areas is described in detail in Appendix A.

Each of the watershed models and the selected groundwater models will be initialized as normally performed for each model. The watershed models require geospatial, meteorology and USGS streamflow records (see section B9). Once acquired, reviewed, and post-processed, each modeling team will initialize and parameterize the watershed models according to best professional judgment. Each study area has lengthy periods of streamflow records within and around the watersheds to calibrate the watershed models.

HSPF and SWAT will then be calibrated utilizing a Monte Carlo scheme to explore parameter combinations to optimize agreement between simulated and observed streamflow at all of the USGS gages in the project areas. The processes will be duplicated in both models and performed separately on the two study watersheds: we will use a multi-objective function, maximum likelihood calibration approach with a minimized number of parameters that range over physically realistic ranges. The multi-objective methods will allow us to customize our calibration target to the most important flow volumes for HF scenario assessment, while retaining a realistic water balance representation of watershed dynamics. Our maximum likelihood approach will produce an optimized or calibrated set of parameter ranges.

The groundwater model will be independently calibrated and validated, and will produce a time series of recharge that can be directly compared to SWAT and HSPF recharge outputs. The GFLOW groundwater model will first be calibrated to: (1) observed/estimated average annual baseflow in the streams, either observed at USGS gages or estimated based on regression at selected river points; (2) observed annual averaged shallow aquifer water levels in wells; and (3) water elevations in perennial streambeds as inferred from USGS topographic maps.

The length of stream defined for the watershed model influences estimates of baseflow volume, so it is important to properly define the upper extent of the stream network. Therefore, we will leverage capabilities of the calibrated groundwater model to define a perennial stream network for the watershed model. We will use the conjunctive groundwater and surface water analytic element modeling technique of Mitchell-Bruker and Haitjema (1996) to define the perennial stream network needed to support long-term average baseflow and transfer this network density to our surface modeling efforts.

We will characterize the quality of model simulations relative to observed using goodness of fit measures, including a weighted Nash Sutcliffe score. The calibration steps and criteria are discussed in detail in Appendix A.

Although calibrated independently, the watershed models and groundwater model will be compared using a stepped and iterative approach. In the model cross-referencing step, we plan to use the groundwater recharge, deep groundwater leakage, and groundwater storage terms from the watershed models to represent boundary conditions and initial conditions for the groundwater model. The groundwater model predicts contributions to baseflow in the streams and the predicted hydraulic head elevations will be checked for consistency with the watershed model. We anticipate that baseflow or subsurface-supplied streamflow will be particularly significant in impacts of HF water acquisition. Criteria will be established in calibration to determine whether sufficient accuracy is achieved by the watershed model recharge terms to represent subsurface water volume on a daily basis. We note that this cross-calibration approach may lower the overall best-fit score achieved for each model independently.

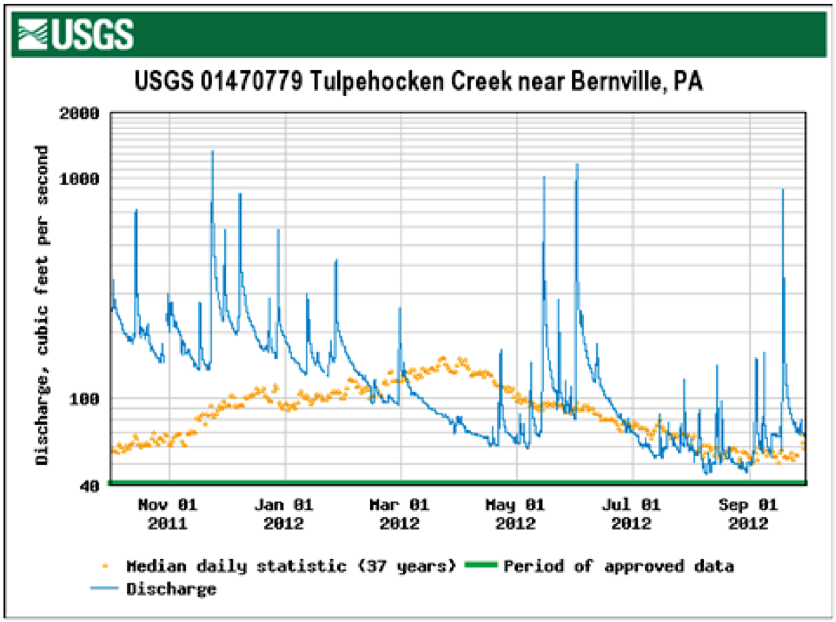
The viability of using SWAT or HSPF parameters from a gauged, calibrated watershed and applying them in an uncalibrated watershed will be examined using a cross-validation approach. This will estimate loss of model accuracy that results from using calibrated parameter values in other watersheds. For example, if we have 10 sets of maximum likelihood parameter calibration for 10 different (but hydrologically similar) watersheds, we also have 10 different estimates of a fit statistic. At each watershed, we can rerun the surface water model 9 times using the calibration parameter sets from the other watersheds and recalculate the fit statistic. This allows us to compare the fit score distribution for the calibrated watersheds to the fit score distribution when "transporting" calibrated parameters across watersheds. If the loss of prediction accuracy is acceptably small, this approach can be used to produce surface water flow estimates at watersheds where we do not have sufficient data for model calibrations.

We will also verify that streamflow and groundwater characteristics fall within published regional variability and that recharge estimates agree reasonably with an independent groundwater model

Model Application. Following iterative calibration, we will then run the watershed model for the 18 years of record producing streamflow and recharge on a daily time step for all selected subwatersheds in the project areas.

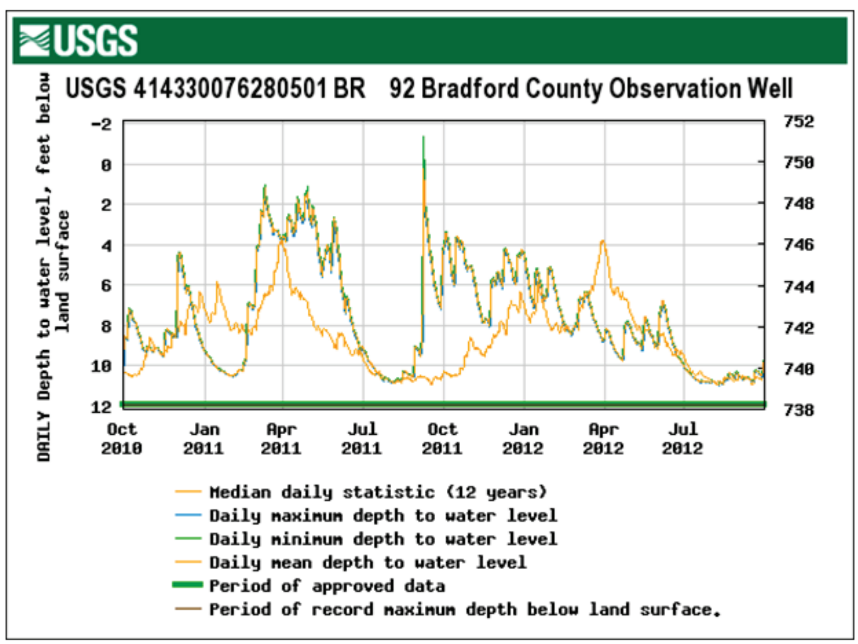
The output from spatially averaged watershed modeling passed to the next steps of analysis is the daily series of streamflow and groundwater storage. An example of the measured daily time series at two USGS monitored sites in Pennsylvania is shown in Figure 15. The streamflow rate will be translated to water volume, such as gallons/day, to facilitate comparison to withdrawals in later steps. The selected groundwater model will also project the saturated zone of the watershed that will translate to available water volume based on watershed area and aquifer base elevation.

Figure 15. Examples of daily flow record using measured daily series of stream flow and water table depth, illustrating data output from watershed and groundwater models passed to the next stage of analysis.



Surface Water

Groundwater



Spatially and Temporally Scaled Analysis

Basin area will be used to organize spatial scale and flow estimates in the project.

Subwatersheds spanning a range of contributing -- basin area will be selected within the larger project area. They will be established by placing pour points at change in stream orders (Strahler, 1957) (Figure 16). This step is accomplished during model initialization where each model uses a process similar to this to delineation watersheds. The groundwater model and field mapped stream networks will be used to help define the accumulation area for the perennial network.

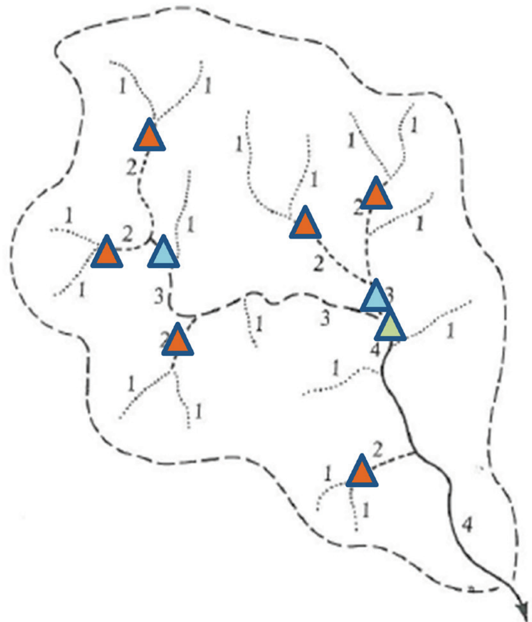


Figure 16. Watershed sketch with pour points established at change in stream order.

This step will create subwatersheds of varying sizes that are candidates for further analysis. The number of potential subwatersheds will decrease as stream order increases, as will the final subsample. A sufficient number of ungaged sites will be selected from the total population to achieve a sample that is as statistically relevant as possible within each stream order. We will use a stratified random sampling approach to build a set of 100+ watersheds with representative basin areas, land uses, and topography.

Temporally Based Flow Parameters. Daily flow characteristics at measured USGS streamflow gaging stations are extended to each of the ungaged watersheds using the spatially-distributed watershed models calibrated to the measured stream data. The watershed model will be run on each ungaged site to generate daily flow rate from each subwatershed using data from 1986 to 2012. Available surface water will be determined for the full range of stream size within project areas represented by the subwatershed sample population created in the previous step.

The ΔS_{GW} time series from the watershed model or the selected groundwater model for each subwatershed will similarly represent the subsurface storage volume on a daily time step for the period of record.

Once calibrated, the models will simulate the 26-year record for each subwatershed to ensure that statistical flow metrics derived from the data set represent a range of climatic conditions. Flow duration curves will be produced from the simulated record for each gaged and ungaged location from the modeled record (Figure 17). This allows results to be broadly extrapolated using common methods of regional hydrologic analysis (e.g. Stuckey, 2006). The streamflow statistics will be collected from the duration curves for all subwatersheds. For purposes of this project, the selected parameters will cover dominant flows, emphasize low flows, and de-emphasize peak (storm) flows in both calibration and application.

Parameters include:

- Q70, Q60, Q40, Q30, Q20
- Annual mean (Q50),
- Annual baseflow
- 30-day low flow, 2-yr, 10 yr
- 7-day low flow, 2-yr, 10 yr
- Annual minimum flow
- Average monthly flow

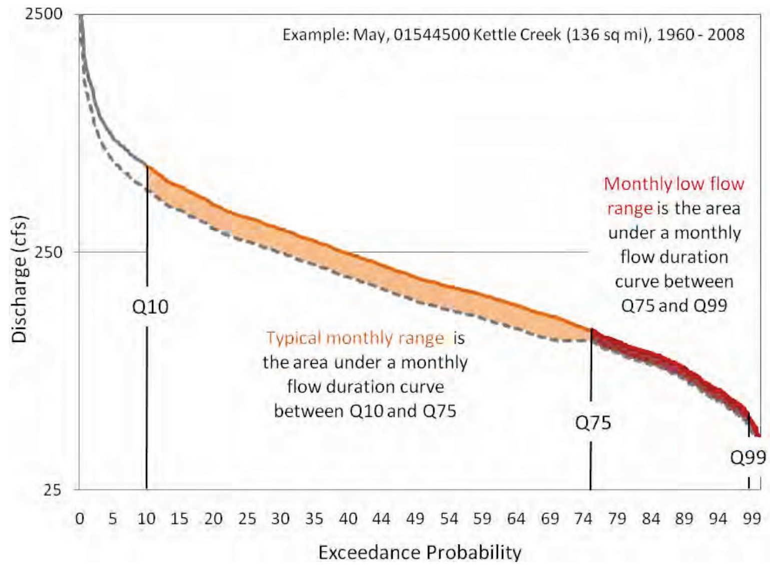


Figure 17. Example flow duration curve that will be calculated from the modeled daily time series steps.

Each of these metrics is a flow volume observed frequently or rarely, though not necessarily sequentially. In the western U.S. there is a strong seasonal pattern in flow volume compared to the humid east where storms and groundwater recharge occur year-round.

The modeled flow duration statistics will be compared and cross-checked with empirical regional discharge relationships computed from USGS stream gaging sites (e.g., Stuckey, 2006). Both methods produce a basin area-to-discharge relationship. Presumably the mechanistic modeled flows will have less variability than in the regional relationships due to the greater spatial resolution of land use, slope, and precipitation that are included as explanatory variables in the regional relationships.

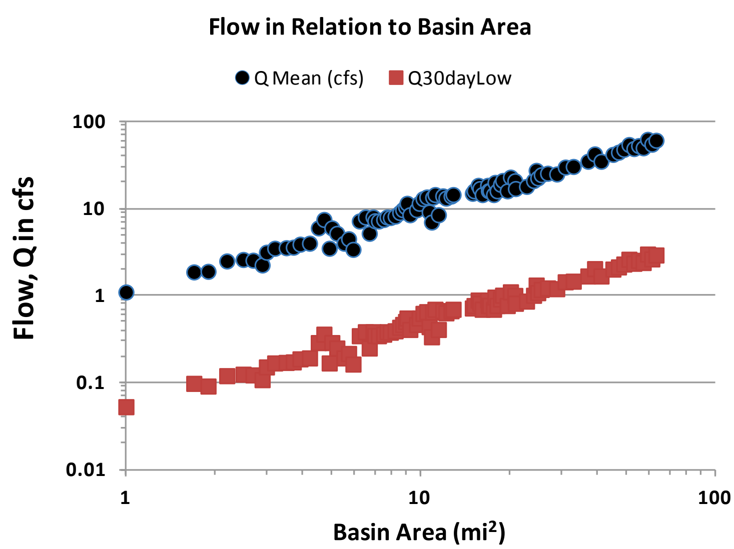


Figure 18. Example relationship between discharge and basin area for two flow-duration statistics computed for Pennsylvania streams from a regional statistical analysis (Stuckey, 2006).

A relationship between each flow statistic determined from the modeled flow record and basin area will represent the expected baseflow for current water use conditions (Figure 18). Similar duration curves will be calculated from the daily time step of subsurface storage volumes produced by the groundwater model. The flow duration statistics from the modeled results represent water availability metrics in Scenario Analysis (Task 3).

Water Consumptive Use Analysis (Task 2)

Consumptive water use must be characterized to complete the supply/demand comparison. Although EPA's HF scenario analysis primarily targets potential effects of hydraulic fracturing on domestic supplies, the combined water consumption by all users must be balanced with combined water availability from all sources in each analysis area. The ability to obtain water for HF from sources other than drinking water sources can reduce pressure on domestic supplies.

Water use from public and self-supplied sources for both surface and groundwater sources will be determined by accessing available information from state and local regulatory agencies, the USGS, and operators. To support the study design, consumptive use must be known or estimated at the same range of spatial scales addressed by the water supply analysis (1st to nth order streams). Similarly, consumptive use must be known or estimated for the fine-tuned temporal resolution in the flow duration periods of analysis. Low flow periods in particular tend to be temporally explicit and likely to coincide with heavier consumption.

The project will assess water consumption at the subwatersheds established in the previous step, as well as at known withdrawal points for domestic water supply (Figure 19). Daily consumptive water use will be compiled for all categories (aquaculture, irrigation, livestock, industrial, self-supplied water supply, mining including HF, and public-supplied water) within the watersheds at selected pour points, at USGS gage calibration sites, and for any public domestic water supplies that occur in the project area are subject to inter-basin transfers or wastewater treatment discharges. Net water withdrawals, expressed as differences between total withdrawals and total discharges, will be estimated. The latest registered and unregistered (estimated) water withdrawals at the daily time scale at each pour point will be used.

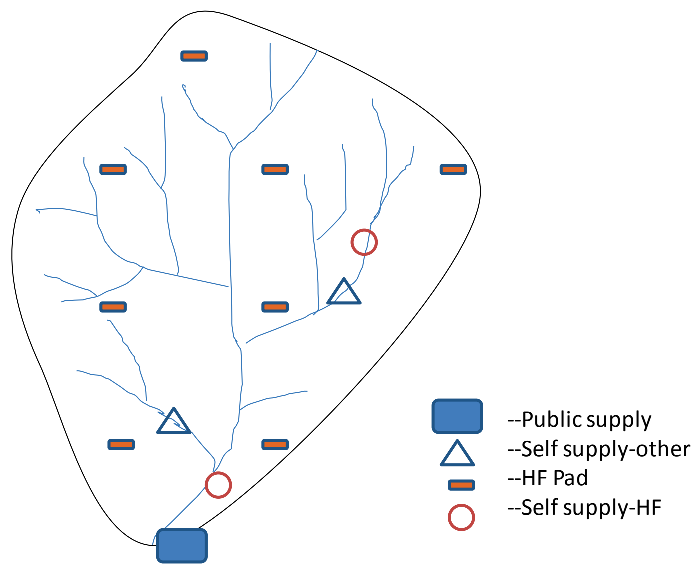


Figure 19. Schematic of water withdrawal and hydraulic fracturing pads within subwatersheds.

Consumptive use data is available from federal, state, and private sources. The USGS performs a detailed water use census every five years and reports at national, state, and county levels. Underpinning data in the census may provide more spatially-explicit data to the EPA project that can enable us to ascribe use at the subbasin scale more accurately. Water use data on all or specific water uses may also be available from state and local authorities and from private industry, including the hydraulic fracturing industry. Much data of this type was obtained in the study areas in Phase 1, although not necessarily at the temporal and spatial scales needed for Phase 2 analysis. Detailed information on gas pad locations and information on withdrawal locations, quantity and timing are key requirements. Known data sources in each study area are provided in Appendix A.

When sufficiently detailed data are not available, we will estimate water use at the subwatershed level, using statistical techniques for downscaling or upscaling to contextualize current uses and potential hydraulic fracturing water withdrawals. To produce estimates of sub-basin water use, these techniques may use spatial weighting based on degree of intersection between county and basin coverages, regression methods that utilize available covariates such as population and land cover characteristics, and/or covariance-based spatial estimation approaches. Such meta-modeling has been applied in climate change assessments (Wimmer *et al*., 2011). Temporal-scaling of water use may also be required.

Ecosystem sustainability in surface waters requires maintenance flows that may also be called "passby" flows. Many states set and enforce flow requirements, although techniques, terminology and/or metrics vary. We will review flow criteria in the study areas and apply them in calculating water use. Should there be no established passby flows, we will review the scientific literature to define and apply an ecosystem flow requirement in calculations.

Consumptive uses from surface water will be summed within each subbasin established for the surface water analysis. Throughout, we will evaluate surface water and groundwater availability and use separately.

SWC = Irrigation + HF + Public Domestic + Industrial + Thermoelectric + Hydraulic Fracturing + ECO **Equation 4**

where SWC is surface water consumption, and ECO are flows designated to protect ecological systems.

GWC = Self_supplied Domestic + Public Domestic + Self supplied Irrigation + Self_supplied HF _ + Self_supplied Industry + Self_supplied Energy **Equation 5**

where GWC is groundwater consumption.

Water withdrawals can be represented in the watershed models, or they can be treated using a "lumped" approach where volumes for all users above the watershed pour point are simply summed and averaged over time. The project expects to use the lumped approach, but during analysis will explore whether there is important loss of resolution that arises from spatial and temporal effects of routing and storage that the models may be able to account for.

As suggested by the overarching hypothesis illustrated in Figure 14, we theorize that withdrawals are more likely to create greater water stress on surface water supplies during low flow periods when irrigation use is higher and water is extracted from smaller streams.

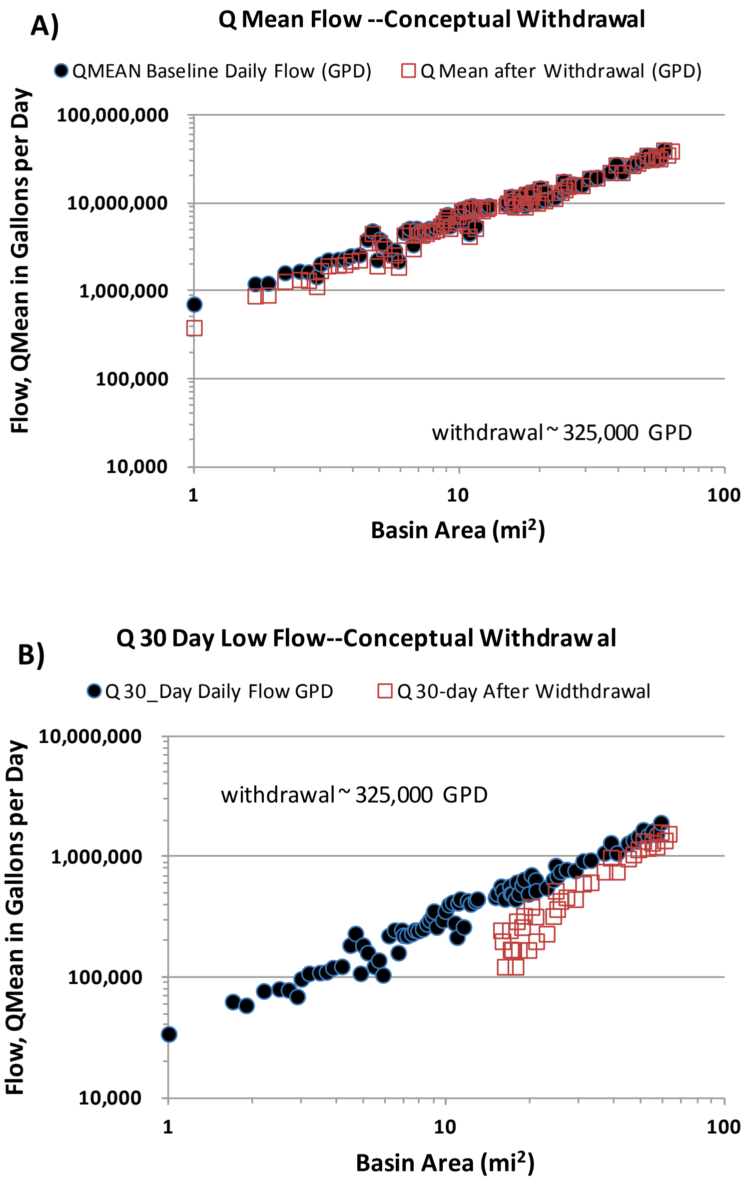


Figure 20. Example calculation of withdrawal of 325,000 GPD on the annual mean flow and 30-day low flow of the Pennsylvania data set generated by the regional regression of discharge in relation to basin area (Stuckey, 2006).

We illustrate this with surface water flow estimates for annual mean flow and 30-day low flow for the Pennsylvania data set computed from regional regressions shown in Figure 18. In this case, we translated instantaneous discharge rate (cfs) to total daily flow in gallons per day (gpd). We apply an arbitrary withdrawal rate of 325,000 gpd to each subwatershed. This volume would provide one well with 2,300,000 gallons of water in a 5-7 day period. There was only a neglible effect on available water at any watershed size where this volume of water was withdrawn from a flow volume equal to the annual mean (Figure 20A). There would be observable and potentially significant effects in the smaller watersheds when the same volume was withdrawn when flow was equal to the 30-day low flow, as evidenced by observable departure from flow levels (Figure 20B). There would be insufficient flow to support any withdrawal in watersheds less than about 10 mi^2.

In the project areas, individual subbasins are likely to have a mix of existing water uses and therefore varying levels of withdrawals from surface and groundwater sources. Actual water use will be key information for the analysis of relative vulnerability to HF withdrawal.

Recognizing there may simply not be enough water, and assuming that low flow withdrawal volumes are limited by regulation, the practical outcome of the example is that potential water supply locations that can serve HF operations may not always be available. There is likely to be more pressure on public supplies from HF during low flow periods, thus diverting to supply systems with adequate storage capacities and/or groundwater sources may be necessary.

Water consumption estimates within the subwatersheds will be constructed from available public records, although there will be no way to verify actual values or impact. The "reasonableness" of consumption estimates will be assessed during calibration steps where current flow records include contemporary uses. The project team will evaluate consumption estimates for reasonableness, but no performance expectations are specified for this step.

Water Stress Indices

The water consumed relative to water availability will be computed separately for surface water and groundwater using a simple comparison index that shows the balance between water consumed and water available, such as that used by Tidwell (2013) in Figure 10. There are several ways such an index can be expressed: the surface water (SWI) and groundwater (GWI) indices can be calculated as:

$$\mathbf{SWI} = \frac{\sum \boldsymbol{SW\ Consumptive\ Use\ Volume}}{\boldsymbol{Available\ Surface\ Water\ Volume}}$$

Equation 6. Surface water index

$$\mathbf{GWI} = \frac{\sum \boldsymbol{GW\ Consumptive\ Use\ Volume}}{\boldsymbol{Available\ Groundwater\ Volume}}$$

Equation 7. Groundwater index

Units will be in gallons/day (or metric equivalent).

Figure 21 displays the example effects analysis for the Pennsylvania streams generated by regional regression equations to compute at Surface Water Stress Index. The available water is as shown in Figure 18 and consumption is 325,000 gallons per day. Figure 21A shows the SWI for the large basins (100 of 1,000 mi^2). The large basins experience no water stress at this withdrawal rate at any level of flow (low SWI). Figure 21B focuses on the size of watersheds to be analyzed in this project. The 200 mi^2 basin shows no stress at the average annual flow, but exhibits low level stress at the lowest annual flows. The smaller watersheds show significant potential for water stress increasing as flows decrease..

Figure 21 demonstrates quantification of the conceptual illustration in Figure 14.

Hydraulic Fracturing Scenario Analysis (Task 3)

The objective of Project 5B is to assess the contribution of hydraulic fracturing activities to potential impacts on domestic water supplies. To accomplish this, we will compare the volume of water consumed by HF plus other use demands relative to the volume of available water within the subwatersheds. Water use for HF will compare water stress indices at a base level of consumptive use prior to HF with three scenarios that reflect the current level of drilling, increased well field development rates projected in the future, and recycling technologies.

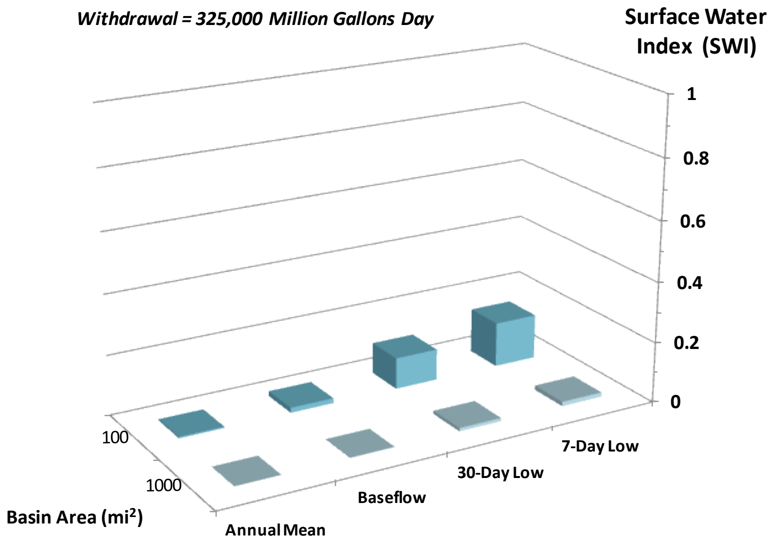
Hydraulic Fracturing Water Consumption. HF future scenario modeling will evaluate the impact of water acquisition on available water supply under various assumptions of intensity of well development activity and ability to augment water withdrawn from surface and groundwater with recycled water within the project areas. The factors determining the total volume of water consumed with hydraulic fracturing will include: 1) the volume of water used per well, and 2) the aggregate number of wells drilled within each area.

HF Volume = # Wells x Volume/well **Equation 8**

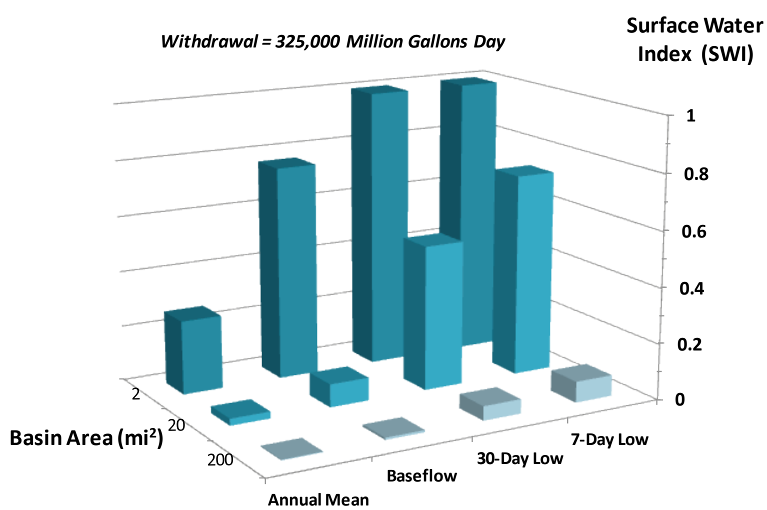
The scenario analysis will evaluate the HF acquisition at two points in time: 1) current rate to represent the current baseline scenario (consumptive use from Task 2), and 2) the future peak year of drilling to represent the maximum impact of HF, termed the"energy plus" scenario. The HF volume will be computed for each subwatershed and specific location point, on an annual or finer timescale if possible.

Figure 21. Surface water index indicating balance between consumption and available surface water based on PA regional hydrology equations and example water withdrawal rate . The illustration shows large watershed relative stress (A) and stress at spatial scales to be emphasized in the Phase 2 refined analysis (B).

A.)



B.)



Drilling Rates. Current drilling rates are tracked by various agencies. Future drilling rates were estimated in the Phase 1 project areas using National Energy Modeling System (NEMS) projections of unconventional gas drilling in the US (Cadmus Group, 2012a,b). NEMS is an energy-economy model designed by the Energy Information Administration of the US Department of Energy to project future US energy production and demand for Annual Energy Outlook reports (USEIA, 2012). NEMS projects annual well starts at the play-level based on play geology and market conditions. Results are aggregated by unconventional gas type (shale gas, tight gas, or coalbed methane) to six regions of the continental US for output. The regional scale analyses were downscaled to the project areas using basin-to-region scaling factors. Future hydraulic fracturing activity estimates were cross-checked with USGS assessments of maximum potential well densities in undiscovered gas resources in the Susquehanna River Basin (USGS Marcellus Shale Assessment Team, 2011) and in the Upper Colorado Basin project areas (USGS Uinta-Piceance Assessment Team, 2003).

Water Use Per Well. There are several sources of available information reporting volume of water used per well, including the industry sponsored data base FracFocus (GWPC and IOGCC, 2013) and various regulatory oversight authorities and land or water management agencies such as the SBRC in Pennsylvania and the BLM in Colorado. The total water used for HF per well is:

HF well volume = "New" water + chemical/proppant volume + recycled water

Equation 8

Key HF characteristics determined in the Phase 1 assessments are provided in Table 4 for the project areas as determined by Cadmus (2012a,b). The total volume of water, proportion of chemical/proppant and use of recycled water varies significantly among plays.

Table 4. HF water use and drilling rate statistics determine for from the Cadmus Phase 1 assessments in the large basin project area (Cadmus Group, 2012a,b)

	Characteristic	**Susquehanna River Basin**	**Upper Colorado River Basin**
Water Use Per Well	Total volume/well (gals)	4.04 million	0.25 million
	Actual new water volume/well (gals)	3.5 million	0.18 million[&]
	Average recycling %	13%	100%
	Green Technology Volume (gals)	2.9 million	0.18 million
Drilling Rate Per Year	Business as usual (current rate)	2370	1884
	Energy plus (maximum yearly rate)	2840	2108
	"Recycling Plus"	2370	2108
Well density	Drainage area per well, acres: Ave (High-end)	149 (80)	80 (20)

[&] water used for drilling, dust abatement hydrostatic testing

The Phase 2 project will utilize similar procedures as used for Phase 1 to determine drilling rates and water use for use in scenario analysis. The Phase 2 project will reassess and check numbers used in Phase 1 and will use downscaling techniques to adjust to the selected subwatersheds within the larger project basins.

The future scenarios used in Phase 1 also consider a potential increase in surface withdrawals for public water supplies due to population growth in the basin. Population growth will be also considered in the consumption applied to future scenarios in the Phase 2 project.

Water Management. Water allocation and management is an important consideration in many areas where hydraulic fracturing is conducted. We use the term "water management" to refer collectively to regulatory and industry practices that determine HF well operations and scheduling that affect when and from where water is sourced.

Water management occurs under the authority of local or state agencies or intra-basin commissions that have top-level responsibility for permitting and regulation. Permitting and regulatory requirements vary regionally and by state, but virtually all potential HF well fields and their potential surface or groundwater sources are covered by a water management authority. These authorities typically permit, track, allocate and, potentially limit water withdrawals. They also may enforce minimum or passby flows in rivers and regulate timing and location of withdrawals.

Operators may proactively apply adaptive strategies to minimize stress on surface and ground water resources in their communities (API, 2010). Individual companies may apply water management strategies for individual wells and their regional areas to minimize costs while providing the necessary flow of water to continue operations. They may reduce water needs with improved technology, control HF scheduling, or distribute acquisition from a mix of sources, to name a few.

It is important for the EPA to fully understand aspects of water management associated with HF as they strongly influence the potential for HF operations to increase or minimize regional or local water stress. Building water management practices into scenario assessment is important but not straight forward. HF operations are dynamic in time, spatially and jurisdictionally variable, and unique to operators.

At this point, the project team is not sufficiently informed on details of water management within the project areas to develop appropriate representation of HF activity for the impact analysis. The project will therefore seek additional information on factors listed in Table 5 within the study areas via information requests to agencies, operators and stakeholders to develop robust management scenarios. Other EPA water acquisition studies are also gathering information as a literature review and for assessment of wells, service companies, and so on. Our project will consult with the companion projects for supplemental information and will coordinate with them in any mutual data acquisition planned for the project areas. Data will be synthesized and built into a set of "decision rules" to define how data and assumptions will be applied for the HF scenario analysis. It is not feasible to define scenarios that will be universally transferable, but we will strive to develop realistic scenarios that incorporate appropriate water management strategies that are informative within and beyond the project areas.

The detailed water management assumptions will be documented once management information is obtained according to the project schedule. Water management details that ultimately are built into the scenarios will be shared with stakeholders for review and input prior to moving on to HF scenario analysis.

Table 5. Elements of water management to be considered in scenario development.

Activity Category	Water Management
Individual Wells	• Continual improvement of technologies that reduce water use • Well volumes • Scheduling sources and well activity • On- site water storage • Off-site water storage
Regional Operations	• Pattern of well development (concentrated, dispersed?) • Factors contributing to development pattern, including thickness of gas resource, proximity to oil and gas pipelines, leasing rights • Number of wells fractured each year • Timing of activities
Water Suppliers	• Availability of public and self-supplied sources • Large-scale storage capacity within public systems • Scheduling • Reuse of flow back and produced water • Substitution and Priorities
Other Users	• Allocations • Public domestic water suppliers • Self-supplied domestic water supplies • Agricultural • Industrial
Regulatory Controls	• Permitting • Closure of sources to maintain minimum ecologic flows

Defining HF Management Scenarios. The scenario analysis will assess the effects of HF water withdrawals within the context of overall water use and availability for four defined scenarios that reflect current and projected future drilling rates and technologies. Scenario analysis will consist of four scenarios reflecting local water use prior to HF activity, current HF use, and future peak use with and without improved technology such as recycling of hydraulic fracturing fluid. The general attributes of the scenarios provided in Table 6 are similar to those developed for Phase 1. All scenarios will utilize existing withdrawals from all users, managed with current water regulatory and best management practices determined during fact-finding trips and information requests.

Baseline Scenarios. The baseline scenario will represent the USGS's major water use categories including the consumptive component of public water supplies, domestic water use, and other major water use categories (irrigation, livestock, industrial, mining, thermoelectric power). The baseline model will assess major consumptive water uses for watershed conditions of the year 2005 for the project areas corresponding with the USGS' water use reports (every five years since 1950) and the National Land Cover Dataset (USGS, 2011). If available, 2010 data will be used. The baseline year predates the significant expansion of hydraulic fracturing in these project areas. Flow values for this scenario for each subwatershed and flow statistic are delivered from the final watershed model run after calibration is completed, along with a sum of consumption volume. Both are entered into an analysis data set as illustrated in Figure 22. The baseline flow volume remains the same through all subsequent scenarios—just the consumption term is changed to reflect the change of assumptions in water management, HF operations, and recycling.

Table 6. Overview of factors considered and definitions that will be included in Scenarios for Impact Analysis. Details will be specified following additional fact-finding.

Scenario	Water Users	HF Source Water Management	Hydraulic Fracturing Well Density and Water Use	Recycling/Reuse produced Water
Historical Baseline	USGS uses + ecologic maintenance	None	None	None
Current Baseline Business as usual	USGS uses + ecologic maintenance	Current Practices (Regulatory + Operator)	Current well numbers (2010) and densities	Current rates
Future Scenarios "Energy Plus"	USGS uses + ecologic maintenance	Current Practices (Regulatory + Operator)	High-end estimate of futured play-level development projections and current total well volume	Current rates
"Recycling Plus"	USGS uses + ecologic maintenance + low-end population growth	Current Practices (Regulatory + Operator)	Median estimate of futured play-level development projections and current total well volume	Recycling volume based on high end rates

HF Scenarios. Three HF scenarios will be applied with results compared to the baseline. The HF scenarios will simulate various rates of drilling, including current and a projected maximum annual rate within a 30-yr planning horizon, determined based on recent drilling trends and projections of natural gas production (USEIA, 2012). HF scenarios will include consideration of "water management" and BMP's that also influence water use. One HF scenario will specifically assess the benefits of water recycling by the industry. Table 6 indicates elements that will included in the scenario definitions. Each scenario assumes distinct levels of natural gas drilling and HF freshwater use and, therefore, will apply distinct hydraulic fracturing water withdrawal time series to modeled stream reaches. Further, significant population growth is projected over the next 30 years in Garfield/Mesa Counties, Colorado (CWCB, 2004) where natural gas extraction within the UCRB has been concentrated in recent years. UCRB future scenarios will therefore consider potential increase in PWS surface withdrawals in the basin.

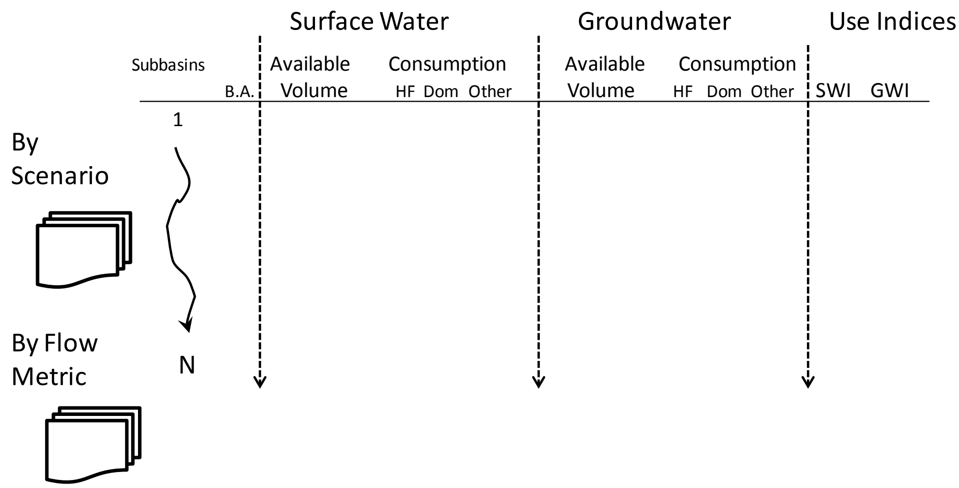


Figure 22. Data elements expected to be included in final scenario analysis.

HF Scenario Analysis.

A single calibrated validated watershed model (based on historical meteorology and streamflow records) provides the water available for all scenarios as developed in Task 1. Available water versus consumption for each flow duration metric and scenario will be computed for surface and groundwater individually.

Existing consumptive use for users other than HF is determined in Task 2. HF water consumption is determined for each future scenario in Task 3. HF assumptions will be determined specific to the project area based on fact-finding on key water management and HF practices (Table 5) and defined in scenarios (Table 6).

The watershed model, consumption analysis and HF scenarios will produce a series of spatial representations of surface water index (SWI) and groundwater index (GWI) which will portray the relative vulnerabilities to overconsumption of subwatersheds of the SBR and the UCRB. We have identified a suite of flow metrics to characterize low and median flows over various timescales. We will generate values of SWI and GWI for each subwatershed, representing each flow metric across the four scenarios of HF consumption.

Each scenario will influence the consumptive use rate and will be reflected in the SWI and GWI computed. The SWI and GWI for the three scenarios that include hydraulic fracturing consumption will be compared to the historic baseline which is the main point of reference. This will be accomplished in a a spreadsheet compilation.

In our models, all other water balance components will remain constant across the scenarios, such that differences in SWI and GWI for a given flow metric will be due solely to differences in HF withdrawals. SWI and GWI will be calculated as the ratio of consumption to availability (Blackmore and Plant, 2008; Tidwell, 2013). Our modeling approach is based on simulation sets and will give us distributions of SWI and GWI to explore variability within the scenarios. Our study design dictates that increasing HF consumption will be associated with greater water stress (higher SWI and GWI). Thus, the focus of our analyses will be on representing vulnerability across spatial and temporal scales, to assess the spatial and temporal dynamics of overconsumption risk due to HF withdrawals.

We will first conduct a multinomial chi-square test on categorized SWI and GWI to determine if differences across the scenarios (for a given flow metric) are statistically significant, by comparing the scenarios to the baseline categorical data (Ott and Longnecker, 2001). Following Tidwell (2012), we will define "high risk" as GWI and SWI values ≥ 0.7. We will then compare the proportion of high risk sub-basins as HF consumption varies across the four scenarios, and we will graphically portray the areas of Towanda Creek and UCRB that are most vulnerable to water stress. Using the suite of flow metrics, which range from extreme lows to median flows, and the range of sub-basin sizes, we can create a three-dimensional plot showing spatial and temporal variability of water stress, for each of the four scenarios. These visualizations will provide guidance on the vulnerability of different parts of the watershed system as a function of upstream drainage area and flow magnitude. Replicating these water stress plots for each scenario will allow us to assess the relative impact of varied levels of HF withdrawal that considers basic scaling dynamics.

We will also report general statistical distributions of these findings and statistically compare hydraulic fracturing scenarios to baseline scenarios. Statistics will be computed with SAS or R.

A6 Project/Task Description and Schedule

The project will assess potential impacts of water acquisition for hydraulic fracturing in case study areas located in western Colorado and the Susquehanna River Basin in Pennsylvania. The project analysis plan described in Section A5 will be implemented in four tasks, three of which address various aspects of the analysis process finalized in report preparation outlined in Task 4. Tasks 1 through 3 will be implemented concurrently. Task activities are summarized below.

Task1. Water Availability Modeling

The product of Task 1 is an estimate of water availability characterized by flow duration statistics computed for general analysis of randomly selected subwatersheds with a range of basin size and at "location-specific" sites centered at public water supplies.

Estimates of water availability are produced using analytical watershed and groundwater models. For Task 1 the project team will acquire, review and prepare spatially-referenced and time series flow and meteorology data necessary to run selected hydrologic models. The GIS specialist will prepare the acquired spatial data necessary for each project area and project scientists will review and prepare the time series data.

Two watershed models will be applied in each of the project areas using standard protocols. One team experienced with the SWAT model will apply it in each project area, and one team experienced with the HSPF model will apply it in each project area. The groundwater specialist will apply the groundwater model in each project area.

Each modeling team will initialize their model. Current consumption estimates needed as part of the calibration parameter set will be acquired from Task 2. Each model will be calibrated at specified USGS stations in cross-reference with the groundwater model (see Section B7) under the leadership of the "analytics team" composed of statistical and spatial analysts assigned to the project. Best parameterization will be selected based on calibration. Steps and outcomes will be documented in a Calibration Report. Additional details on parameterizing and calibrating the models in the project areas is provided in Appendix A.

Subwatersheds to be used for further analysis will be randomly selected from candidate sites determined during the watershed delineation step of model initialization to represent a range of basin size and land use within each project area for "general" analysis. Public water supply sources will also be identified for "location" analysis.

The calibrated model will be run with "current consumption" minus any HF influence to represent the "Historical Base Case" (HBC) scenario to produce daily time series of streamflow (Q) and subsurface water table depth for the modeling period of record for each subwatershed and location point. This model run will be archived.

Output from the HBC scenario model run is a text-delimited file of the daily time series of discharge and subsurface water depth. The model file is post-processed using SAS or R to produce monthly mean statistics and various cumulative distribution statistics as described in Section A5 for the period of record such as mean baseflow, 7-day low flow, etc., for each subwatershed and location point. All of these data are the “statistical” data and represents “available water” in all subsequent scenario comparisons. The statistical data are kept in the “analytics file” stored in a Microsoft®Excel file and/or in a batch readable text file.

Task 1 is complete when the analytics file for both project areas is created and populated with the synthesized surface water and groundwater volumes.

Task 2. Water Consumption

The product of Task 2 is an estimate of the volume of water consumed from surface water and groundwater storage by all users, other than hydraulic fracturing operations, within each subwatershed and expressed for each period defined in the flow duration statistics (e.g., 7-day low flow, mean annual flow).

Within Task 2 the project team will acquire, review and prepare consumption data obtained from external sources. Data on water consumption by all sources will be sought from the U.S. Geological Survey, water utility districts, and other authorities such as the Susquehanna River Basin Commission and Colorado Oil and Gas Conservation Commission. Data on HF consumption will also be sought from public and industry sources. The project team will review the data for quality assurance. To the extent possible, the project team will also use guidance and information from the Phase 1 study.

The project team will assess the data and determine how to estimate consumption within watersheds in the general analysis and at supply location points. If data are spatially referenced with good location estimates, Task 2 can proceed. If data must be downscaled or upscaled, the analytics team will develop a statistical procedure to accomplish this task. The consumptive withdrawals are parsed to the subwatersheds according to land use and demographics supplemented by available information from known withdrawal points.

The water volumes needed for maintenance of ecologic services are estimated based on existing regulatory procedures for each project area. If ecologic flows are not specified, the project team will review the scientific literature and develop a method to determine this value.

Consumptive use volumes are summed for each subwatershed and location point for each flow level representing SWC and GWC described in A5 (Equations 3 and 4). There is no hydrology modeling conducted to complete this task. These values are added to the analytics file and represents current baseline consumptive use. Baseline consumption represents “current conditions” and individually tracks domestic supplies as part of the all users.

Task 2 is complete when consumption values are added to the analytics file.

Task 3. HF Scenario Analysis

The product of Task 3 is the Scenario Analysis comparing effects of HF scenarios on water stress indices (see section A5). The water stress indices (Section A5: equations 5 and 6) are calculated as the ratio of water consumption determined in Task 2 to water availability determined in Task 1.

Task 3 will require the project team to define management details of the scenarios that will be used to characterize HF operations including water management practices (regulatory and industry practices), regional and individual well development (i.e., volume of water use per pad, number of pads, and pad distribution within the project area). These elements will inform the “future energy plus” scenario. The project team will also determine reasonable estimates of use of recycled waters for the “recycling plus” scenario evaluation.

The project team will conduct a fact-finding step seeking information on all of these factors within the project areas utilizing information generated in Phase 1 by the extramural contractors to the extent possible. This will entail querying regulatory authorities, the HF industry, and others. The project team will synthesize the information into “decision rules” or assumptions used to frame details of the HF scenarios. These will address spatial location of wells and supplies and temporal periods of withdrawal. The rules will be written and offered for review and comment prior to completing the scenario analysis.

The volume of water required for HF operations in the future scenarios will be determined for each subwatershed and location point. These values will be a summation of estimated withdrawal, as in the consumption step. There is no hydrologic modeling included in this step. These values will be input to the analytics file to be added to existing water user consumption.

The project team will analyze differences in water stress indices among the four scenarios statistically, graphically, and empirically with guidance from the analytics team.

Task 3 is complete when the scenario analysis is complete.

Task 4. Reports

The project will deliver peer-reviewed findings to the EPA Hydraulic Fracturing Synthesis Report, scheduled to be submitted in draft form to the EPA Science Advisory Board in December 2014. The EPA National Center for Environmental Assessment (NCEA) is assembling the synthesis report which will include a dedicated chapter to the question of potential impacts of large volume water withdrawals of groundwater and surface waters on drinking water resources.

Expected Products from Project 5B Phase 2

1. Report, “Modeling the Impact of Hydraulic Fracturing on Drinking Water Resources Based on Water Availability Scenarios in Case Study Catchments in Garfield County, Colorado and Bradford County, Pennsylvania (May 2014)

2. Journal manuscript, “Modeling the Potential Impact of Water Acquisition Activities of Hydraulic Fracturing for Unconventional Natural Gas on Drinking Water Resources” (December 2014)

Project Team Milestones

At minimum, this project will deliver products according to the schedule described below.

Major project team and product milestones are provided in Table 7. Intermediate steps and milestones are detailed in Table 8.

Table 7. Project 5B Phase 2 refined analysis major milestones.

Milestone	What	Date
Project Team Milestones	**QAPP completed**	**Aug 30, 2013**
	Water Availability Modeling Complete (Task 1)	**Dec 12, 2013**
	Consumption Analysis Complete (Task 2)	**Dec 12, 2013**
	Scenario Analysis Complete (Task 3)	**Jan 17, 2014**
	Internally reviewed draft journal article complete	**Mar 14, 2014**
	Submit manuscript to journal	**Apr 10, 2014**
	Internally reviewed draft EPA report	**Jun 1, 2014**
Major Milestones: Final Products following peer review	**QAPP Approved**	**Aug 30, 2013**
	Approved EPA Project Report Complete:	**Aug 29, 2014**
	Journal article accepted:	**Dec 1, 2014**
	QA Documentation Complete	**Dec 15, 2014**

Table 8. Project 5B Tasks and Schedule.

Project 5B. Refined Water Acquistion HF Scenarios	8/9/2013	8/16/2013	8/23/2013	8/30/2013	9/6/2013	9/13/2013	9/20/2013	9/27/2013	10/4/2013	10/11/2013	10/18/2013	10/25/2013	11/1/2013	11/8/2013	11/15/2013	11/22/2013	11/29/2013	12/6/2013	12/13/2013	12/20/2013	12/27/2013	1/3/2014	1/10/2014	1/17/2014	1/24/2014	1/31/2014
Quality Assurance Project Plan																										
Complete Draft QAPP		■																								
EPA Review			■																							
Final QAPP				■																						
Task 1--Water Availability																										
Acquire spatial data				■	■	■																				
Acquire time series data (Q, weather)				■	■	■																				
Intialize Water model						■	■	■																		
Major Calibration Process (B7)								■	■	■	■	■														
Identify Subwatersheds						■	■	■																		
Init. Model Run--historical base case													■	■	■											
Post process model output															■	■	■	■	■							
Compute flowdur statistics															■	■	■	■	■							
Process data to analytics file															■	■	■	■	■							
Archive acquired data and model output																				■	■	■	■			
Task 2--Water Consumption																										
Acquire water use, census, parcel data			■	■	■	■	■																			
Review data and quality assure			■	■	■	■	■																			
Determine methods for extending to subwat							■	■	■	■	■	■														
Determine method for consumption calc							■	■	■	■	■	■														
Decision: run model for scenarios or done													■													
Process data to analytics file														■	■	■	■	■								
Task 3--HF Scenario Analysis																										
Conduc watershed management fact finding				■	■	■	■	■	■	■																
Develop HF projects from NEMS				■	■	■	■	■	■	■																
Define Scenario Details											■	■	■													
Review as needed														■	■	■	■									
Compute CSI indices for Scenarios																		■	■							
Conduct statistical analysis																				■	■	■	■	■		
QUALITY ASSURANCE	ongoing																									

Table 8. Tasks and Schedule, continued

Project 5B. Refined Water Acquistion HF Scenarios	1/31/2014	2/14/2014	2/28/2014	3/14/2014	3/28/2014	4/11/2014	4/25/2014	5/9/2013	5/23/2014	6/6/2014	6/20/2014	7/4/2014	7/18/2014	8/1/2014	8/15/2014	8/29/2014	9/12/2014	9/26/2014	10/10/2014	10/24/2014	11/7/2014	11/21/2014	11/28/2014	12/12/2014	12/26/2014	1/9/2015
EPA Reports																										
EPA Project Report																										
Draft report complete							■																			
Initiate internal review EPA and revisions							■	■	■																	
Initiate clearance NERL EPA and revisions										■	■															
Initiate external review IOAA EPA and revisions												■	■	■	■											
Revised Report to IOAA																■										
Journal Article																										
Draft journal manuscript			■																							
Initiate internal review man. and revisions				■																						
Initiate clearance NERL journal man. & revisions					■																					
Submit manuscript to journal						■																				
Revised manuscript to IOAA																	■									
Revised manuscript to journal for acceptance																			■							
Quality Assurance																										
QA documentation																			■	■	■	■	■	■		
Review by ORD Quality Team																			■	■	■					
Project data archival record																							■	■	■	

Project Subteams:

- The SWAT modeling team includes M. Gabriel and C. Knightes
- The HSPF modeler is K. Kim with guidance from Y. Mohamoud
- The Groundwater modeler is S. Kraemer
- The Analytics team includes K. Price, M. Cyterski and T. Purucker
- The GIS analyst is L. Prieto
- The Fact-finding team includes K. Price, M. Gabriel, S. Kraemer and K. Sullivan

Primary responsibility for reports and journal articles:

- K. Sullivan and S. Kraemer and T. Purucker are responsible for the journal manuscript
- K. Sullivan C. Knightes and T. Purucker are responsible for the EPA report.
- K. Sullivan is responsible for the quality assurance documentation

A7 Data Quality Objectives and Criteria

This project is primarily a modeling study designed to assess the potential of water acquisition for HF activities that increase stress on available water resources by comparing relative changes in the balance between water availability and consumption under current use patterns and future HF activity scenarios. The project is expected to provide general and relative assessments of the likelihood of increased water stress relative to watershed conditions, time and seasonality, and demographic factors driving current water use. The project does not intend to quantify volumetric impacts on specific surface water or groundwater supplies within the study areas.

The project relies on external data provided by a number of federal, state, and local agencies and other sources to model water availability and estimate water use. The project will also generate a large amount of data from mechanistic and empirical hydrologic models that will be further analyzed. A number of steps in this process address data and modeling objectives and quality assurance criteria that will be applied during the project. The following are key aspects of the study that contribute to meeting data quality objectives:

- Estimation of water availability in surface water and subsurface volumes by mechanistic models will use existing, widely-used, well-tested and documented models. No changes to code will be used.
- Estimation of water availability based on statistical analysis of measured flow data will be based on published and peer-reviewed sources.
- Time series of meteorology and flow records are sufficiently long to establish long-term averages and incorporate periods of droughts into the summary statistics.
- Time series of meteorology and flow records are sufficiently complete that filling any gaps is minor
- Modeled flow record statistics are within the statistical variability of regional flow statistics based on measured records such as regional regression analysis
- Calibrated models achieve satisfactory fit statistics
- GW recharge estimates from watershed models are reasonably similar to GW models
- Selected subwatersheds are representative of the regional mix of land use and demographics
- Consumption data is sufficiently detailed within project areas to estimate consumption from domestic uses within reasonable expectations based on demographics
- Reasonable maximum and minimum consumption estimates are achieved for scenario development
- Hydraulic fracturing future scenarios are based on informed estimation methods
- Water stress determinations are supported by documented restrictions applied by regulatory agencies

These objectives and criteria are discussed in greater detail throughout this QAPP, including Appendix A.

A8 Special Training Requirements/Certification

All EPA members of the project team will be informed on the content and expectations outlined in the EPA Hydraulic Fracturing Study Quality Management Plan (QMP) and will thoroughly understand their responsibilities as defined in this project QAPP.

EPA, postdoctoral, and contractor personnel involved in modeling and analysis in this project hold advanced degrees from academic programs that are well-known for excellence in geology, hydrology, statistical analysis, and watershed modeling. Project personnel collectively have professional experience in watershed characterization, data management, advanced hydrological modeling, geospatial analysis, statistics, ecological risk assessment and operational management of water acquisition for industrial activities.

The GIS analyst supporting the project has successfully completed advanced training in ARC/INFO software products at the Center for Remote Sensing and Mapping Science at the Univ. of Georgia, and has routinely updated skills by completing onsite and online training.

No additional special training or certification is required for participants in this project beyond the already high degree of academic training and professional experience in hydrology obtained to fulfill job requirements commensurate with their current assignments.

A9 Documentation and Records

The project team will develop EPA reports and journal articles from the results of the scenario analysis.

The EPA project report and journal article(s) products will be reviewed by EPA for clearance and internally and externally peer-reviewed following the Quality Management Plan for the Study of the Potential Impacts of Hydraulic Fracturing on Drinking Water. Documentation for this project will be project reports and journal articles, records of peer review, and quality assurance activities, and archiving electronic and hard copy data files and work products. These records will be reviewed by the ORD QA team assigned to the project.

Project Report. This report will provide detailed project results for both project areas, including water availability assessments, consumption methods, fact finding results and scenario definitions, and HF scenario results. Results will also be synthesized for the physiographic, demographic and HF management factors that may lead to increased or adverse impacts on water supply vulnerability.

The Project 5B Phase 2 project will rely on other EPA projects to review and synthesize the pertinent scientific literature and important information on HF management practices and water management activities on a national scale and will not include a substantive literature review on these topics.

Journal article(s). The project will produce one or more journal articles summarizing key synthesis findings on the subjects addressed in the project report.

Quality Assurance Documentation

Accomplishment of quality assurance measures for the project will be included in project documentation, including results of model calibration and model input, output for inclusion in EPA quality assurance records.

Documentation of research outcomes will follow the ORD peer review process detailed for the EPA HF drinking water study. A flow chart specific for manuscripts targeted for peer-review journals is shown in Figure 23. Reports will be reviewed by extramural experts through a contractor-led letter review process as charted in Figure 24. The EPA Science and Technical Information Clearance System (STICS) will be used to manage routing and approval of manuscripts. Peer review will be coordinated internally by the project lead and Technical Information Manager (TIM) in Athens, and external peer review will be coordinated by the ORD Peer Review and Clearance team.

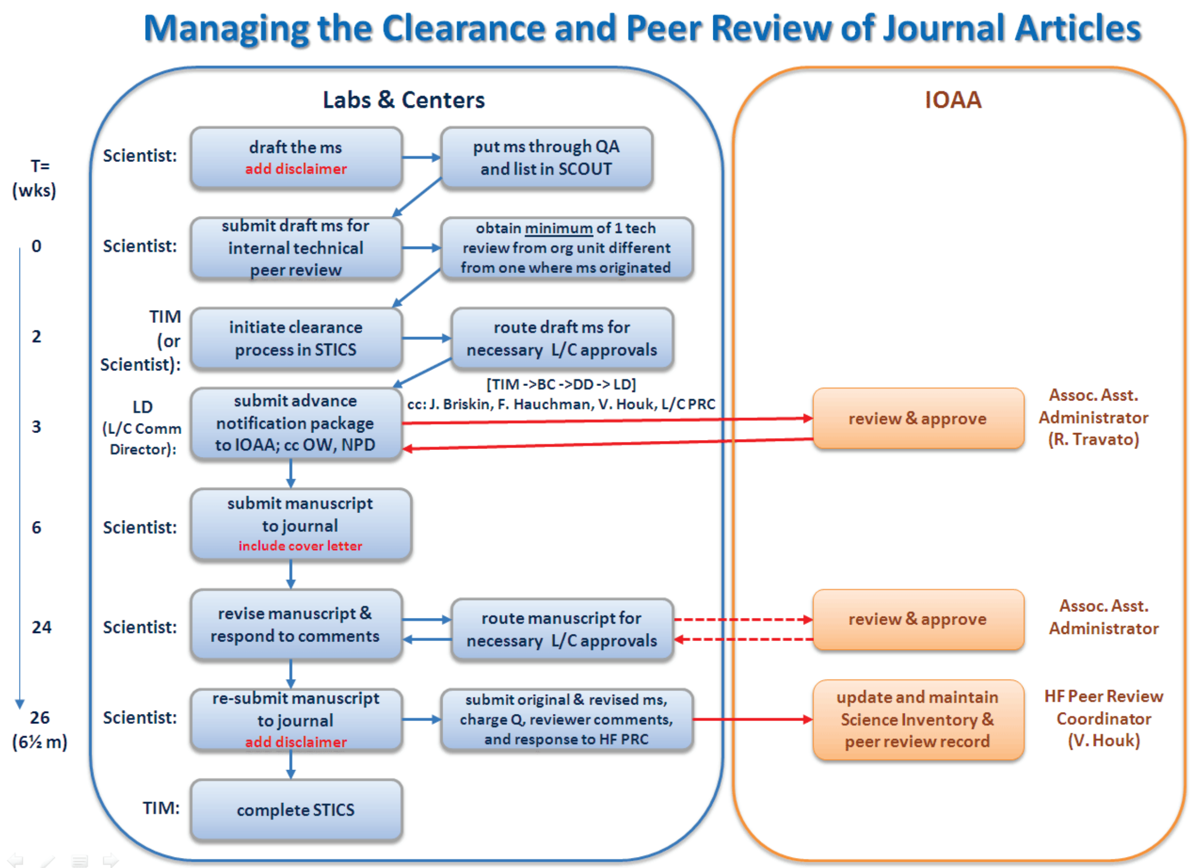


Figure 23. Flowchart for peer review of journal articles.

Quality Assurance Project Plan. The project lead will control review, revision, and distribution of the most recent version of the QAPP. Document control information (i.e., version and date) will appear in the upper right hand corner of each page of the QAPP. A signed approval form will accompany the approved QAPP. The final approved version of the QAPP will be distributed by the project lead to all project staff and to the ORD Management Team and ORD Quality Team. Any revisions to the approved QAPP will be circulated for review and approval. Approval is evidenced by signatures.

Project Records. The project team will generate research notes and quality assurance records during the conduct of the project as well as for data products. Research notebooks will be utilized by the staff and periodically audited by the project lead and onsite Quality Assurance Manager. Data documentation is further detailed in Sections B and C. Notes on those audits will be maintained.

Data manipulation and quality assurance procedures are defined in Sections B9 and B10. These activities will be documented and kept as part of the project quality assurance implementation files within the project folders.

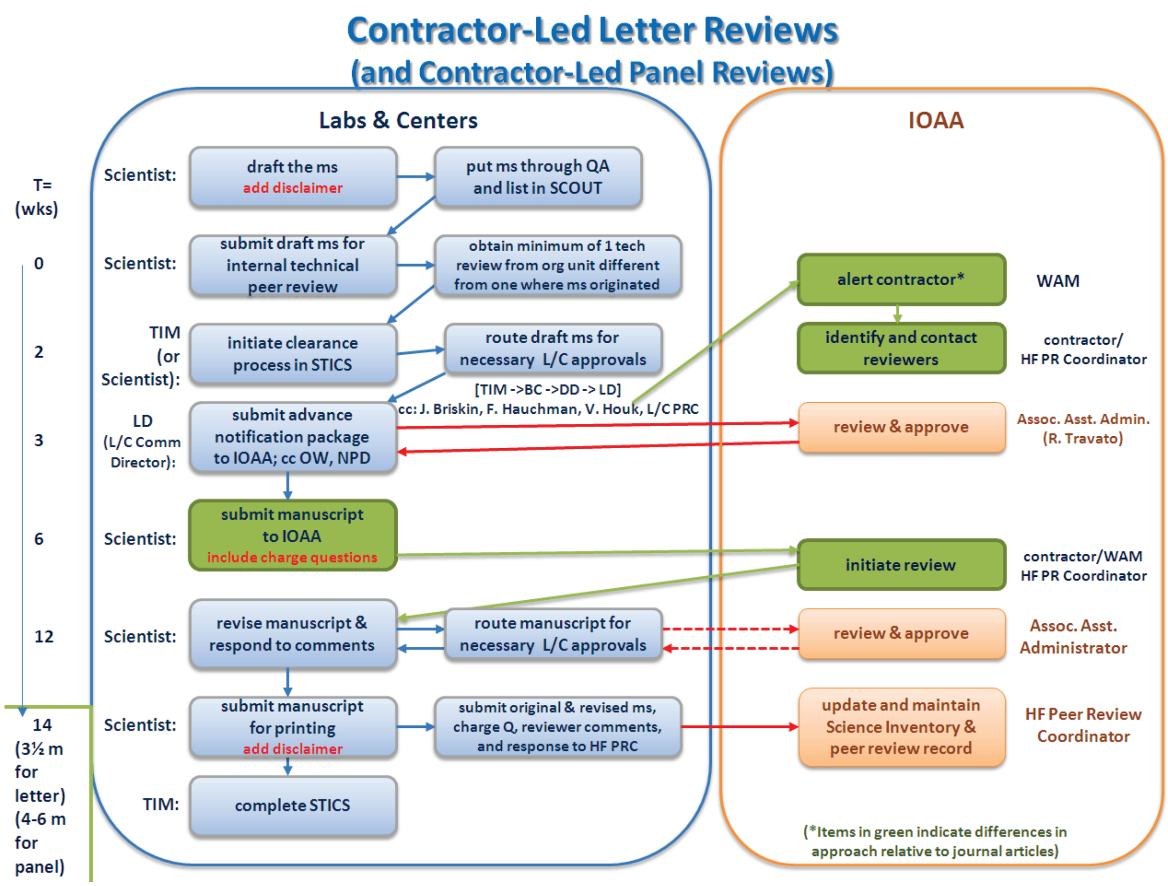


Figure 24. Flow chart for peer review of reports.

The project team will conduct all work and store all electronic files on a common ORD server at the ERD/NERL Athens facility on the L drive. The folder L:/Priv/HF has restricted access with permissions requested by the project lead and authorized by the OSIM site manager at the NERL/ERD facility in Athens, GA. This server is backed up weekly and maintained by EPA's ITI contractors under the supervision of OSIM. The electronic archives on the L: drive will be maintained at least until 2020.

All electronic data and work products are stored and archived according to the file structure and naming convention specified and maintained by the project lead. These data management structures and processes are described in greater detail in Section B9. The project leader maintains a version control system to ensure integrity of interim and final work products.

Hardcopy records such as research notebooks will be archived in the ORD/NERL/ERD repository.

All project products, documents and records will also be provided to the ORD Quality Team at the completion of the project for archiving in the location: O:\Priv\NRP_SSWR_HF. As Category 1 research, all project records will be permanently stored following EPA Records Schedule 501.

Any documents received as hard copy will be scanned, converted to PDFs and stored with all other electronic files.

Section B – Measurement and Data Acquisition

This project is based entirely on modeling with existing publicly available mathematical and empirical hydrologic models and therefore follows the Guidance for Quality Assurance Project Plans for Modeling EPA QA/G-5m (USEPA, 2002). The project relies exclusively on secondary sources of data, also referred to as non-direct measurements, for watershed and subsurface groundwater modeling, streamflow, and consumptive water use. Most of this information is provided to the public by federal and state agencies or local regulatory authorities or public utility district with documented quality control procedures. Some data will be acquired from public and private sources that may have less well documented quality assurance procedures.

The project generates no data from field or laboratory experiments defined as primary or direct data. Sections B1- B8 address quality assurance procedures for primary data and are not applicable to this project (Table 9).

Table 9. Section B elements that address primary data.

Section Element	Application in this Project
B1 Sampling Process Design	This section pertains to the acquisition of primary data and is not applicable to this project.
B2 Sampling Methods	This section pertains to the acquisition of primary data and is not applicable to this project.
B3 Sample Handling and Custody	This section pertains to the acquisition of primary data and is not applicable to this project.
B4 Analytical Methods	This section pertains to the acquisition of primary data and is not applicable to this project.
B5 Quality Control	This section pertains to the acquisition of primary data and is not applicable to this project.
B6 Instrument/Equipment Testing, Inspection, and Maintenance	This section pertains to the acquisition of primary data and is not applicable to this project.
B7.1 Instrument/Equipment Calibration and Frequency	No calibration of instrumentation. Model calibration is an activity within this project. See Appendix A for a description of model calibration as applied in this project.
B8 Inspection/Acceptance of Supplies and Consumables	This section pertains to the acquisition of primary data and is not applicable to this project.

B9 Non-direct Measurements

The project will use secondary (non-direct) data from many sources and will generate large volumes of modeled data during implementation. Each data source will undergo quality assurance analysis per project objectives. A general list of acquired data to support the three project tasks is provided in Table 10.

The project will require geospatial data for all or part of the Susquehanna and Upper Colorado River basins and specifically for subareas within Bradford County Pennsylvania and Garfield and Mesa counties in Colorado. Within these geographic regions, the project will select one basin at the 12-Digit HUC scale for analysis. A large portion of geospatial data needed will be obtained from national databases created by or for federal agencies (Table 11).

Table 10. List of anticipated secondary data by subtask

Task	Subtask	Anticipated Data
1	**Quantify available water resources**	• Hydrography • Land cover raster • Current land use • Digital elevation models (DEMs) • Detailed and generalized soil polygons • Groundwater levels from wells • Meteorological data (precipitation, air temperature, evapotranspiration) • Water storage data • High- and medium-resolution hydrography • 12-digit hydrologic units • Digital topographic maps (current and historical) • Political boundaries: state, county • Locations of roads, pipelines • USGS flow records
2	**Quantify Consumptive Use**	• Surface water and groundwater usage and return flow information (volume, timing, location) for major usage categories (public water supply, irrigation, industrial water supply, thermoelectric cooling water, other mining) • Surface and groundwater usage and return flow information (volume, timing, location) for HF operations (withdrawal, disposal, recycling) • Current HF well locations • Census and parcel data • Projected population data for 'future' simulations
3	**Scenario Analysis**	• Specific activities in HF water acquisition operations related to sourcing and scheduling • Regulatory water management considerations including flow criteria for environmental flows • Company water management best management practices including but not limited to recycling flow back and produced water. • Information needed to project future HF development activity using the U.S. EIA project methods: thermal maturity, depth and thickness of natural gas deposits.

Table 11. Secondary geospatial data that will be obtained from national databases

Data set	Scale or raster resolution	Source
National Elevation Dataset (NED) http://nationalmap.gov/elevation.html Ground elevation data in raster format. Various resolutions available depending on location.	1 arc-second (about 30 meters) 1/3 arc-second (about 10 meters)	The National Map Viewer http://viewer.nationalmap.gov/viewer/
National Land Cover Database 2006 (NLCD2006) http://www.mrlc.gov/nlcd2006.php Land cover information in raster format. It uses a 16-class classification scheme	30 meters spatial resolution	The National Map Viewer http://viewer.nationalmap.gov/viewer/
National Hydrography Dataset Plus (NHDPlus Version 2) http://www.horizon-systems.com/NHDPlus/index.php Medium resolution vector stream network with a suite of "value-added attributes".	1:100,000	Horizon Systems Corporation http://www.horizon-systems.com/NHDPlus/NHDPlusV2_data.php
National Hydrography Dataset (NHD) http://nhd.usgs.gov/data.html High resolution vector stream network.	1:24,000 or higher	USGS FTP website ftp://nhdftp.usgs.gov/DataSets/Staged/SubRegions/PersonalGDB/HighResolution
U.S. General Soil Map (STATSGO2) http://soils.usda.gov/survey/geography/SSURGO/description_statsgo2.html Generalized polygon-based soils information.	Compiled at a 1:250,000 scale.	USDA Geospatial Data Gateway http://datagateway.nrcs.usda.gov
Soil Survey Geographic (SSURGO) http://soils.usda.gov/survey/geography/SSURGO/description.html Detailed polygon-based soils information.	Compiled at scales ranging from 1:12,000 to 1:63,360.	USDA Geospatial Data Gateway http://datagateway.nrcs.usda.gov
Watershed Boundary Dataset (WBD) http://nhd.usgs.gov/wbd.html 12-digit hydrologic unit polygons (subwatersheds).	Minimum scale of 1:24,000	USDA website ftp://ftp.ftw.nrcs.usda.gov/wbd

Census Bureau 2012 TIGER/Line Shapefiles http://www.census.gov/geo/maps-data/data/tiger.html 2012 census block geography	Not available	http://www.census.gov/cgi-bin/geo/shapefiles2012/main
Population data from 2012 census http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml Demographic data from the 2012 census	Census block	http://factfinder2.census.gov/faces/nav/jsf/pages/download_center.xhtml
One Million-Scale State Boundaries State boundaries	1:1,000,000	National Atlas of the United States http://nationalatlas.gov/atlasftp.html#statep0
One Million-Scale County Boundaries of the United States http://nationalatlas.gov/mld/1cntyp.html County boundaries	1:1,000,000	National Atlas of the United States http://nationalatlas.gov/atlasftp.html#countyp
USGS US Topo Quadrangles http://nationalmap.gov/ustopo/index.html Digital topographic maps.	1:24,000	The National Map Viewer http://viewer.nationalmap.gov/viewer
USGS Historical Topographic Map Collection http://nationalmap.gov/historical/index.html Digital version of historical printed topographic maps	Most maps will likely have a scale of 1:24,000.	The National Map Viewer http://viewer.nationalmap.gov/viewer

The project will primarily use an Albers equal-area conic projection (see below) for model input, but may choose other projections for map creation or to accommodate different modeling and solution technologies. Specifications include:

Albers Equal-Area Conic Projection – Contiguous USA, USGS version
Spheroid: Geodetic Reference System of 1980 (GRS 1980)
Datum: North American Datum of 1983 (NAD 83)
Central meridian: -96.0°
1st standard parallel: 29.5°
2nd standard parallel: 45.5°
Latitude of origin: 23.0°
False easting: 0
False northing: 0
Linear unit: Meter

Geospatial data listed in Table 11 is provided by federal agencies with documented quality assurance procedures as specified by each source. This project will accept the data as qualified for use as is, although we will examine metadata associated with these databases to assess their quality and appropriateness for the project.

The metadata of secondary geospatial data will also be tested with the validation tool of the EPA Metadata Editor (EME) (https://edg.epa.gov/EME/) to determine if it meets the minimum requirements of the Federal Geographic Data Committee's (FGDC) Content Standard for Digital Geospatial Metadata (FGDC, 1998) and the EPA Geospatial Metadata Technical Specification (USEPA, 2007). In instances where these minimum requirements are not met, all efforts will be made to contact the data originator/provider to obtain missing information. The EME will then be used to update the metadata record. If the missing information cannot be obtained, the project leader will decide if the data can still be used. The rationale for any such decision will be recorded and included in the quality record.

Other geospatial data may be obtained from state, county, or local government agencies and non-spatial data may be obtained from peer-reviewed journals, academic institutions, federal, state, county, or local government agencies, or private sources. Table 12 contains a list of other anticipated secondary data sources.

Table 12. Other anticipated secondary data sources

Source	**Information to be obtained**
USGS National Water Information System http://waterdata.usgs.gov/nwis/rt	Streamflow and aquifer water level elevation in wells.
EPA Safe Drinking Water Information System http://www.epa.gov/enviro/facts/sdwis/search.html	Information on public drinking water systems.
Susquehanna River Basin Commission Water Resources Portal http://gis.srbc.net/wrp/	Information on and visual locations of proposed and approved HF well pads, water withdrawals, and water consumptive uses in the Susquehanna River Basin.
NOAA National Climatic Data Center http://www.ncdc.noaa.gov/cdo-web/	Meteorological data (precipitation, air temperature, evapotranspiration)
FracFocus 2.0 Database http://fracfocus.org/	HF well information (water volumes, water depth)

Publicly available information will be checked for history of quality assurance procedures by the GIS analyst. The project will obtain metadata for these sources and review them for quality assurance procedures. If QA procedures are specified for the sources, they will be accepted as qualified. Suitability for use will also be verified by the project leader in consultation with the modeling and analytical experts.

Data such as consumptive water use or well data may be obtained from private companies. Any data classified as confidential business information (CBI) will be treated confidentially following EPA's CBI procedures. The location information associated with Community Water Systems is also considered sensitive and is controlled by EPA. The FracFocus Database maintained by the oil and gas industry accepts information from well operators can be a valuable source of information to this project and is publicly accessible (GWPC and IOGCC, 2013). The EPA project [FracFocus Analysis] is reviewing the quality of this data for use in the suite of EPA Hydraulic Fracturing studies. This project will follow EPA guidance on the quality and acceptability if FracFocus data for use in this project.

All secondary data will also be assessed against the following general acceptance criteria:

- Reasonableness - Data sets will be checked for reasonableness. When appropriate, graphical methods will be used to evaluate potential anomalous entries that may represent data entry or analytical errors. If major data anomalies are found, the project will seek clarification from the agency from which the data were collected. The project will correct and document minor deviations.

- Representativeness - The data provides the information needed at the required spatial and/or temporal resolution. Some requirements will be specified for each subtask as analysis proceeds. If more than one dataset is available, we will select the one with the most appropriate resolution for modeling purposes.

- Completeness - The dataset encompasses the designated study areas and has little or no missing information that cannot be interpolated from existing data. Data that has gaps will be used only if it is the best, most complete information available. Inevitably, there may be data gaps in some types of data (e.g., weather). If the missing information cannot be obtained, the project leader in consultation with the project team will determine if the data can still be used. The rationale for any such decision will be recorded and included in the quality record.

- Currentness– The most current versions of geospatial data will be used.

- Comparability - Data sets will be checked with respect to variables of interest, commonality of units of measurement, and similarity in analytical and QA procedures. The project team will evaluate comparability of data by similarity in geographic, seasonal, and sampling method characteristics. Units will be carefully checked when compositing data from models and various secondary data sources in analysis steps.

Though not identified as explicit acceptance criterion for this project, priority will be given to peer-reviewed data and data that have undergone documented QA procedures by their sources.

Data sources will be appropriately cited in project deliverables.

Quality Assurance for Secondary Data. It is anticipated that all secondary data will be acquired in electronic format. Acquisition of data will be documented in dedicated laboratory notebooks using archival-quality pens to record the following information: download website address, download date, filename(s), location where the data have been stored, and name of person who acquired the data. If data were provided via personal contact, that person's name and information will be recorded and any correspondence concerning the data will be maintained. Original acquired data will be archived prior to manipulation.

The project will also maintain a Microsoft®Excel spreadsheet to track the metadata information described above. The GIS analyst will maintain and control the spreadsheet to track secondary data and will be responsible for their upkeep. Information stored in the Excel metadata library file is listed in Table 13. If data were provided via a personal contact, the person's name and information will be recorded and any correspondence concerning the data will be maintained. To make it easy to recognize the most current version of the file, the date and time will be part of the filename.

Table 13. Metadata of secondary data that will be tracked in a the Microsoft®Excel spreadsheet.

Field	Description
Data type	General category the data represents (elevation, soils, streams, etc.)
Data format	Data structure (vector, raster, tabular, etc.)
Data filename(s)	Name used to store the data
Data path	Computer drive and directory where data is located
Acquisition method	How data was obtained (internet download, personal contact, etc.)
Acquisition date	Date file was acquired
Acquisition person	Person responsible for the acquiring the data
Notebook number	Research notebook number where acquisition was recorded
Data scale or raster resolution	Scale of vector data, resolution of raster data, or not applicable for tabular data (NA)
Metadata status	Available, unavailable, in process of being acquired
Intended use	Way data is going to be used (model input, map creation, reference, etc.)

B10 Data Management

Project 5B Phase 2 of the water acquisition project will acquire data from secondary sources, manipulate secondary data to meet use requirements, generate data from model output, and create summary and processed files for further statistical and analytical treatment. It is anticipated that the project will acquire or generate little or no hardcopy data. The project will maintain a data management system that protects integrity of the data received and generated throughout the project. This includes file management systems, version control, archiving procedures, and quality assurance activities, some of which were discussed in section B9.

The project team will conduct all work and store all electronic files on a common ORD server at the ERD/NERL Athens facility on the L drive. The folder L:/Priv/HF has restricted access with permissions requested by the project lead and authorized by the OSIM site manager. This server is backed up weekly and maintained by EPA's ITI contractors under the supervision of OSIM. All electronic project work will be conducted within L: drive folders.

Data will be maintained in an organized electronic filing system for maintaining control and integrity of the variety of data and manipulations required. This file management system will accommodate:

- Original secondary data
- Manipulations to secondary data
- Model Output Data
- Processed Model Data
- Analytical Processing Files
- Statistical Processing

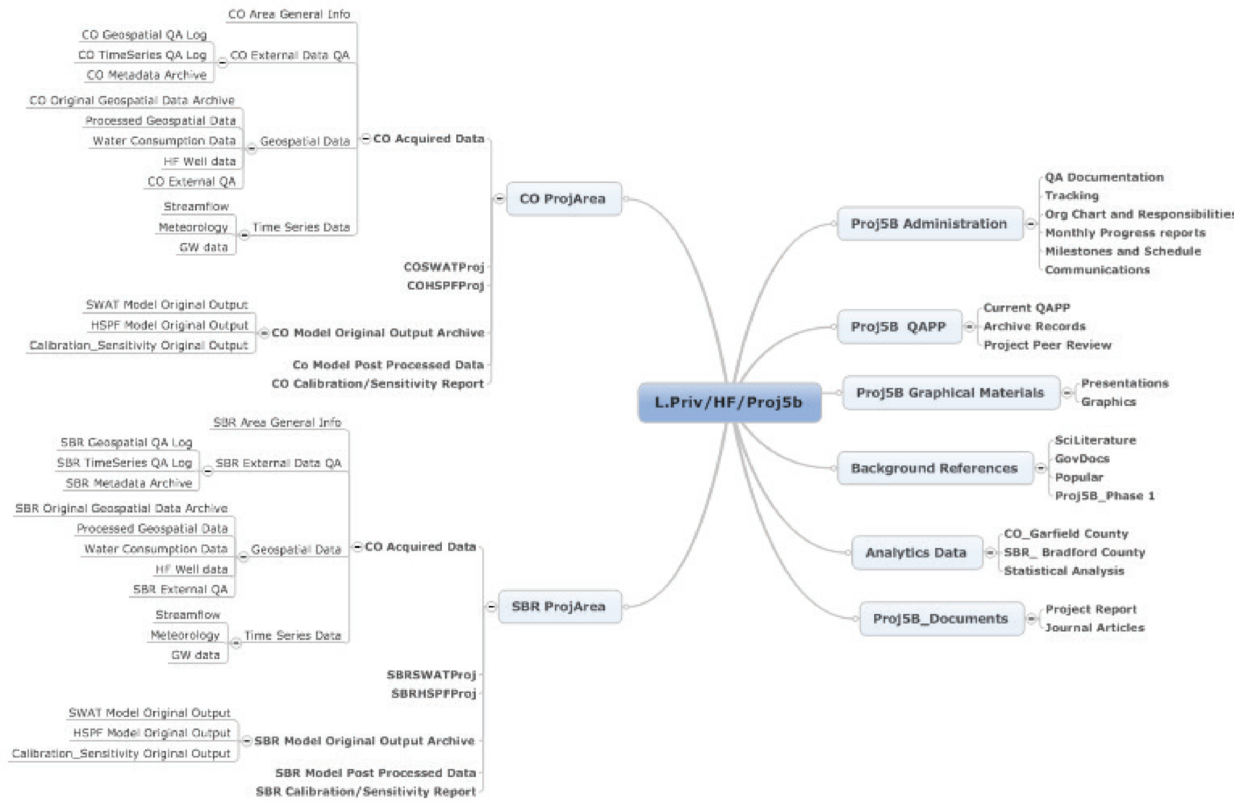
Shared Files. Documents and data files will have shared access among the project team. The project team has unique filename structures and version control systems when working on common documents. Conventions are referenced in the Admin corner of the project file for reference and will be documented in a file accessible to all users. The project lead will monitor the orderliness of the shared drive system.

Filename_initials1_counter1_initials2_counter2MMDDYYYY.extension

File organization of electronic files on the Shared Drive (L:/Priv) folder is shown in Figure 25. There are storage areas for original data received from external sources, original model output runs, calibrations, and processed data from original model data to files suitable for statistical analysis.

Acquired data. Original copies of all acquired data are stored. Manipulations and processing of geospatial or time series data are documented. Processed data is stored separately from original data.

Figure 25. Project 5B ORD/NERL/ERD L/Priv/HF folder organization.



***GIS Data Processing and Manipulations**.* Any geo-processing or manipulation applied to the secondary GIS data (such as reprojections, clips, or field calculations) will be documented initially in laboratory notebooks as described in Section B9. The process date and person doing the manipulation will be recorded, as well as the GIS tool and specific parameter(s) used. A unique name will be given to the modified dataset and its location recorded.

Modified datasets in vector format or in raster formats that allow long names (e.g., TIFF or JPEG) will be given descriptive names to indicate the geography they represent and dataset from which they were derived. For example, a stream shapefile for Bradford County extracted from the National Hydrography Dataset will be named *bradford_NHDHigh.shp*, a name that will carry over to the next step. If this shapefile is reprojected to Albers, it will be renamed *bradford_NHDHigh _alb.shp*. Under some circumstances, it may not be necessary to specify the dataset: for example, HUCs extracted from the Watershed Boundary Dataset will simply be named with the HUC number (i.e., *0205010601.shp* and *0205010601_alb.shp*). Raster datasets in ESRI grid format are limited to filenames of no more than 9 to 13 characters, depending on the number of bands in the raster. In these situations, unique names will be created to be as descriptive as possible. If necessary, an incremental numbering system will be used (e.g. mosaic_01, mosaic_02, and so forth). To avoid altering original data, processes that do not create a new file will be performed in a copy of the original dataset. The copy will be given a unique name as explained above. The EPA metadata editor (EME) will be manually updated in those instances where GIS software does not automatically record the change.

Vector national GIS datasets will be clipped to the study areas or a buffered study area, as necessary, and areas of interest will be extracted from raster datasets as needed. For some elevation datasets it may be necessary to merge (mosaic) two or more data tiles before areas of interest can be extracted. When extracting or merging raster data, random visual checks will be performed to insure cells of the newly-created raster align with the original cells and cell values have not changed. Any field calculations will be checked by randomly selecting a subset of the values and hand-calculating them. If needed, area calculations will be performed while the dataset is being stored or displayed in a projection that preserves area such as the Albers Equal Area projection. Area calculations will be checked by randomly selecting several polygons and calculating the area with the GIS measuring tool.

Any changes made to secondary data when preparing data for project needs will be recorded in a Data Manipulation Log file. Information tracked in the Data Manipulation Log is listed in Table 14. There will be one log file per dataset. The file will be stored in the project's network drive: L:/HF/XXX , named HF5.B Secondary_data_tracker_[Original Data set name], and controlled by the GIS analyst.

Table 14. GIS data information for quality assurance, including metadata received with the data and data manipulation tracking.

Metadata stored for each file (see section B9)	**Data Manipulation Log entries:**
• Data type (elevation, soils, streams, etc.) • Data format (vector, raster, tabular, etc.) • Data filename(s) • Data path (computer drive and directory) • Acquisition method (if downloaded, website will be included) • Acquisition date • Metadata status (validated, in process of being validated, waiting to be validated, or no metadata available) • Data scale or raster resolution	• **Data manipulations (reprojections, clips, calculations, etc.)** • **Manipulation date** • **Person responsible for the manipulation and manipulation date** • **New file name used for each processing step** • **Location of processed data** • **Data use (model input, map creation, reference, etc.)**

Model output data. Original model output with metadata on parameterization applicable to each model is stored for archival purposes. Original modeled output consisting of daily flow values is post-processed to flow statistics, sent to analytics data file, and archived. Model inputs and outputs and other data calculations will be verified by initial and final reviews. This will include checking to ensure that model input files contain intended input values and that model outputs correspond to the correct set of inputs.

Analytics data. Data on hydrology statistics by subwatershed are composited with consumption data in the Analytics data folder for scenario analysis. Statistical analysis is conducted on files in this folder. Data transferred to analytical files will be checked for correct reference to original model input parameterization as metadata, and for correct data entry and identification of units and parameter names. Records will be checked by a second party and a log maintained of file checks.

Hardcopy data. Any data received in hard copy form will be manually entered into Excel spreadsheets or model input files. One hundred percent of the data will be checked to ensure accuracy. Detected errors will be corrected. Source documentation used for manually entered data will be scanned, converted to PDFs, and electronically stored with all other electronic files.

Project scientists will maintain research notebooks following EPA standards, including use of archival-quality ink. The research notebooks will be reviewed periodically by the project lead and the QA Manager.

Section C – Assessment and Oversight

C1 Assessments and Response Actions

The EPA Quality Assurance Manager will conduct a Technical Systems Audit (TSA) to ensure this QAPP is being followed during execution of the research project. Work conducted for this project will undergo ongoing technical review by personnel at EPA/ORD/NERL/ERD who are implementing the project.

The project lead will have responsibility for monitoring project activities and identifying or confirming quality problems. Any problems will be brought to the attention of the ORD Management Team and the ORD QA Team, who will document the nature of the problem, initiate corrective actions, and ensure the recommended corrective action is carried out.

This QAPP describes processes for model sensitivity and uncertainty analysis (Section A5, Appendix A), data quality assessment (Section B9), data management and error checking (Section B10), and model performance evaluations (Appendix A). The project team will assess model sensitivity to parameters in calibration steps as well as analysis steps to understand model performance relative to modeling objectives. It has provisions for data validation and usability (Section D)

Many technical problems that might occur such as modifying the technical approach or correcting errors or deficiencies in documentation can be solved immediately by the technical staff. The project team is responsible for documenting a response to any significant findings. Immediate corrective actions are part of normal operating procedures and noted in project records. Problems that cannot be solved in this way require more formalized corrective action. If quality problems that require such attention are identified, the QA Officers will determine whether attaining acceptable quality requires short- or long-term actions.

The project lead will perform surveillance activities to ensure that management and technical aspects are implemented properly according to the schedule and quality requirements specified in this QAPP. These activities will include assessing how project milestones are achieved and documented, corrective actions are implemented, budgets are followed, reviews are performed, and data are managed.

The technical systems assessment will include assessment of data collection activities, documentation, quality checks, record management, and reporting.

Only peer-reviewed and documented publicly available models are used in this project. Therefore, this QAPP does not plan activities for hardware or software testing, code verification, checking for programming or mathematical errors, or other quality assurance activities associated with technology development and deployment. No new modeling products or changes to existing models will be implemented during this study.

C2 Reports to Management

The project lead serves as the primary liaison between the ORD Management Team and the ORD QA Team and assigned personnel conducting the study. The project lead will participate in bi-weekly conference calls with NERL management led by the Associate Director. The project lead will also represent the project in monthly conference calls with the EPA HF Study Lead. The project lead will participate in the annual All Investigators meeting if appropriate, and in EPA workshops or other events at the request of the ORD Management Team.

The project team will participate in weekly meetings with ERD management and as needed to accommodate team activities.

Documentation to be submitted for quality assurance and review purposes includes:

- Draft and final QAPP
- Monthly progress reports
- Project reports
- Journal manuscript(s)
- Final model and other data files (upon completion of the project or upon request of EPA).

Section D – Data Validation and Usability

This project uses well-known watershed and groundwater models to estimate streamflow and groundwater storage that have a history of extensive use by public agencies and universities and documentation in peer-reviewed literature. The modeling requires input data that must be acquired from federal, state, and local authorities, as well as private entities. The quality of data used for and generated during modeling will be reviewed and verified at multiple levels by project technical staff and QA officers, as described in detail in other sections (Table 15).

Table 15. Summary of Section D1 And D2-- Model and Data Review, Verification and Validation.

Topic	Validation Methods and Criteria	Additional Description
Review of model components	Computational models used in project are widely used, peer-reviewed, well documented and curated. Watershed models: SWAT, HSPF Groundwater models: GFLOW, MODFLOW	A5
Review of Input Data	Hydrologic. meteorologic, soils, land use data for modeling obtained from qualified secondary data acquired from federal agencies. Metadata documented. Streamflow and weather data obtained from USGS and NOAA accepted as qualified, but examined for completeness. Data not used if gaps > 10% of record. Consumptive water use data obtained from federal, state, local and other regulatory authorities. Data accepted as qualified, metadata documented. Well data obtained from industry data base FracFocus: project will use data qualified by EPA project tasked to review this data-source.	B9
Review of Model Performance Tests	Multi-objective parameter optimization used for calibration and validation steps for each model application. Multiple models used in calibrating. Determination of goodness-of-fit with modified Nash/ Sutcliffe metrics. Comparison of statistics computed from modeled streamflow to streamflow records to empirically modeled estimates of streamflow statistics from published relationships.	A5, Appendix A

D1 Data Review, Verification, and Validation

Data review, verification, and validation will focus on acceptability of the input data used for calculations and modeling. All original and modified data files will be reviewed for input, handling, and calculation errors. Any issues identified through this review process will be evaluated and, if necessary, data will be corrected and analysis carried out using corrected data.

Deviations from the approved QAPP could occur as the project proceeds. The need for adjustments may arise based on data validation and quality assurance checks, outcomes of model initialization and calibration steps, scenario development, and so on. Deviations will be documented in writing and reviewed by the ORD Management Team and the ORD Quality Team. The QAPP will be revised accordingly and recirculated for Quality Assurance review and approval.

D2 Verification and Validation Methods

The integrity of model output data will be verified and validated through a review of data files by project technical staff. Reviews may include a thorough evaluation of content and/or a "spot-check" of calculated values. Should a review identify an aberration from established data quality objectives and criteria (Section A7), the reviewer will notify those responsible for taking corrective actions. QA officers will be notified if corrective action may be required.

Evaluation of whether model components and their outputs are satisfying the DQOs will be an ongoing process for QA personnel during model calibration and validation stages of the project. In-progress assessments of validation issues will be discussed by a team including technical and QA representatives from EPA. The final authority for resolving validation issues will be the Quality Assurance Manager for EPA ORD (see Section A4).

Results of performance evaluations will be logged and integrated into project documentation at the conclusion of the project, as well as any corrective actions that were implemented.

D3 Reconciliation with User Requirements

The objective of the project is to assess potential impacts on water acquisition to support current and future scenarios of increased hydraulic fracturing activity that could affect domestic water supplies. The water balance method that is the basis for this project compares consumption by all water uses, including domestic uses, with water that is available at selected withdrawal sites in a catchment and at specific public supply locations. The study is designed to assess water availability and consumption in a range of watershed sizes -- from headwaters to 200-300 mi^2 basins -- and on short-to-moderate time increments (less than one year). A number of challenges can affect development of an effective water accounting method.

The project anticipates no issues with achieving project objectives for assessing surface water availability at appropriate spatial and temporal resolution due to availability of measured data or analytical capabilities with available watershed models and published regional empirical relationships that can corroborate modeled results.

Although sound mechanistic groundwater models will be used in the project, assessing subsurface water storage is inherently more uncertain due to lack of data on subsurface physical characteristics that determine groundwater movement and few observations to validate models. The project expects to address subsurface water withdrawals per project objectives. Due to data limitations, however, the analysis can only provide quantitative estimates at a very coarse scale of resolution and with large uncertainties. Limitations of the analysis will be fully addressed in project reports and journal articles.

Potential for water stress as evaluated in the project, requires reasonable estimates of water consumption at the same temporal and spatial scales determined for water availability. Consumptive use data is collected by federal and state agencies as well as by public and private suppliers and users. Most information is aggregated to the county level for reporting, which is much larger than the spatial scales of analysis used in the study. The availability of spatially-explicit data will determine how precisely water consumption can be determined within subwatersheds and at selected time scales of resolution. The project has provisions for downscaling aggregated data to provide necessary consumptive use estimates, but any statistically-derived values will have greater uncertainty. Uncertainties in consumption estimates will be discussed in project reports and journal articles.

Similarly, the project recognizes that hydraulic fracturing operations are conducted with considerable attention to water management, both in interaction with regulatory authorities and in company operations, to manage costs and address environmental needs. These operational aspects are important to minimizing impacts from HF operation, but are locally variable due to differences in regulatory authorities and operators. Defining the water management and use details in the HF scenarios represents a static definition of management practices that are actually applied in a complex dynamic operating environment. The project will seek information from operators and regulatory authorities to include in scenario development, recognizing it is not possible to analyze all possible situations or define scenarios that cover the full range of possibilities.

The objective of the Phase 2 project is to identify the circumstances under which current and projected water use for hydraulic fracturing have potential to increase local water stress and potentially impact domestic water supply. Despite data or scenario resolution issues that may arise as described above, the project results should still be instructive in identifying general watershed, seasonal, and demographic factors that contribute to higher potential water stress when mixed with hydraulic fracturing operations. Any data limitations encountered in consumptive use representation or in applying management factors in scenario development should not prohibit meaningful generalized interpretations of them. Any data quality issues determined to affect the conclusions or recommendations of this project will be discussed in the project report and journal articles.

References

Abbaspour, K., 2009. SWAT-CUP2: SWAT Calibration and Uncertainty Programs – A User Manual. Department of Systems Analysis, Integrated Assessment and Modeling, SIAM, Dübendorf, Switzerland: 95 pp. <http://www.eawag.ch/forschung/siam/software/swat/index>.

American Petroleum Institute (API), 2010. Water management associated with hydraulic fracturing. API Guidance Document HF2, First Edition, June, 2010. American Petroleum Institute, Washington D.C. 26 pp.

Bekele, E.G., Nicklow, J.W., 2007. Multi-objective automatic calibration of SWAT using NSGA-II. Journal of Hydrology, 341(3-4): 165-176.

Beven, K., 2006. A manifesto for the equifinality thesis. Journal of Hydrology, 320(1-2): 18-36.

Bicknell, B.R., J.C. Imhoff, J.L. Kittle Jr., T.H. Jobes, and A.S. Donigian, Jr. 2001. Hydrological Simulation Program - Fortran (HSPF). User's Manual for Release 12. U.S. EPA National Exposure Research Laboratory, Athens, GA, in cooperation with U.S. Geological Survey, Water Resources Division, Reston, VA.

Blackmore, J.M., Plant, R.A.J., 2008. Risk and resilience to enhance sustainability with application to urban water systems. Journal of Water Resources Planning and Management-Asce, 134(3): 224-233.

Brathwaite, L.D., 2009. Shale-Deposited Natural Gas: A Review of Potential. California Energy Commission, CEC-200-2009-005, 32 pp.

Cadmus Group, 2012a. Modeling the impact of hydraulic fracturing (HF) on water resources based on water acquisition scenarios Susquehanna River Basin case study. Prepared by Aqua Terra Consultants. Subcontract # 002-AT-1. Draft, December 31, 2012. 82 pp.

Cadmus Group, 2012b. Impact of water use and hydro-fracking on the drinking water resources of the Upper Colorado River basin. Prepared by Deb Debjani, Srinivasan, R. and Shiremen, J. Texas A&M University. Draft, December 26, 2012. 54 pp.

Colorado Water Conservation Board (CWCB), 2004. Appendix A: State of Colorado Population Projections 2000 to 2030. Colorado Department of Natural Resources, Statewide Water Supply Initiative Report, < http://cwcb.state.co.us/public-information/publications/Documents/ReportsStudies/SWSIAppendices/Appendix%20A%20State%20of%20Colorado%20Population%20Projections%202000%20to%202030.pdf>

Doherty, J., Johnston, J., 2003. Methodologies for calibration and predictive analysis of a watershed model. Journal of the American Water Resources Association 39(2): 251-265.

Donigian, A.S. Jr., 2000. HSPF Training Workshop Handbook and CD. Lecture #19. Calibration and Verification Issues, Slide #L19-22. EPA Headquarters, Washington Information Center, 10-14 January, 2000. Presented and prepared for USEPA, Office of Water, Office of Science and Technology, Washington, DC.

Dunne, T., Leopold, L.B., 1978. Water in Environmental Planning. WH Freeman and Company, New York, 818 pp.

Efstratiadis, A., Koutsoyiannis, D., 2010. One decade of multi-objective calibration approaches in hydrological modelling: a review. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 55(1): 58-78.

FGDC, 1998. Content Standard for Digital Geospatial Metadata, FGDC-STD-001-1998 <http://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata/base-metadata/index_html>

Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., 2007. The Soil and Water Assessment Tool: Historical development, applications, and future research directions. Trans. ASABE, 50(4): 1211-1250.

Ground Water Protection Council (GWPC) and ALL Consulting, 2009. Modern Shale Gas Development in the United States: A Primer. US Department of Energy, 98 pp.

Ground Water Protection Council (GWPC), 2012. A White Paper Summarizing the Stray Gas Incidence and Response Forum, GWRC, 48 pp.

Ground Water Protection Council (GWPC) and Interstate Oil and Gas Compact Commisison (IOGCC), 2013. FracFocus Chemical Disclosure Registry. <http://fracfocus.org>

Gupta, H.V., Wagener, T., Liu, Y., 2009. Reconciling theory with observations: elements of a diagnostic approach to model evaluation. Hydrological Processes, 22(18): 3802-3813.

Guzman, J.A., D.N. Moriasi, P.H. Gowda, J.L. Steiner, J.G. Arnold, P.J. Starks, R. Srinivasan, 2013. A fully integrated SWAT-MODFLOW hydrologic model, in preparation.

Haitjema, H.M., 1995a. On the residence time distribution in idealized groundwatersheds. Journal of Hydrology, 172(1-4): 127-146.

Haitjema, H.M., 1995b. Analytic Element Modeling of Groundwater Flow. Academic Press, San Diego, CA, Out of Print (contact www.haitjema.com).

Harbaugh, A.W., 2005, MODFLOW-2005, The U.S. Geological Survey modular ground-water model—the Ground-Water Flow Process: U.S. Geological Survey Techniques and Methods 6-A16 <http://water.usgs.gov/nrp/gwsoftware/modflow.html>

Hettinger, R.D., Kirschbaum, M.A., 2002. Stratigraphy of the Upper Cretaceous Mancos Shale (upper part and Mesaverde Group in the southern part of the Uinta and Piceance basins, Utah and Colorado. U.S. Geological Survey IMAP: 2764.

Hill, M.C., 1998. Methods and Guidelines for Effective Model Calibration. U.S. Geological Survey Water-Resources Investigations Report 98-4005.

Hill, M.C., Tiedeman, C.R., 2007. Effective Groundwater Model Calibration,with Analysis of Data, Sensitivities, Predictions, and Uncertaintiy. New York: John Wiley and Sons, 384 pp.

Hunt, R.J., 2006. Ground water modeling applications using the analytic element method. Ground Water, 44(1): 5-14.

INTBM, 2013. Integrated Northern Tampa Bay Model Application: Peer Review Report, prepared for Tampa Bay Water (Clearwater, FL) and South Florida Water Management District (Brooksville, FL) by WEST Consultants, Inc Bellevue, WA, Earthfx, Inc, Toronto, Canada and Hydrocomp, Inc (Menlo Park, CA).

Johnson, T.E., Butcher, J.B., Parker, A., Weaver, C.P., 2012. Investigating the sensitivity of U.S. streamflow and water quality to climate change: The EPA Global Change Research Program's "20 Watersheds" Project. Journal of Water Resources Planning and Management 138(5), 453-464.

Kenny, J.F., Barber, N.L., Hutson, S.S., Linsey, K.S., Lovelace, J.K., Maupin, M.A., 2009, Estimated use of water in the United States in 2005: U.S. Geological Survey Circular 1344, 52 pp.

Kent, H.C., 1968. Biostratigraphy of Niobara-Equivalent Part of Mancos Shale (Cretaceous) in Northwestern Colorado. AAPG Bulletin 52(11): 2098-2115.

Klemes, V., 1986. Operational testing of hydrological simulation models. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 31(1): 13-24.

Kraemer, S.R., 2007. Analytic element ground water modeling as a research program (1980-2006). Ground Water 45(4): 402-408.

Krysanova, V., Arnold, J.G., 2008. Advances in ecohydrological modeling with SWAT: A Review. Journal of Hydrologic Sciences, 53: 939-947.

Kuuskraa, V.A., Ammer, J., 2004. Tight gas sands development: How to dramatically improve recovery, GasTIPS 10(1): 15-20.

Matott, L.S., Babendreier, J.E., Purucker, S.T., 2009. Evaluating uncertainty in integrated environmental models: A review of concepts and tools. Water Resources Research, 45.

Mitchell-Bruker, S., Haitjema, H.M., 1996. Modeling steady state conjunctive groundwater and surface water flow with analytic elements. Water Resources Research, 32(9): 2725-2732.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models, Part 1: A discussion of principles. Journal of Hydrology, 10: 282-290.

National Atlas of the United States, 2012. 1:1,000,000-Scale County Boundaries of the United States. National Atlas of the United States, July 2012. <http://nationalatlas.gov/mld/1cntyp.html>

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and Water Assessment Tool Theoretical Documentation Version 2009, Texas A&M University System, College Station, TX.

Nicot, J.P. 2013. Sources of data for quantifying hydraulic fracturing water use in Texas. Presented at the U.S. E.P.A. Water Acquisition Workshop, held June 4, 2013, Washington D.C.

Ott, R., Longnecker, M., 2001. An Introduction to Statistical Methods and Data Analysis. Duxbury, Pacific Grove, CA, 1152 pp.

Pranter, M J. and N. K. Sommer, 2011. Static connectivity of fluvial sandstones in a lower coastal-plain setting: An example from the Upper Cretaceous lower Williams Fork Formation, Piceance Basin, Colorado, AAPG Bulletin, vol. 95, p. 899-923.

Price, K., Purucker, S.T., Kraemer, S.R., Babendreier, J., Knightes, C.D., 2013. Comparison of radar and gauge precipitation data in watershed models across varying spatial and temporal scales. Hydrological Processes, Published online (Wiley Online Library).

Price, K., Purucker, S.T., Kraemer, S.R., Babendreier, J.E., 2012. Tradeoffs among watershed model calibration targets for parameter estimation. Water Resources Research, 48.

Richenderfer, J. 2013. Water acquisition for unconventional natural gas development within the Susquehanna River Basin. Presented at the U.S. E.P.A. Water Acquisition Workshop, held June 4, 2013, Washington D.C.

Robson, S.G., Banta, E.R., 1995. Ground Water Atlas of the United States: Arizona, Colorado, New Mexico, Utah. U.S. Geological Survey HA-730-C.

Rojas, N.A., 2008. Spectral Decomposition Applied to Time-Lapse Seismic Interpretation at Rulison Field, Garfield County, Colorado. M.S. Thesis, Colorado School of Mines, Golden, Co., 111 pp.

Rutledge, A.T., (1998). Computer programs describing the recession of ground-water discharge and for estimating mean ground-water recharge and discharge from streamflow records --- update, US Geological Survey Water Resources Investigation Report 98-4148.

Satterfield, J., Kathol, D., Mantell, M., Hiebert, F., Lee, R., Patterson, K. 2008. Managing water resource challenges in select natural gas plays. Presented at GWRC Anuual Forum, September 20-24.

Schaefli, B., Gupta, H.V., 2007. Do Nash values have value? Hydrological Processes, 21(15): 2075-2080.

Schaefli, B., Hingray, B., Niggli, M., Musy, A., 2005. A conceptual glacio-hydrological model for high mountainous catchments. Hydrology and Earth System Sciences, 9(1-2): 95-109.

Shenk, G.W., J, W., L, L., 2012. Enhanced HSPF Model Structure for Chesapeake Bay. Journal of Environmental Engineering, 138(9): 949-957.

Soil Survey Staff, Natural Resources Conservation Service, 2013.United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database. < http://soildatamart.nrcs.usda.gov>

Strack, O.D.L., 1989. Groundwater Mechanics. Prentice Hall.

Strahler, A.N., 1957. Quantitative analysis of watershed geomorphology. Transactions of the American Geophysical Union, 38(6): 913-920.

Stuckey, M.H., 2006. Low-flow, base-flow, and mean-flow regression equations for Pennsylvania Streams. U.S. Geological Survey Scientific Investitations Report 2006-5130, 84 p.

Susquehanna River Basin Comission (SRBC), 2008. Consumptive Use Mitigation Plan. SRBC Publicaiton No. 253, 62 pp.

Tetra Tech, Inc. 2009, Susquehanna River Model Configuration, calibration and validation: Draft report. Prepared for Global Change Research Program, USEPA Office of Research and Development, Washington, D.C.

Thorndahl, S., Beven, K.J., Jensen, J.B., Schaarup-Jensen, K., 2008. Event based uncertainty assessment in urban drainage modelling, applying the GLUE methodology. Journal of Hydrology, 357(3-4): 421-437.

Tidwell, V.C., Kobos, P.H., Malczynski, L.A., Klise, G., Castillo, C.R., 2012. Exploring the water-thermoelectric power nexus. Journal of Water Resources Planning and Management, 138(5): 491-501.

Tidwell, V.C. 2013. Mapping water availability and cost in the western U.S. Presented at the U.S. E.P.A. Water Acquisition Workshop, held June 4, 2013, Washington D.C.

Uhlenbrook, S., Sieber, A., 2005. On the value of experimental data to reduce the prediction uncertainty of a process-oriented catchment model. Environmental Modelling & Software, 20(1): 19-32.

USEIA. 2012. Annual Energy Outlook 2012. Available online at www.eia.gov/forecasts/aeo.

USEPA, 2001. Fact sheet: Better assessment science integrating point and nonpoint sources (BASINS, version 3.0). EPA-823-F-01-0121. Washington, D.C.: USEPA Office of Water.

USEPA, 2002. Guidance for Quality Assurance Project Plans for Modeling. EPA QA/G-5M.

USEPA, 2007. Geospatial Metadata Technical Specification Version 1.0, Office of Information Collection Office of Environmental Information, Environmental Protection Agency, November 2007. http://www.epa.gov/geospatial/policies.html

USEPA, 2012. Quality Assurance Project Plan (QAPP): Modeling the Impact of Hydraulic Fracturing (HF) on Water Resources Based on Water Acquisition Scenarios, Prepared by the Cadmus Group, Inc, in conjunction with AQUA TERRA Consultants and Texas A&M University, Version 1.0 updated September 4, 2012, <http://www2.epa.gov/hfstudy/qapp-modeling-impact-hydraulic-fracturing-hf-water-resources-based-water-acquisition>

USGS, 2011. NLCD 2006 Land Cover.

USGS Marcellus Shale Assessment Team, 2011. Information relevant to the U.S. Geological Survey assessment of the Middle Devonian Shale of the Appalachian Basin Province, 2011. U.S. Geological Survey Open-File Report 2011–1298, 22 p.

USGS Uinta-Piceance Assessment Team, 2003, Petroleum systems and geologic assessment of oil and gas in the Uinta-Piceance Province, Utah and Colorado. U.S. Geological Survey Data Series 69-B.

Van Griensven, A., Breuer, L., Di Luzio, M., Vandenberghe, V., Goethals, P., Meixner, T., Arnold, J., and Srinivasan, R., 2006. Environmental and ecological hydroinformatics to support the implementation of the European Framework Directive for river basin management. Journal of Hydroinformatics, 8(4): 239-252.

Vidic, R.D., Brantley, S.L., Vandenbossche, J.M., Yoxtheimer, D., Abad, J.D., 2013. Impact of Shale Gas Development on Regional Water Quality. Science, 340(6134).

Watermark Numerical Computing, 2004. PEST model-independent parameter estimation, User Manual: 5th Edition.

Weston Solutions, 2012. Evaluation of Geology and Water Well Data Associated with the EPA Hydraulic Fracturing Retrospective Case Study, Bradford County, Pennsylvania, prepared for Chesapeake Energy, Report 41312.

Williams, John H., 1998. Hydrogeology and Groundwater Quality of the Glaciated Valleys of Bradford, Tioga, and Potter Counties, Pennsylvania, prepared for the US Geological Survey, Pennsylvania Geological Survey Water Resource Report 68.

Wimmer, F., Eisner, S., Flörke, M., 2011. Development and validation of the WaterGAP water resources and water use meta-models. In: Holman, I., Harrison, P. (Eds.), The CLIMSAVE Project: Climate Change Integrated Assessment Methodology for Cross-Sectoral Adaptation and Vulnerability in Europe. Cranfield University, Cranfield, UK, pp. 38-44.

Wöhling, T., Samaniego, L., Kumar, R., 2013. Evaluating multiple performance criteria to calibrate the distributed hydrological model of the upper Neckar catchment. Environmental Earth Sciences, 69(2): 453-468.

Appendix A. Project Area and Model Implementation

Quality Assurance Project Plan

Appendix A.

Project Area Information and Model Implementation and Calibration

.

Appendix A. Table of Contents

List of Figures—Appendix A

List of Tables—Appendix A

Overview

Appendix A to the QAPP for Modeling HF Scenarios Water Acquisition Phase 1 provides additional detail on the study project areas and the application of the hydrologic models in these areas. This appendix addresses portions of Task 1 that were briefly described in Section A5.

Section 1. Project Areas

The potential impacts of water acquisition for HF activities will be studied in two watersheds where hydraulic fracturing activity has increased in recent years. Phase 1 of the EPA study assessed relative impacts of hydraulic fracturing water withdrawals in two major river basins:

- Susquehanna River Basin (SRB) overlying the Marcellus Shale gas reservoir is located primarily in Pennsylvania and New York and represent humid eastern climate
- Upper Colorado River Basin (UCRB) in arid western Colorado overlies the Piceance structural basin and tight gas of the Williams Fork formation

These major basins were selected because of high current and projected rates of hydraulic fracturing activity over the next several decades, and because the EPA has previously calibrated and tested watershed models in these areas to investigate future climate change impacts on watershed hydrology the "20 watersheds study"(Johnson *et al*., 2012). Each of these study areas exceeds 18,000 square miles. Phase 1 was conducted by extramural contractors who completed a draft report in 2012 currently under review (Cadmus Group, 2012a,b). Most elements of this Phase 2 project build on the Phase 1 study and we utilize information from these projects to inform next steps in the Phase 2 refinement.

The Phase 2 study of HF water acquisition scenarios to be conducted by EPA/ORD will increase the spatial resolution of the water stress analysis by narrowing geographic scope within the major basins to watersheds less than 250 mi^2 located primarily in Bradford County within the Susquehanna basin and in focus watersheds that will be identified in the 2020 mi^2 area of Garfield County within the Upper Colorado basin.

Smaller project areas within the large basins were identified within the larger basins based on a set of selection criteria as follows.

We established lists of mandatory and desired attributes for study watersheds in the UCRB and SRB, such that results would be most relevant to the questions at hand and most transferrable to areas of high HF potential. Our mandatory criteria were:

- The watershed contains a currently-active, long-term (multi-decade) USGS streamflow gage at the outlet.
- A currently-active, long-term USGS groundwater monitoring well exists in the vicinity of the watershed.
- The watershed a) contains a currently-active, long-term NWS hourly station inside the watershed boundary, and/ or b) contains multiple, distributed stations within a 20 km radius of the watershed center.

- The watershed is located within an active HF region, where HF growth is forecasted to continue.
- The watershed contains no major surface water reservoirs.

Additional desired attributes were:

- The study watershed has one or more additional USGS gages, though not necessarily long-term or currently-active, within the main watershed boundary.
- The study watershed has numerous groundwater monitoring wells, including one within the watershed boundary.
- The study watershed contains a drinking water supply intake with detailed consumption data.
- The study watershed overlaps with external research projects.
- The study watershed is located within a "hot spot" of the current and projected regional HF activity, which we identified as Garfield County, CO in the UCRB and Bradford County, PA in the SRB.

While there were no sites that contained all desired attributes, we identified multiple candidate locations in both study areas meeting all of the mandatory site selection criteria. Our final site selection was based on which of the candidate watersheds offered the greatest number of desired attributes.

Next, we provide information on the project areas to the extent that we researched it for site selection or it was included in the Cadmus Group reports. Considerable more information will be acquired during the project.

Table A1. Comparison of Phase I and Phase 2 project area details.

Project 5B	Study Feature	Susquehanna River Basin (SRB)	Upper Colorado River Basin (UCRB)
Large Basin Study (Phase 1)	Total Basin Area	27,510 mi^2	17,800 mi^2
	Focal Area	Total basin	Total basin
	Project Team	Extramural (Cadmus + Aqua Terra)	Extramural (Cadmus + Texas A&M)
	Analytical Model	HSPF	SWAT
	Status	Draft Report in review Cadmus (2012a)	Draft Report in Review Cadmus (2012b)
Refined Analysis (Phase 2)	Focal Area	Bradford County—Towonda Creek	Garfield County
	Total Basin Area	215 mi^2	2021 mi^2
	Watershed size	Small headwaters to HUC 12	Small headwaters to HUC 12
	Temporal Scale	Short-term to annual	Short-term to annual
	Project Team	EPA-ORD	EPA-ORD
	Hydro Model	SWAT, HSPF	SWAT, HSPF
	Status	Start August 2013	Start August 2013

Section 1.1 Garfield County CO

Key information from the Phase I project in the UCRB is provided here. This project assessed the potential effects of HF for the Upper Colorado Basin using the SWAT model. Phase 2 of the project to be conducted by EPA/ORD/NERL/ERD will focus on Garfield County within the UCRB. The information in this section is synthesized from the Cadmus Group Phase 1 report produced by D. Deb, R. Srinivasan, and J. Shireman, Texas A & M University for the Cadmus Group, Draft Report submitted to EPA, Dec 2012.

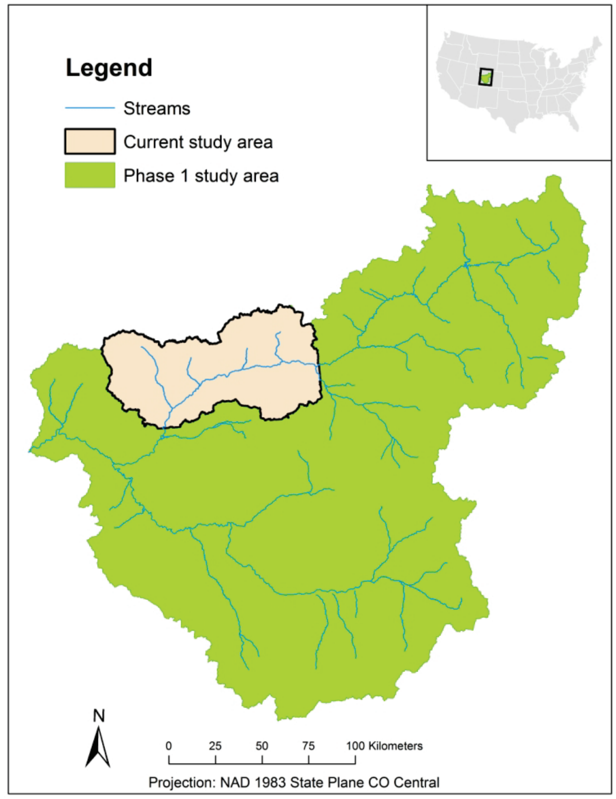


Figure A1. Garfield County project area within the Upper Colorado River Basin.

Colorado is one of the major states for natural gas production in the United States and nearly all active wells in Colorado are hydraulically fractured. The Upper Colorado River Basin (UCRB) encompasses about 17,800 square miles (Figure A1) upstream from the Colorado-Utah state line. The primary river in the basin, the Colorado River, originates in the mountains of central Colorado and flows about 230 miles southwest into Utah. The UCRB basin provides water supply, flood control and hydropower to a population of approximately 308,000 within the basin. Garfield County is located in the western portion of the UCRB with about 10% of the UCRB area (Figure A1). Topography in this area is high plateaus bordered by steep cliffs along valleys. Climate in the western portion of the UCRB is semiarid to arid. Precipitation in the Garfield County area ranges from about 25 to10 inches per year. The major use of water is irrigation, but trans-mountain diversions within the UCRB , especially in the high elevation alpine region in eastern portion of the basin, provide water to more than 1 million people in the eastern part of Colorado (outside of the study area).

Geology. The Upper Colorado River Basin is within Piceance Basin natural resources area (Figure A3). The Piceance Basin is a source of natural gas derived from tight gas and oil shale plays including the Williams Fork, and Mancos Formations. The Upper Cretaceous Williams Fork Fm., a thick section of shale, sandstone, and coal, has been recognized as a significant source of gas since 2004 (Kuuskraa and Ammer, 2004). The Piceance Basin was originally exploited for its coal resources. Gas plays in the Williams Fork are in lenticular and discontinuous alluvial sand bodies occurring mainly in the lower portion of the Formation were coal horizons are common. The Williams Fork is underlain by the Iles Formation of which such as the Rolins, Cozzette, and Corcoran members are also gas producing and are often penetrated and developed with the Williams Fork for production in a single well (Rojas, 2008). The Iles overlies the Mancos Shale, a marine shale with a gradational interbedding transition between the two (Hettinger and Kirschbaum, 2002). The discontinuous sands in the Williams Fork have lateral dimensions of approximately 1,000 feet, but may be offset and stacked, therefore exploitation of the Williams Fork for gas production requires a dense well spacing (Pranter and Sommer, 2011), and wells are drilled either vertically or directionally rather than horizontally.

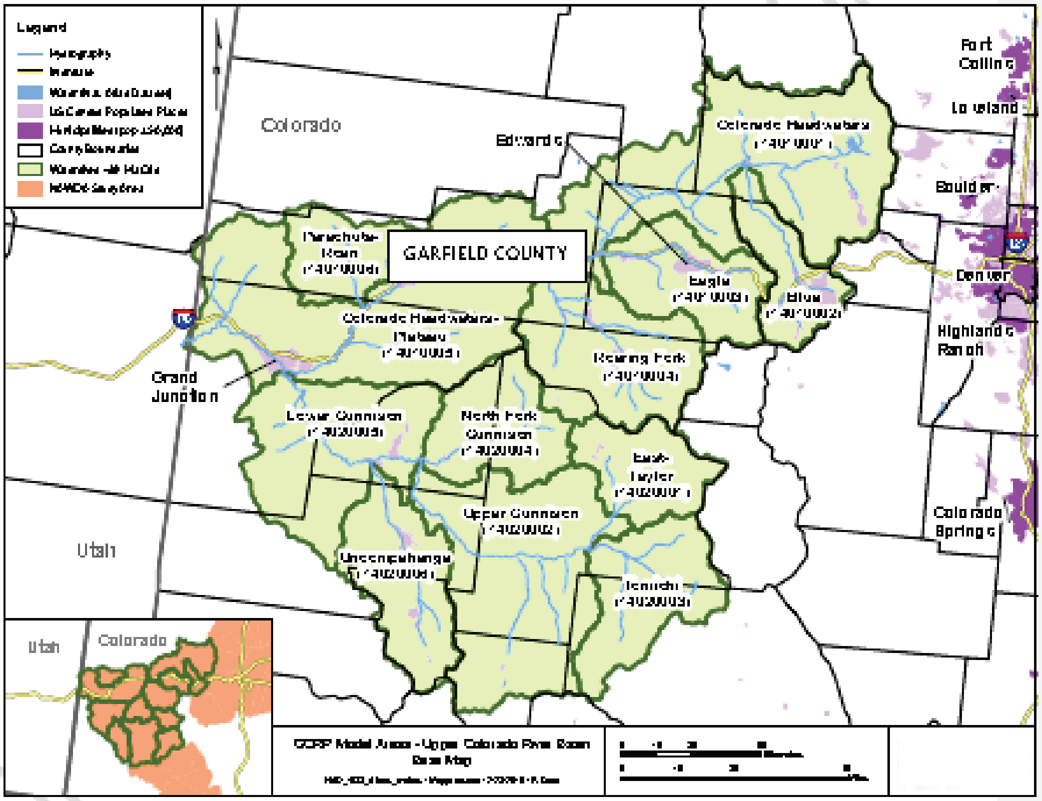


Figure A2. Location of the Upper Colorado River basin.

The Mancos and Niobrara shale gas plays flank the Piceance Basin; and the units are geologically correlative (Kent, 1968). Mancos occurs in western Colorado and is considered the second largest producer of shale gas in the Rocky Mountains (Brathwaite, 2009). The Cretaceous Mancos and Niobrara Shale Formations are continuous shale units which can be developed using conventional well spacing and horizontal drilling methods, similar to development of the Marcellus Shale in the Susquehanna River Basing in Central Pennsylvania and New York states.

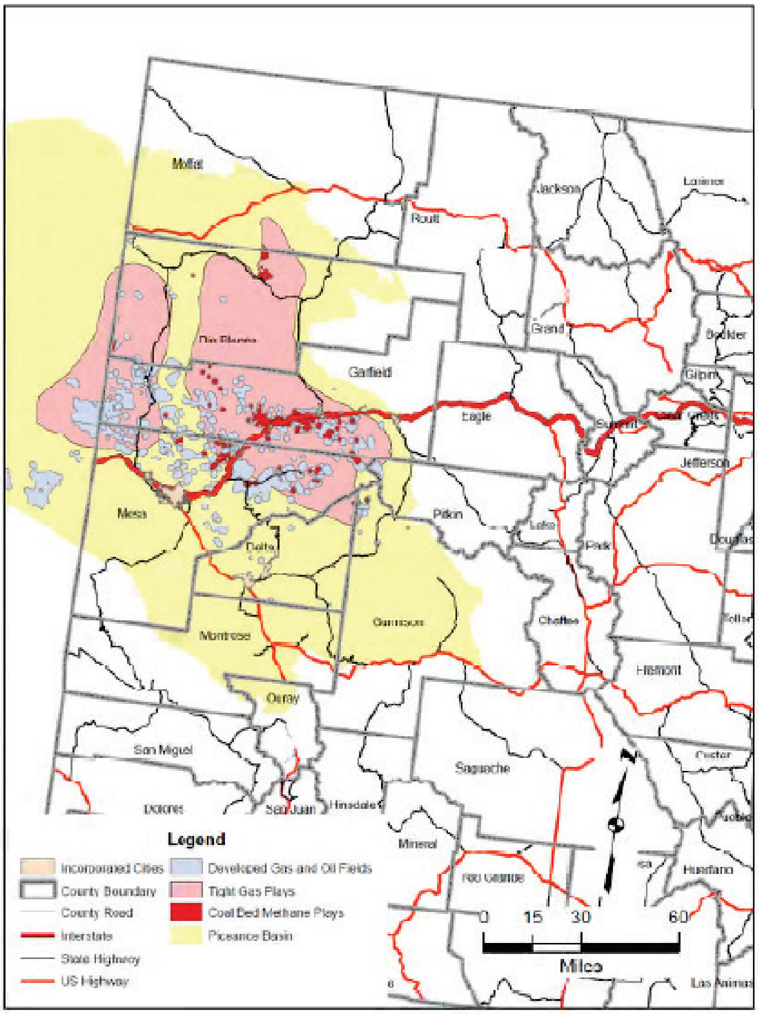


Figure A3. Oil and gas plays in the Piceance Basin in the Upper Colorado Basin (yellow). Within the Piceance are tight gas plays (pink) that require close well spacing.

Garfield County is one of the most intensively developed areas (gray).

Figure A4. High aerial view of the western region of the UCRB and Garfield County (faint county line is shown). Image from Google Earth.

The hydrogeology of aquifers of Garfield County

The aquifers of significance in Garfield County are beneath the Colorado Plateau geographic province and include additional alluvial aquifers associated with the Colorado River. See Figures A5 and A6. From shallowest to deepest, the aquifers include the Unita-Animas aquifer, the Mesaverde aquifer, and the Dakota-Glen Canyon aquifer system. Water quality ranges from 1,000 mg/L dissolved solids in the Unita-Animas to over 35,000 mg/L in the Glen Canyon.

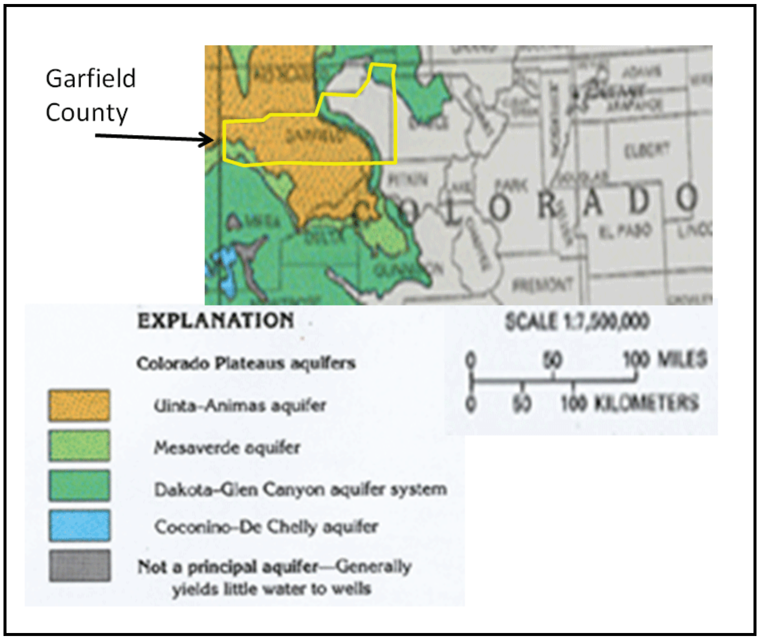


Figure A5. The bedrock aquifers of Garfield County, Colorado (from Robson and Banta, 1995).

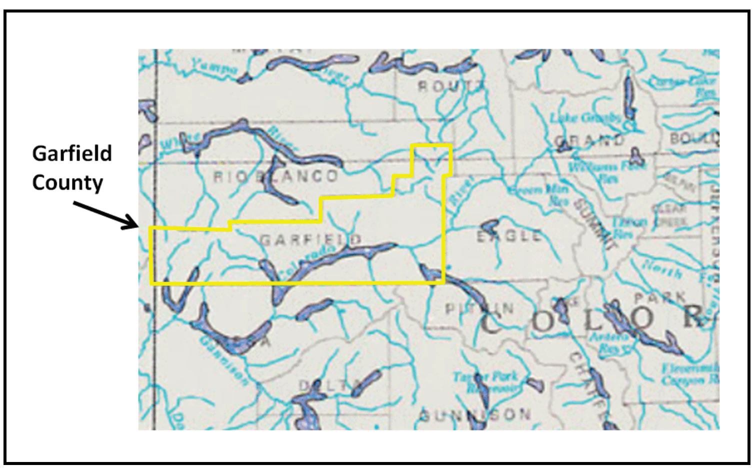
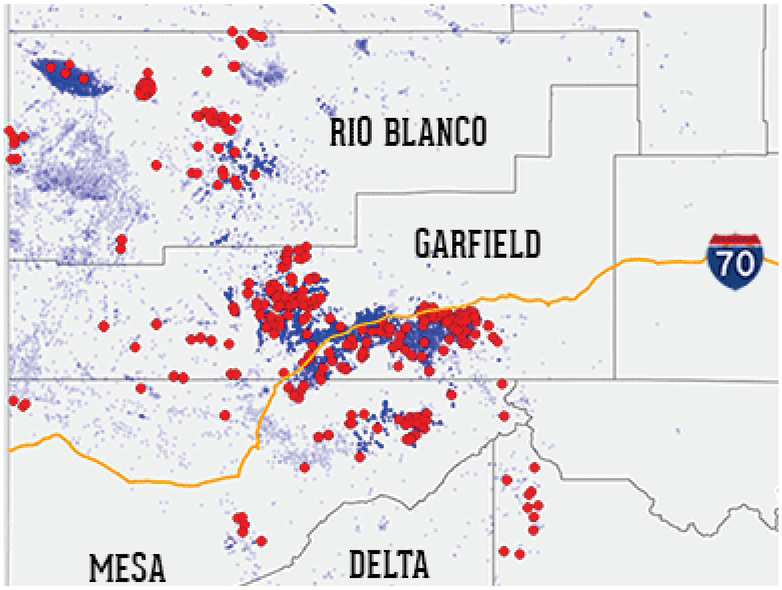


Figure A6. The alluvial aquifers for Garfield County, Colorado (from Robson and Banta, 1995).

Figure A7. Garfield County well locations. Data from COGCC.



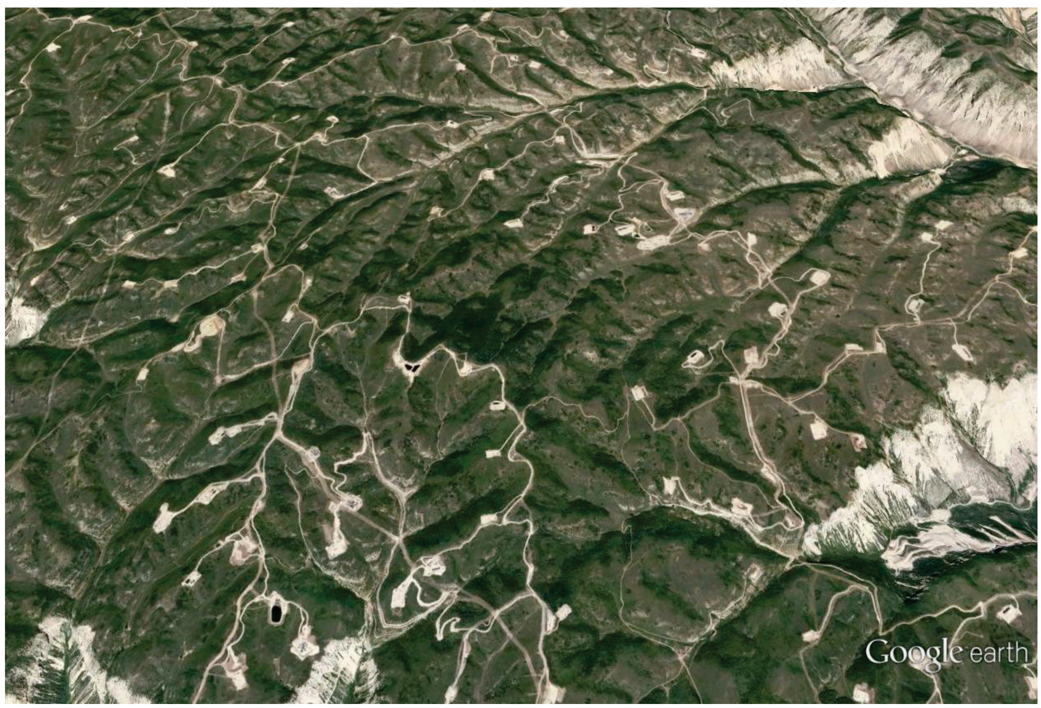


Figure A8. Lower aerial imagery of developing field within Garfield County. Image from Google Earth.

Baseline Information and Sources of Information from UCRB

Hydraulic Fracturing Activity. HF well information is available from the Colorado Oil and Gas Conservation Commission (COGCC). They maintain a number of on-line databases that provide spatial data, some construction details, production data, and well status of the oil and gas wells in the State of Colorado. The well data source is the wells completion table included with the production records Access™ databases (http://cogcc.state.co.us/ >Library> Production and Prices).

Well development in Garfield County has been increasing since 2000 and after the economic downturn curtailed drilling activity in 2008 (Figure A9). Projected maximum drilling rate for future scenarios was estimated by the Cadmus Group (2012b) to be 2108 wells per year. Final well density is expected to be as low as 1 per 20 acres and as high as 1 per 80 acres.

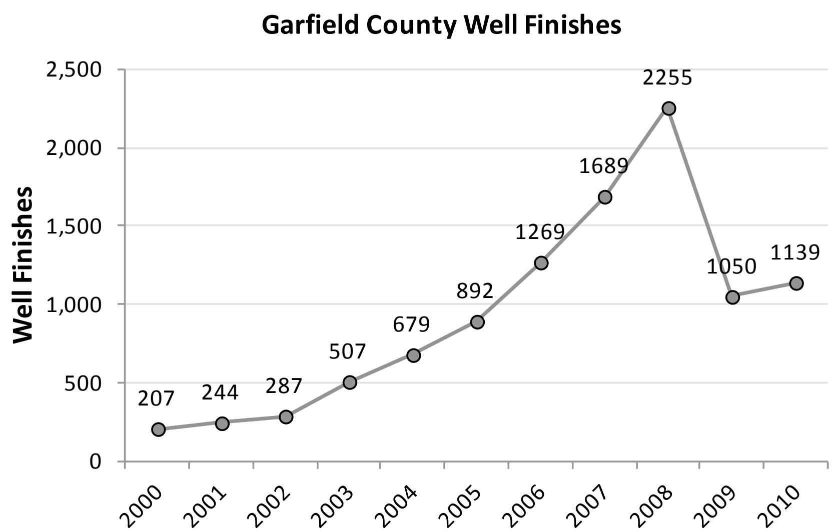


Figure A9. Well finishes in Garfield County. Data from COGCC.

Operators within the UCRB include:

- UCRB
- Berry Petroleum,
- Williams Production RMT (now WPX),
- EnCana Oil and Gas

The Bureau of Land Management, Colorado River Valley Field Office provides oversight to activities on BLM lands.

Water use estimates per well were developed by the Cadmus Group (2012b) as described in QAPP Section A5 and shown in Table A2. These numbers for well use will be checked in Phase 2.

Table A2. HF Water use characteristics determined by the Cadmus Group (2012b)

	Characteristic	Upper Colorado River Basin
Water Use per Well	Total volume/well (gals)	0.25 million
	Actual new water volume/well (gals)	0.18 million[&]
	Average recycling %	100%
	Green Technology Volume (gals)	0.18 million

Water Use

Water use information can be obtained from:

(http://co.water.usgs.gov/nawqa/ucol/pdf/circ1214.pdf).

The dominant use of water within Garfield County is irrigation (Figure A7). There are public supplies within Garfield County.

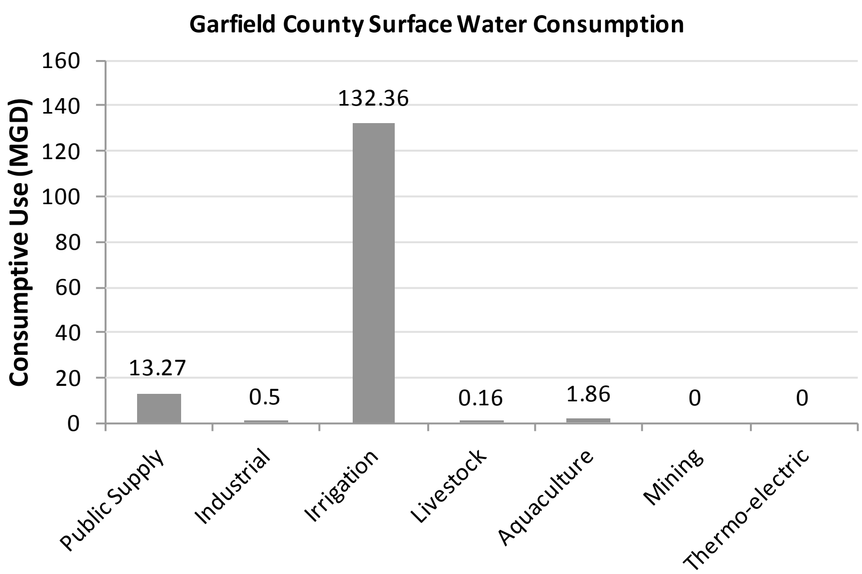


Figure A10. Consumptive water use for Garfield County (from Cadmus Group, 2012b).

Water rights structure spatial data is included in an online geodatabase at Colorado's Decision Support System (CDSS) site: http://cdss.state.co.us/GIS/Pages/GISDataHome.aspx by Water Division and structure type (diversions and wells). To obtain the authorized use type of a Water Right structure, the CDSS one line tools search tools must be used to query by Division or District.

Public water supply use in Garfield County is projected to increase by 71% in the HF futuring "energy plus" scenario.

Municipal use locations should be queried for HF use because it is known that municipalities in the Upper Colorado River Basin have sold water to Oil and Gas Operators for drilling and completion operations (Cadmus Group, 2012b).

Key Water Stress Findings from Phase Water Acquisition Modeling Study

Results of Phase 1 HF scenario analysis were primarily reported for the entire UCRB (Cadmus Group 2012b). However, the Phase 1 report also provided some information at a subbasin level (e.g. Figure A11. Although the project found no impacts from HF water acquisition when aggregated for the entire UCRB, impacts were identified at the subbasin level, especially in Garfield County. Figure A9 shows one example of subbasins flagged as particularly high impact.

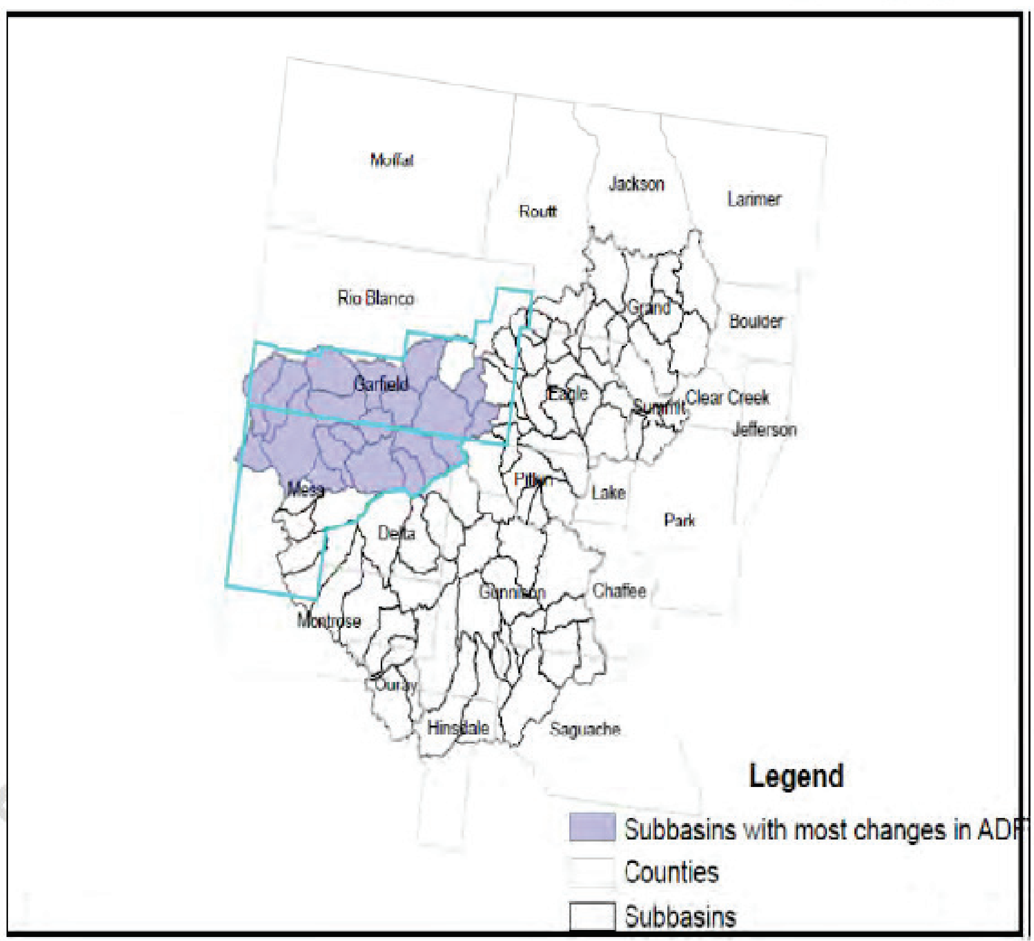
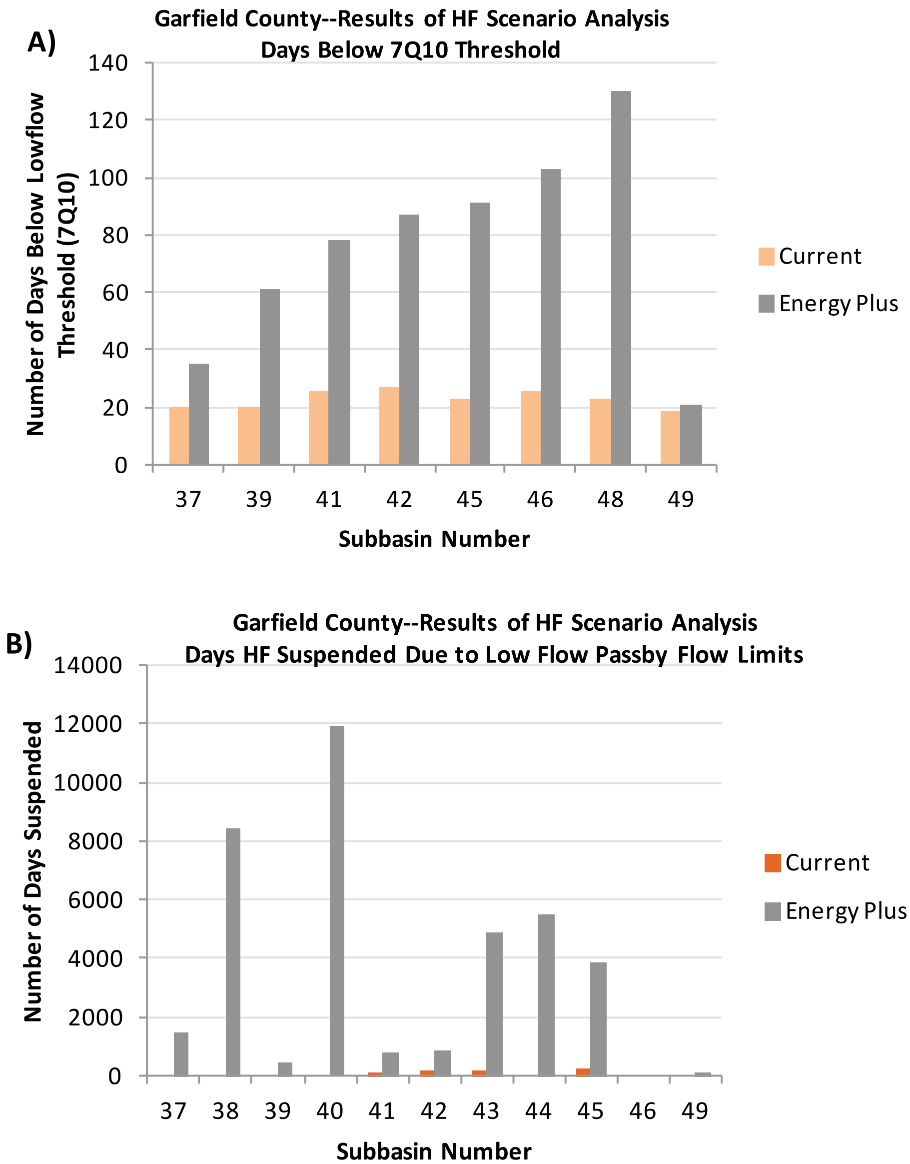


Figure A11. Location of subbasins within the UCRB with the most significant changes to Annual Daily Flow during the futures simulation as an example of scenario analysis impacts at the subbasin scale .

A variety of flow metrics were analyzed in the scenario analysis. Figure 12 shows the change in several low flow related conditions in the simulated in the "current" scenario and in the "energy plus" scenario in subbasins located in Garfield County. Figure 12A is the number of days less than a 7Q10 threshold and 12B is the number of days HF withdrawals would be suspended due to passby flows. Low flow metrics were particularly impacted in the "energy" plus scenarios. Several Not all subbasins were equally affected for all metrics.

Figure A12. Impact of HF scenario analysis on subbasins of the UCRB located in Garfield County.



Section 1.2 Bradford County PA

Key information for the Susquehanna River Basin (SRB) from the Phase 1 project is provided here. This project assessed the potential effects of HF for the Susquehanna River Basin using the HSPF model. Phase 2 of the project to be conducted by EPA/ORD/NERL/ERD will focus on Bradford County within the Susquehanna River Basin. This information is synthesized from the Cadmus Phase 1 report produced by Aqua Terra for the Cadmus Group, Draft Report submitted to EPA, Dec 2012 a. In turn, that report extracted, edited and supplemented information from Tetra Tech *et al*., 2009:

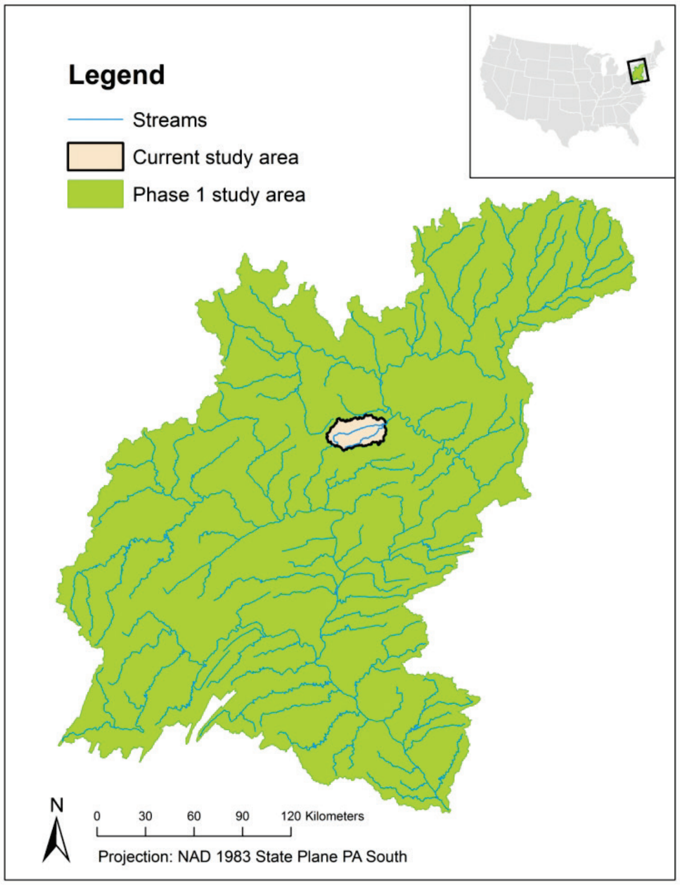


Figure A13. Location of Towanda Creek with the Susquehanna River Basin

The Susquehanna River drains about 27,000 square miles (mi^2) of New York, Pennsylvania, and Maryland. The watershed makes up 43 percent of the Chesapeake Bay's drainage area and consists of six major sub-watersheds (Chemung, Upper Susquehanna, West Branch Susquehanna, Middle Susquehanna, Juniata, and Lower Susquehanna). The Susquehanna River flows about 444 miles from its headwaters at Otsego Lake in Cooperstown, New York to Havre de Grace, Maryland, where the river flows into the Chesapeake Bay (Figure A13 and A14). The river is the largest tributary to the Chesapeake Bay, providing 50 percent of its freshwater flows (SRBC, 2008).

Geology. The Marcellus Shale is a black shale of Middle Devonian age that underlies much of Pennsylvania, New York, Ohio, West Virginia and adjacent states. Geologists have long known that the Marcellus contains natural gas, however, the depth of the rock unit and its low permeability made the Marcellus an unconventional exploration target. Within the past few years hydraulic fracturing and horizontal drilling, have been tested in the Marcellus resulting in some of the most productive wells in the eastern United States. These developments triggered an explosion of drilling and leasing activity in the areas above this rock unit since 2008.

Figure A14. Susquehanna River Basin

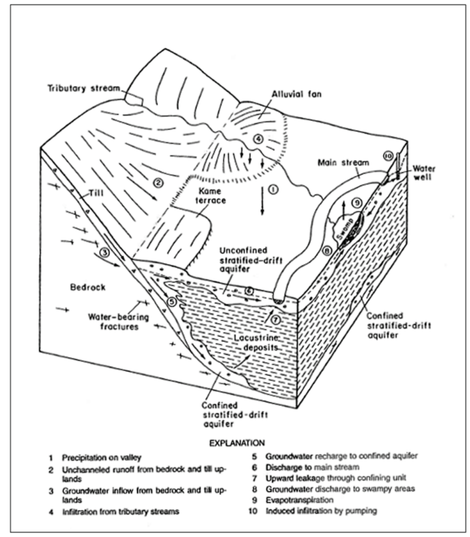
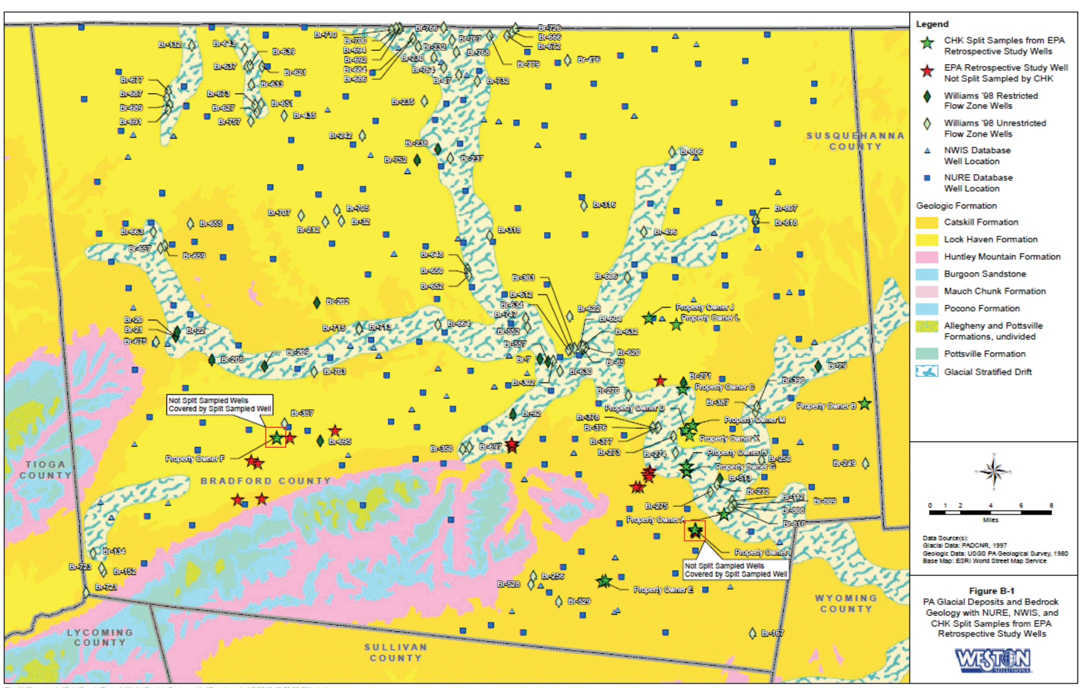


Figure A15. Generalized hydrogeology of the stratified-drift aquifers (Williams, 1998). The USGS maintains a long-term observation well in Bradford County that measures water levels in the Lock Haven Formation (May 1966 to current year) .

Figure A16. Bedrock geology of Bradford County (Weston Solutions, 2012)



Baseline Information and Sources of Information from SRB.

Hydraulic Fracturing Activities

Hydraulic fracturing has been increasing in most of the Susquehanna River basin since 2008 (Figure A17).

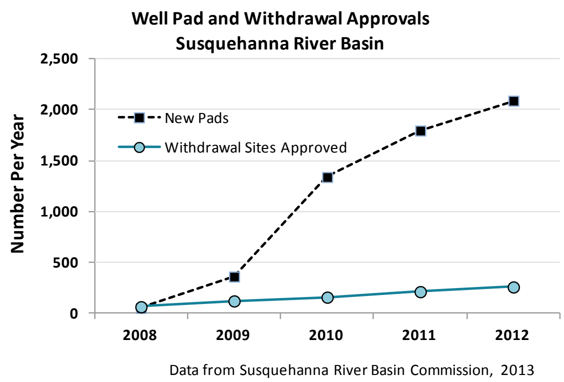


Figure A17. Well pad starts in the entire Susquehanna River Basin.

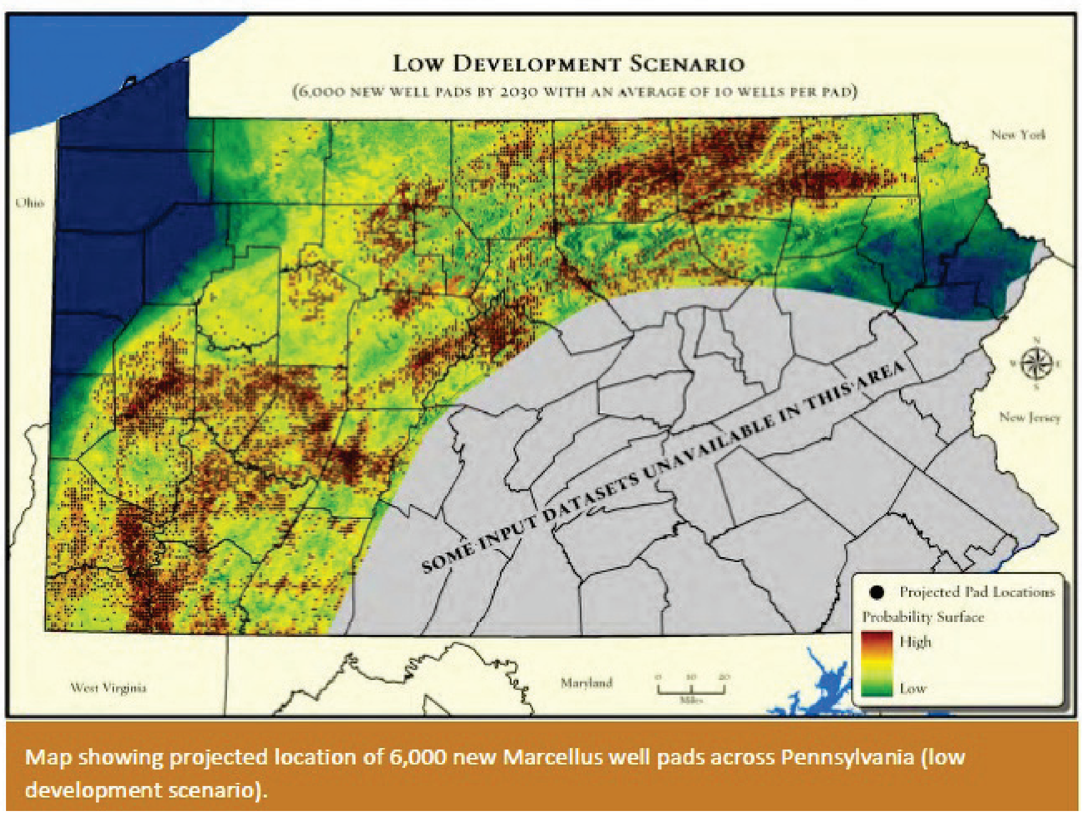


Figure A18. Projected well pad density and distribution for low development scenario reported in Cadmus Group, 2012a. Bradford County is located in the Northeast corner of Pennsylvania bordering New York.

Table 3. Water use for HF activities determined in Phase I for the entire SRB area.

	Characteristic	Susquehanna River Basin
Water Use per Well	Total volume/well (gals)	4.04 million
	Actual new water volume/well (gals)	3.5 million
	Average recycling %	13%
	Green Technology Volume (gals)	2.9 million
Drilling Rate per Year	Business as usual (current rate)	2370
	Energy plus (maximum yearly rate)	2840
	”Recycling Plus”	2370
Well density	Drainage area per well, acres: Ave (High-end)	149 (80)

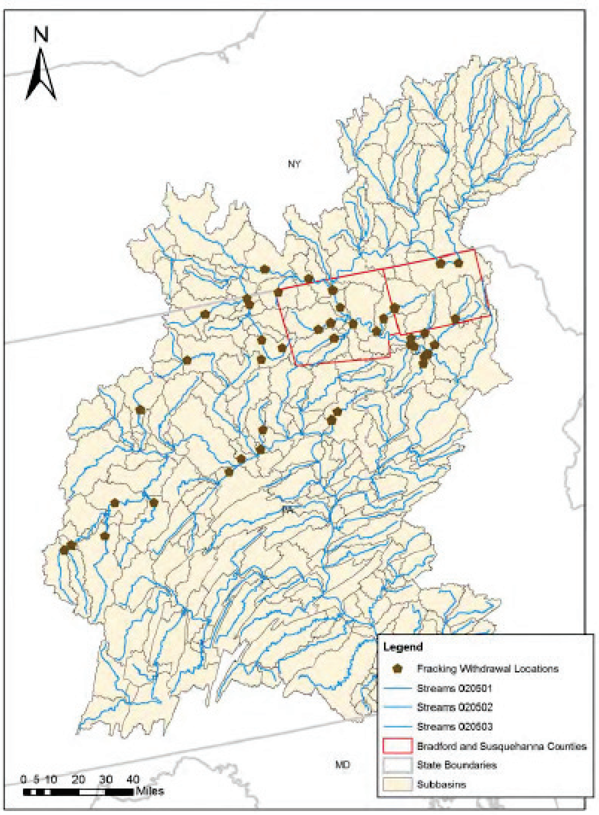


Figure A19 HF 2010 withdrawal sites in Bradford County (shown in red outline.

The SRB water acquisition data base can be mined for for facilities with NAICS =211111 (Crude Petroleum and Gas Extraction industry sector). For Bradford County, 47 geo-referenced, daily time interval withdrawal records were identified. The source of all withdrawals was determined as surface water. HF withdrawal points in Bradford County identified in Phase 1 are shown in Figures A19 and A20. HF operations are known to purchase water from other suppliers.

Water Use.

Data on water use is available from:

- PA DEP PWS data base
- SBR water acquisition data base (PASDA).

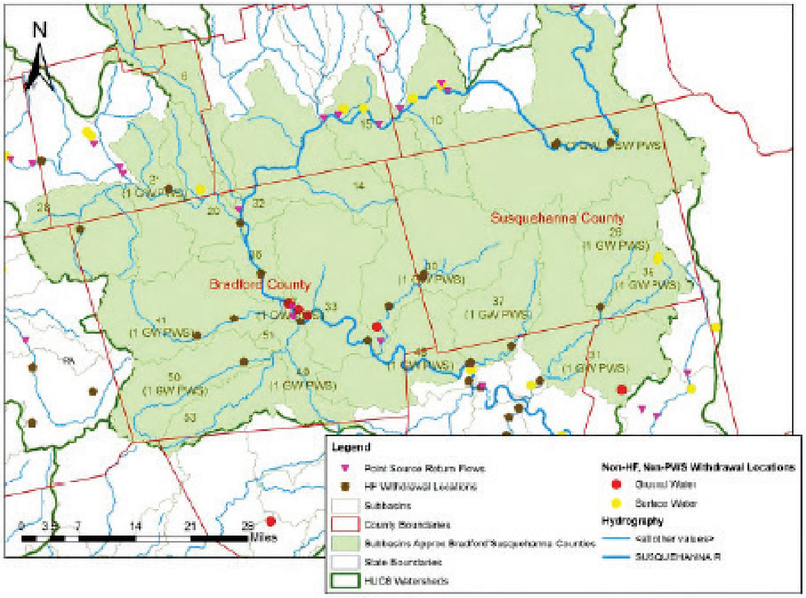


Figure A20. HF water withdrawal points in Bradford County in 2010.

HF Scenario Key Findings

The Phase 1 project team performed a zoomed in analysis on Bradford County during scenario testing. These efforts provided graphics but no data could be mined from the Cadmus report (Cadmus Group, 2012a) report to summarize scenario analysis at a more local level as was possible for Garfield County.

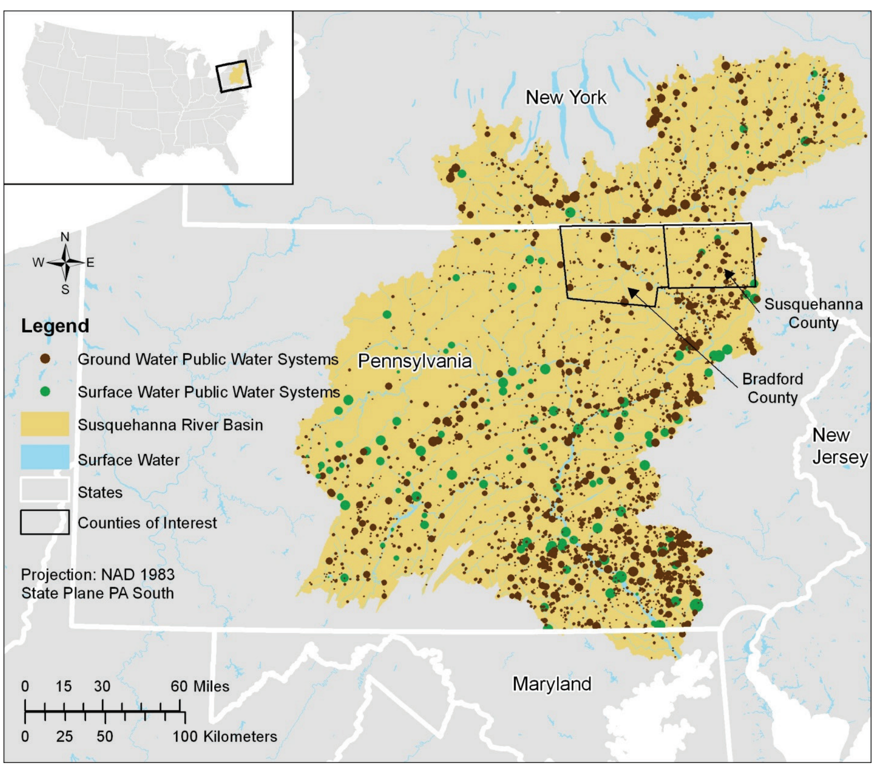


Figure A21. Public withdrawal locations in the SRB. Bradford County is outlined.

Figure A22. HF Withdrawal locations experiencing closure to maintain passby flows.

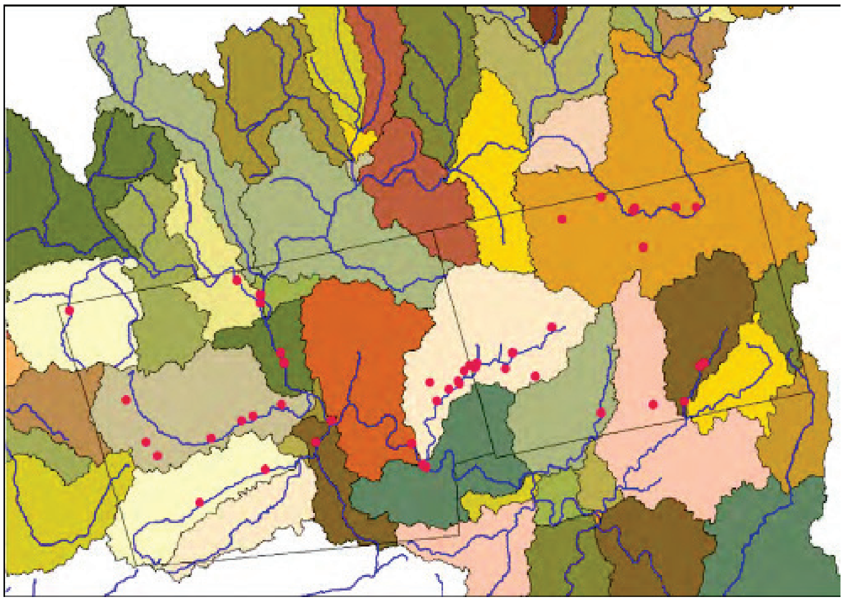


Figure A23. High level aerial view of HF pads distributed in rural area of Bradford County. Image from Google Earth.

Section 2. Hydrologic Modeling

Hydrologic modeling is the primary activity to determine water availability in the project areas in Task 1. The refined analysis of Phase 2 will use watershed and complementary groundwater models to complete the study design for Task 1 described in Section A5. Two watershed models will be applied in each of the two project areas, along with a groundwater model in the calibration step. In this section of Appendix A, we provide a more detailed description of the models, their application within the project areas and the methods of calibration and validation. Data acquisition was discussed in detail in Section B9 of the QAPP, and is briefly summarized here.

Appendix A includes a discussion of:

- Data acquisition
- SWAT
- HSPF
- Groundwater Modeling with GFLOW
- Calibration and Validation

Section 2.1 Data acquisition

The first step of watershed modeling requires obtaining the necessary spatial and time series data bases to run the models. Both HSPF and SWAT are spatially distributed models that have most of the same data requirements. Geospatial and time series data will be acquired for both study areas from sources provided in Section B9 and managed according to section B10 of this QAPP.

The models manage data input, overlay spatial data (Figure 24) and integrate time series data for hydrological simulations. Data needs are listed in Table A4 and A5.

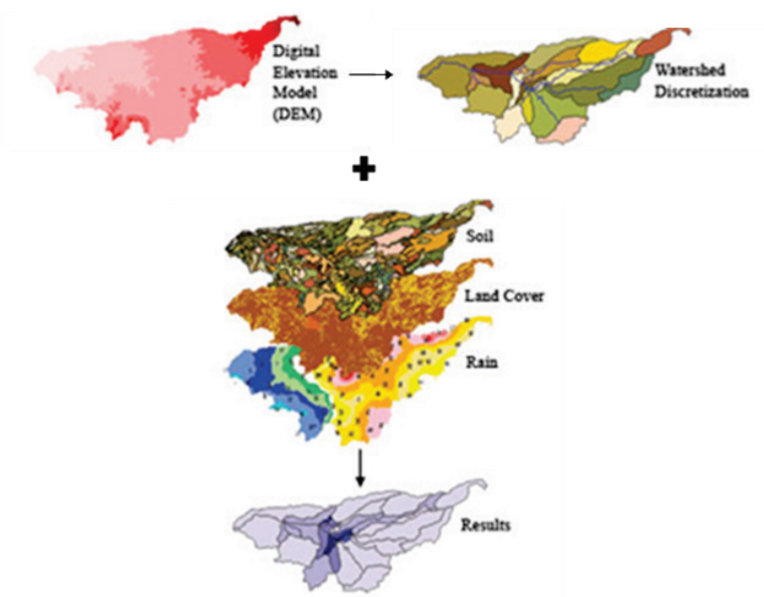


Figure A24. Schematic of data input layers used in SWAT and HSPF

Table A4. Data requirements for models. See Section B for information on source locations and management of data for project areas.

Data Category	Specific Data Needs
Geospatial Data	• Hydrography • Land cover raster • Current land use • Digital elevation models (10-m DEMs) • Detailed and generalized soil polygons • High- and medium-resolution hydrography • 12-digit hydrologic units • Digital topographic maps (current and historical) • Political boundaries: state, county • Locations of roads, pipelines
Meteorology (NOAA) (Time Series)	• Precipitation, • Air temperature, • Evapotranspiration
Hydrologic (USGS) (Time Series)	• USGS Flow records • Groundwater levels from wells • Water storage data
Supplemental to Hydrologic Models	• Current HF well locations • Census and parcel data • Projected population data for 'future' simulations

Data retrieval sources are summarized here.

Table A5. Source of geospatial, meteorology and hydrology data. See Section B9 for additional source information.

Item	Retrieval Source
DEM	http://nationalmap.gov/viewer.html
Soil data	http://soils.usda.gov/survey/geography/SSURGO/description_statsgo2.html
Land cover data	http://www.mrlc.gov/nlcd06_data.php
Meteorological (rainfall and ambient temperature) data	http://www.ncdc.noaa.gov/data-access/land-based-station-data/find-station
Gauge station flow data	http://waterdata.usgs.gov/nwis/rt

Garfield County , Colorado Hydrology and Met Data

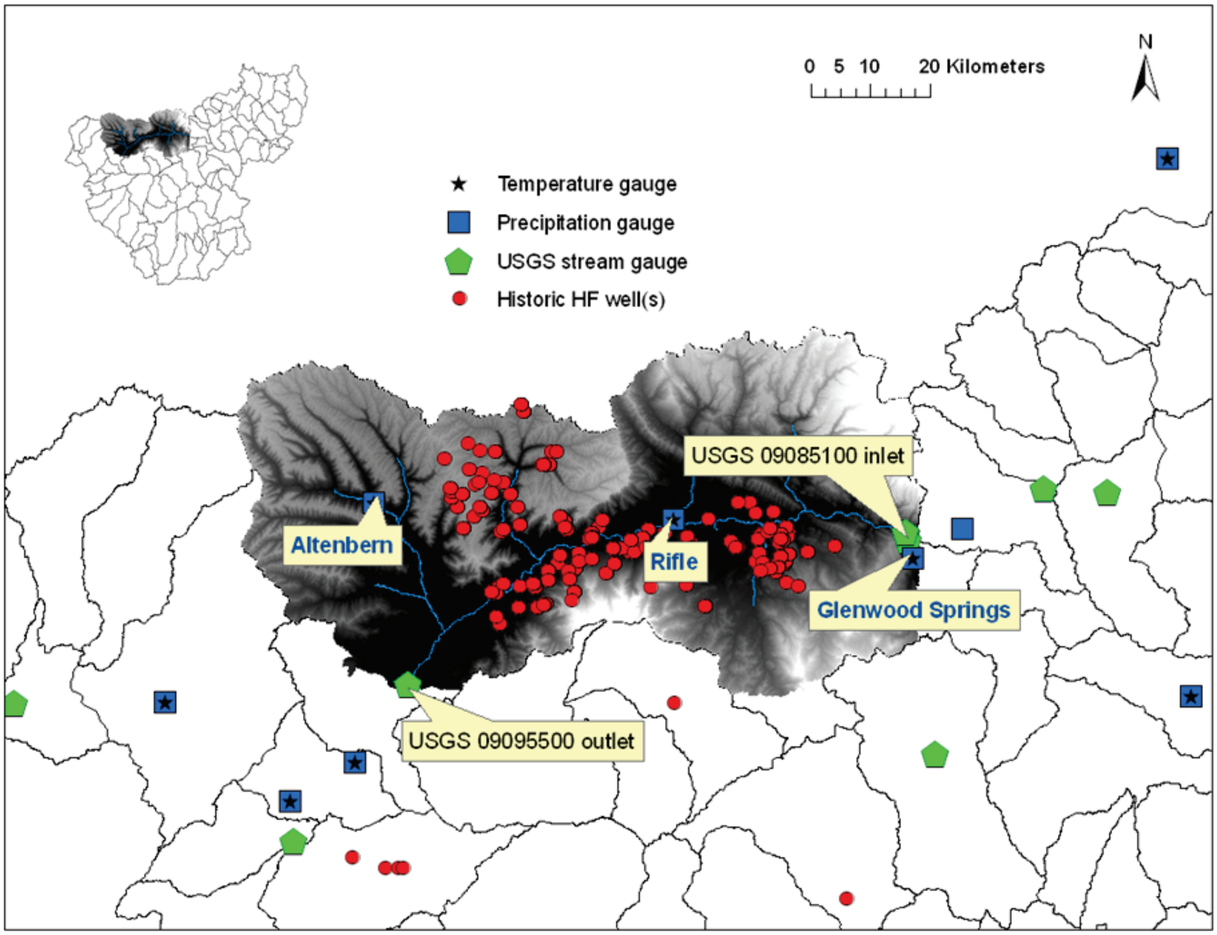
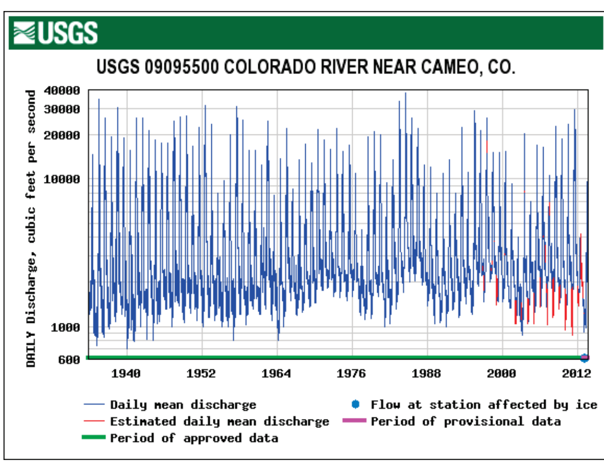


Figure A25 The Garfield County watershed within the Upper Colorado River Basin with location of available data.



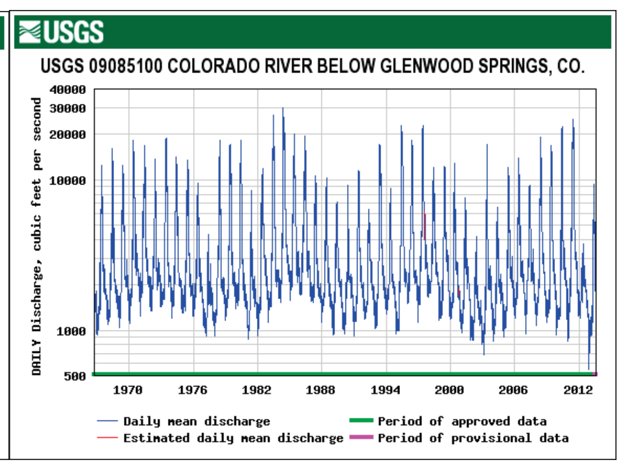


Figure A 26. USGS gauge station data for the periods of record at sites in Garfield County.

Table A6. Surface flow, precipitation and ambient temperature records for each station in Garfield County, CO.

NOAA, NCDC Station	Daily Precipitation (mm)	Daily Temperature (min, max)
"Altenbern": GHCND:USC00050214	<1960 to Present	<1960 to Present
"Rifle": GHCND:USC00057031	<1960 to 2009	<1960 to 2009
"Rifle": GHCND:USR0000CRIF	none	1984 to Present
"Glenwood Springs": GHCND:USC00053359	<1960 to Present	<1960 to Present
USGS Gauge Station	**Flow (cfs)**	
USGS 09095500	1933 to Present	
USGS 09085100	1966 to Present	

Bradford County/Towanda Creek Pennsylvania Hydrology and Met Data

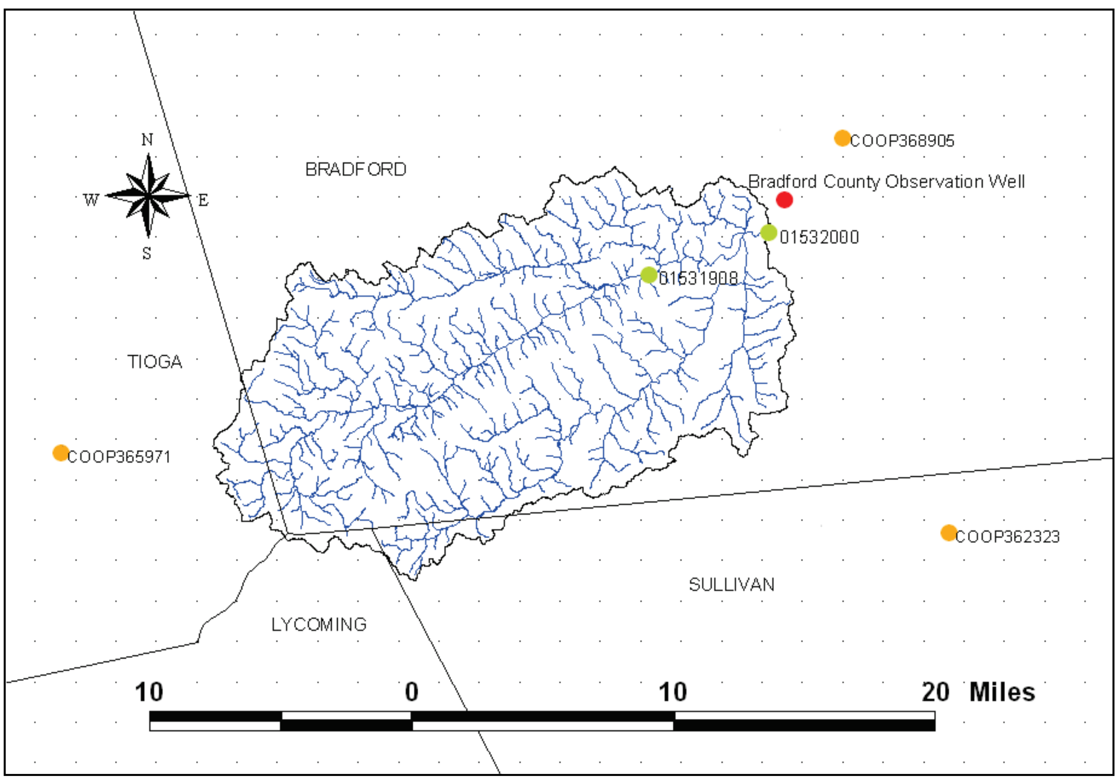
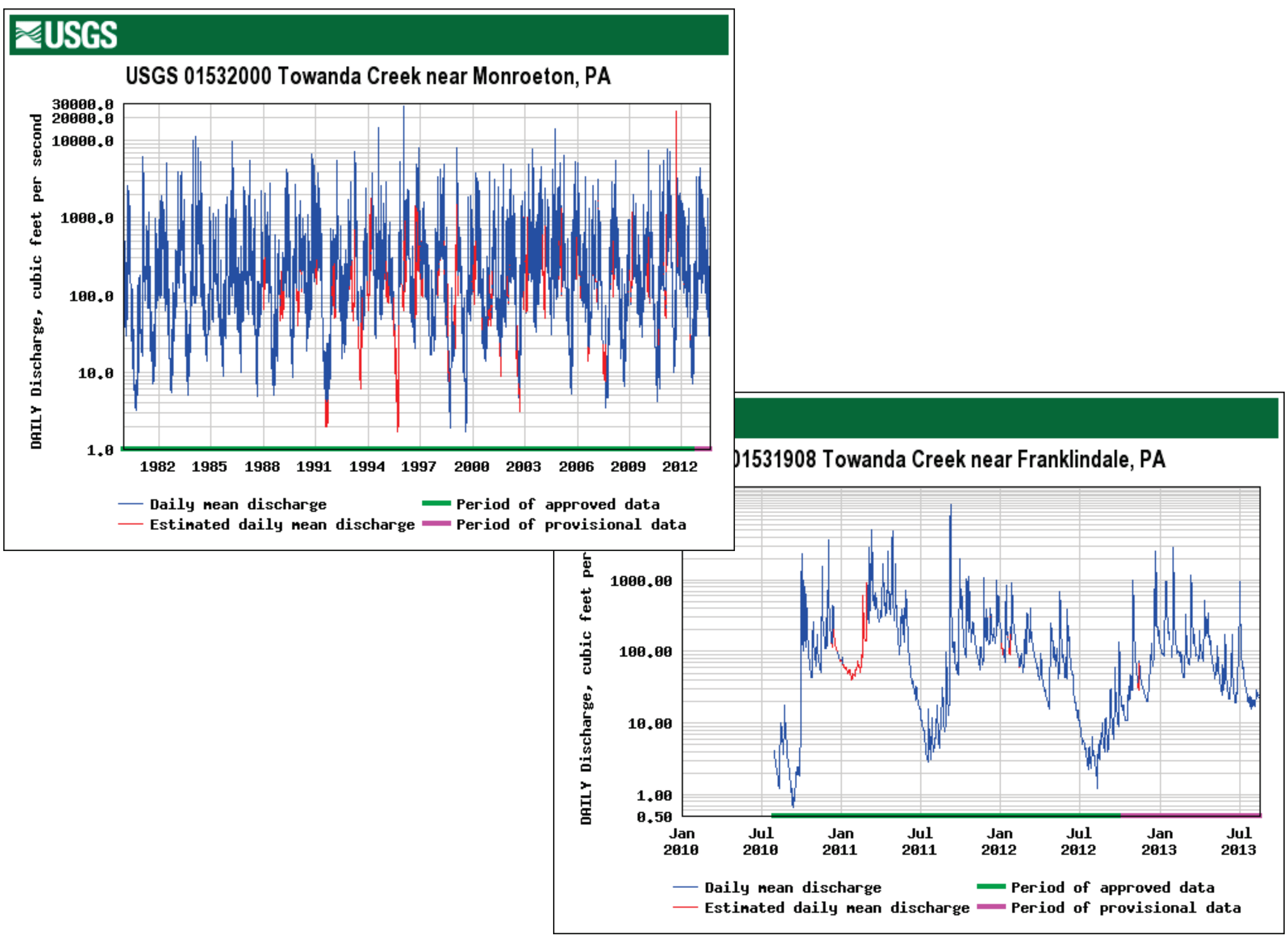


Figure A 27. The Towanda Creek watershed within Bradford Countywith location of available data.

Table A7. Surface flow, precipitation and ambient temperature records in Towanda Creek.

NOAA, NCDC Station	Daily Precipitation	Daily Temperature (min, max)
"Towanda 1 S": GHCND:USC00368905	<1960 to Present	<1960 to Present
"Troy 1 NE": GHCND:USC00368959	1986 to 2011	NO DATA
"Dushore": GHCND:USC00362323	2003 to Present	2003 to Present
"Canton": GHCND:USC00361212	1976 to Present	1976 to Present
"LaPorte": GHCND:USC00364815	1991 to Present	1991 to Present
USGS Gauge Station	**Daily Flow**	
USGS 01532000	1914 to Present	
USGS 01531908	2010 to Present	

Figure A28 The Bradford County study watershed Towanda Creek with location of available data.



Section 2.2 SWAT Model

The Soil and Water Assessment Tool was selected as one of the watershed models to evaluate the impact of hydrofracking on drinking water resources. SWAT is a basin-scale, continuous-time model that operates on a daily time step and is designed to predict the impact of management on water, sediment, and agricultural chemical yields in ungauged watersheds. The SWAT model is physically based, computationally efficient, and capable of continuous simulation over long time periods. SWAT was chosen because of its wide user base for hydrologic simulations and its appropriate application for the selected watersheds which are largely agricultural and forest-based. SWAT performs calculations on a daily time step.

SWAT Description. SWAT is a semi-empirical, semi-distributed model which can be operated through an ArcGIS interface (ArcSWAT version 2009.93.5). In SWAT, larger watersheds are separated into sub-basins and "hydrologic response units" (HRUs). HRUs are land surface areas that contain the same combination of soil, land cover and slope class (Neitsch *et al*., 2005). The main components of SWAT are: climate, hydrology, land cover/plant growth, with capability to predict erosion, and channel routing of nutrients, pesticides, and to simulate effects of management practices.

SWAT was designed to determine impacts of climate and land management on water supply and water quality (Krysanova and Arnold, 2008). The Soil Conservation Services (SCS) Curve Number is used to predict runoff and infiltration; the SCS Curve Number method is a function of soil permeability, land cover and antecedent soil water conditions.

In SWAT, a watershed is divided into multiple subwatersheds, which are then further subdivided into hydrologic response units (HRUs) that consist of homogeneous land use, management, and soil characteristics. The HRUs represent percentages of the subwatershed area and are not identified spatially within a SWAT simulation. Alternatively, a watershed can be subdivided into only subwatersheds that are characterized by dominant land use, soil type, and management.

Climatic inputs used in SWAT include daily precipitation, maximum and minimum temperature, solar radiation data, relative humidity, and wind speed data, which can be input from measured records and/or generated. Customized climatic input data options include: (1) simulation of up to ten elevation bands to account for orographic precipitation and/or for snowmelt calculations, (2) adjustments to climate inputs to simulate climate change, and (3) forecasting of future weather patterns.

Important features of SWAT model structure include the following:

1. SWAT is a watershed-scale, continuous-time model that operates on a daily time step and capable of continuous simulations over long time periods.
2. In order to adequately simulate hydrologic processes in a basin, the basin is divided into subbasins through which streams are routed. The subunits of the subbasins are referred to as hydrologic response units (HRUs) which are the unique combination of soil and land use characteristics and are considered to be hydrologically homogeneous. The model calculations are performed on a HRU basis and flow and water quality variables are routed from HRU to subbasin and subsequently to the watershed outlet.

3. The SWAT model simulates hydrology as a two-component system, comprised of land hydrology and channel hydrology. The land portion of the hydrologic cycle is based on a water mass balance.
4. Within each HRU, the major hydrological processes simulated by SWAT include canopy interception of precipitation, evapotranspiration, infiltration, surface runoff, lateral flow (subsurface flow), return flow (shallow ground water flow or baseflow), soil moisture redistribution, and percolation to deep aquifer.
5. Water enters the SWAT model's watershed system boundary predominantly in the form of precipitation. Precipitation is partitioned into different water pathways depending on system characteristics.
6. The water balance of each HRU in the watershed contains four storage volumes: snow, the soil profile (0-2 m), the shallow aquifer (2-20 m) and the deep aquifer (>20 m). The soil profile can contain several layers.
7. The soil-water processes include infiltration, percolation, evaporation, plant uptake, and lateral flow. Surface runoff is estimated using the SCS curve number or the Green-Ampt infiltration equation. Percolation is modeled with a layered storage routing technique combined with a crack flow model. Potential evaporation can be calculated using Hargreaves, Priestly-Taylor or Penman-Monteith method (Arnold et al., 1998).

SWAT Pre-Processing. The following provides an abbreviated description of the SWAT parameterization process.

The first step is to load a DEM. The DEM provides land surface slope information which defines model system flow paths.

Following DEM loading, the user initiates an automated watershed delineation function which will define the boundaries of the catchment and sub-basin outlets/pour points. Once the catchment has been delineated, the user overlays the soil layer file and the land cover file.

After the soil and land cover layers, the user initiates an automated procedure that defines the hydrologic response units (HRU) within the watershed. An HRU is a computational unit that is derived from areas that contain the same land cover, soil type and surface slope. HRUs will be developed for land covers, soil and surface slopes that meet a minimum area threshold. A threshold scheme is used to neglect HRUs with small areas since they typically have negligible impact on hydrologic simulation and can slow computation time. A commonly used threshold scheme is 5% for surface slope, 10% for land cover and 5% for soil cover.

Finally, the user loads daily meteorological data for multiple stations (typically NCDC stations) that surround the watershed. After loading the meteorological data, an automated gauge selection procedure in SWAT selects the closed station(s) to the watershed.

SWAT Sensitivity Analysis. After model pre-processing has been completed, the uncalibrated model setup will be used for identification of sensitive parameters. SWAT streamflow calibration allows adjustment of 27 parameters (Table A8), but we will ultimately reduce these to a small subset, to minimize problems of overparameterization (Beven, 2006; Matott *et al.*, 2009).

We will first define the allowable range for each parameter, all of which have a physical meaning and will be constrained using information from existing published literature and government agency data. For streamflow, the "Sensitivity Analysis" routine provided within SWAT performs 280 runs, comparing iterations of runs with varied parameter ranges to the mean simulated flow value. Latin Hypercube (LH) sampling and the one-factor-at-a-time (OAT) method will be applied (Van Griensven *et al.*, 2006) to determine parameter sensitivity.

The process results in a ranking of the parameter sensitivities, along with the total flow variability each parameter accounts for. For calibration, we will retain the subset of parameters accounting for 90% of the total variability. In our experience, a small handful of the parameters explain almost all of the variability. We will limit our calibration to ten parameters, regardless of the total variability explained.

SWAT Simulations and Output Data Processing. Calibration and validation will be used to identify value ranges for the most sensitive flow parameters identified in the sensitivity analysis. This process is described in the model calibration process (Section 2.5).

Once calibration and validation is complete, values for these parameters will be loaded into SWAT. At this point, SWAT will be ready to use for assessing the potential impact of large volume withdrawals (consumptive use of freshwater) from Hydraulic Fracturing (HF) operations on drinking water resources for the Garfield County area of the Upper Colorado River Basin (UCRB) (See QAPP section A5).

SWAT performs computations on a daily basis and we will extract model output results on a daily basis over the simulation period(s). There are several different types/forms of SWAT output data. We will place attention to output data located in the output.hru, output.rch, output.std and output.sub files (see http://swat.tamu.edu/documentation/ for explanation of these files). In particular, we will focus on the following hydrologic-based parameters: IN/OUT FLOW (.rch file), GW_Q (.sub file), GW_RCHG (.hru file), and DA_RCHG (.hru file). IN/OUT identifies flow into and out of a watershed or sub-basin. GW_Q is ground water contribution to streams flow (baseflow), GW_RCHG is the amount of water entering the shallow and deep aquifers and DA_RCHG is amount of water entering the deep aquifer. The ouput.std file provides a summary of basin-wide averages. See Table A9 for the SWAT output variables passed to the next analysis step.

Table A8. Flow-related input parameters that will be used in the sensitivity analysis: SWAT will run a sensitivity analysis on these parameters to determine their relative sensitivity on flow simulation. The parameters that account for 90% of the total flow variability will be used in the calibration and validation phase.

Input Parameter	Description
CN2.mgt	SCS runoff curve number for moisture condition II
ALPHA_BF.gw	Baseflow alpha factor
CH_N2.rte	Manning's " n" value for main channel
SURLAG.bsn	Surface runoff lag coefficient
ESCO.hru	Soil evaporation compensation factor
GWQMN.gw	Threshold depth of water in the shallow aquifer required for return flow to
CANMX.hru	Maximum canopy storage
SOL_AWC.sol	Available water capacity of soil layer
GW_REVAP.gw	Groundwater "revap" coefficient
REVAPMN.gw	Threshold depth of water the shallow aquifer for "revap" or percolation to the
GW_DELAY.gw	Groundwater delay time
RCHRG_DP.gw	Deep aquifer percolation fraction
SLSUBBSN.hru	Average slope length
CH_K2.rte	Effective hydraulic conductivity in main channel alluvium
SOL_K.sol	Saturated soil hydraulic conductivity
BLAI.CROP.dat	Maximum potential leaf area index
SOL_Z.sol	Depth from soil surface to bottom of layer
SOL_ALB.sol	Moist soil albedo
TIMP.bsn	Snow pack temperature lag factor
SFTMP.bsn	Snowfall temperature
SMFMN.bsn	Melt factor for snow on December 21
SMFMX.bsn	Melt factor for snow on June 21
SMTMP.bsn	Snow melt base temperature
TLAPS.sub	Temperature lapse rate

Table A9. SWAT Output Time Series sent to Consumptive Use processing files.

Output	Parameters	Parameter Description
Output.sub	"GW_Q"	Groundwater contribution to stream flow (baseflow) (mm/day)
Output.rch	"IN/OUT_FLOW"	Flow into and out of a watershed or sub-basin (average cms/day)
Output.hru	"GW_RCHG"	Water entering the shallow and deep aquifers (mm/day)
Output.hru	"DA_RCHG"	Water entering the deep aquifer (mm/day)
Output.std	All water balance parameters	See file

Section 2.3 HSPF Model

The Hydrological Simulation Program – Fortran (HSPF) simulates hydrology and water quality at various temporal and spatial scales. HSPF is the core catchment model in BASINS and it is widely used for EPA's regulatory (e.g., Chesapeake Bay TMDL) and policy decisions. HSPF model uses hourly historical precipitation and evapotranspiration data to predict hourly streamflow time series data. Detailed description of HSPF can be found in HSPF model's User's Manual (Bicknell et al., 2001).

HSPF Description (from USGS, 2013)

HSPF uses continuous rainfall and other meteorologic records to compute streamflow hydrographs and pollutographs. HSPF simulates interception soil moisture, surface runoff, interflow, base flow, snowpack depth and water content, snowmelt, evapotranspiration, ground-water recharge, dissolved oxygen, biochemical oxygen demand (BOD), temperature, pesticides, conservatives, fecal coliforms, sediment detachment and transport, sediment routing by particle size, channel routing, reservoir routing, constituent routing, pH, ammonia, nitrite-nitrate, organic nitrogen, orthophosphate, organic phosphorus, phytoplankton, and zooplankton. HSPF can simulate one or many pervious or impervious unit areas discharging to one or many river reaches or reservoirs. Frequency-duration analysis can be done for any time series. Any time step from one minute to one day can be used, and any period from a few minutes to hundreds of years may be simulated. HSPF is generally used to assess the effects of land-use change, reservoir operations, etc. Separate programs are available to support data preprocessing and post-processing for statistical and graphical analysis of HSPF output. A conceptual diagram for HSPF is presented in Figure A29. The model contains hundreds of process algorithms developed from theory, laboratory experiments, and empirical relations from instrumented watersheds.

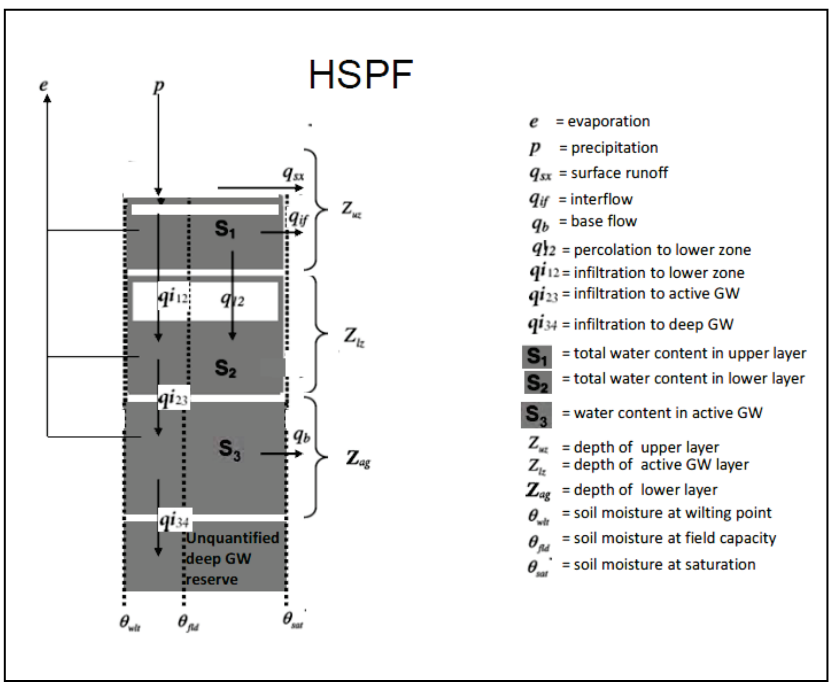


Figure A29 HSPF schematic

Important features of HSPF model structure include the following:

1. Water movement is not computed using a scheme that differentiates between free water and tension water.
2. Compartments include upper zone, lower zone, active groundwater and deep groundwater. Movement of water to deep groundwater is uni-directional; deep groundwater is a sink with unspecified depth or volume.
3. During precipitation events infiltration fluxes are calculated that move water from the upper zone, to the lower zone, to active groundwater, and to deep groundwater.
4. During period of no precipitation the opportunity for percolation from the upper zone to the lower zone is simulated.
5. Lateral fluxes include surface runoff, interflow from the upper zone, and base flow from the active groundwater.
6. Evaporation fluxes are computed from the upper zone, the lower zone and from active groundwater.

HSPF Data Requirements and Model Setup

Meteorologic records of precipitation and estimates of potential evapotranspiration are required for watershed simulation. Air temperature, dewpoint temperature, wind, and solar radiation are required for snowmelt. Air temperature, wind, solar radiation, humidity, cloud cover, tillage practices, point sources, and (or) pesticide applications may be required for water-quality simulation. Physical measurements and related parameters are required to describe the land area, channels, and reservoirs.

Table A10 lists hydrologic and meteorological data required by HSPF to estimate streamflow (water availability) at the selected pour point.

Additional input data requirements are spatial datasets representing land use, soil types, topography (DEM), and hydrography. HSPF setup can be automated using BASINS (USEPA, 2001), or manually initialized by loading the spatial and time series data into the HSPF interface. Detailed instructions for HSPF operation are provided in the HSPF User's Manual (Bicknell *et al.*, 2001).

Table A10. Time series data required to simulate streamflow with HSPF model

Data Type	Time Resolution	Units
Precipitation	1 hour	inches
Minimum Air Temperature	1 hour or day	degree Fahrenheit
Maximum Air Temperature	1 hour or day	degree Fahrenheit
Potential evapotranspiration	1 hour or day	inches
Streamflow	1 day	cfs
Groundwater levels	1 day or month	ft

HSPF Sensitivity Analysis

After model setup has been completed, a sensitivity analysis is performed to evaluate which model parameters have the greatest influence on simulated streamflow. We will identify the most sensitive parameters to minimize the number of calibrated parameters (Table A11). We will first define the allowable range for each parameter, all of which have a physical meaning and will be constrained using information from existing published literature and government agency data. Parameter estimation model (PEST) will be used for one-at-a-time (OAT) sensitivity analysis on a minimum number of 200 model runs (Watermark Numerical Computing, 2004). After identification of the most sensitive parameters, model setup is complete, and we will proceed to our model calibration process (Section 2.5).

After the model is calibrated, simulations per the study design are performed. HSPF output variables passed to the next analysis step are listed in Table A12.

Table A11 Candidate HSPF streamflow simulation parameters.

Input Parameter	Description
LZETP	Lower zone evapotranspiration parameter
INFILT	Index to mean soil infiltration rate (in/hr)
LZSN	Lower zone nominal soil moisture storage (inches)
INTFW	Interflow inflow parameter
IRC	Interflow recession parameter
DEEPFR	Fraction of groundwater inflow to deep recharge
AGWRC	Groundwater recession rate
UZSN	Nominal upper zone soil moisture storage (inches)
BASETP	Fraction of remaining evapotranspiration from baseflow
KVARY	Groundwater recession flow parameter
CEPSC	Interception storage capacity

Table A 12. HSPF Output Time Series Sent to Consumptive Use processing files.

Variable	Output file location
Daily streamflow at main outlet	
Daily streamflow at sub-basin outlet	
Daily recharge over entire watershed	

Section 2.4 GFLOW Groundwater Model

Traditional watershed models, such as SWAT and HSPF, have advanced capabilities to represent the surface water and soil water components of the hydrologic cycle, while traditional groundwater models such as GFLOW have advanced capabilities to represent the saturated subsurface components of the hydrologic cycle. Ground water models and watershed models will be built and tested along parallel tracks. At appropriate moments in the workflow the conceptualizations and parameterizations will be compared and evaluated for reasonableness. The primary objective of the ground water modeling is to support the watershed model in its capability to represent the change in groundwater storage time series that feeds the impact assessment. A secondary objective is to build a ground water model for representation of the site specific influences on the water table (aquifer depletion) and on local impacts on stream reaches (streamflow depletion). In this project we will use field observations of stream discharge and piezometric head (or water table elevation for unconfined aquifers) to test the performance of the ground water simulation model.

The proposed step-wise and progressive groundwater modeling strategy will progress in three steps: (1) baseflow separation; (2) single-layer ground water model; and (3) multi-layer ground water modeling (Table A13). The baseflow separation will provide an estimate of the averaged infiltration recharge over the entire watershed draining to the pour point or outlet. Deep leakage is assumed negligible during baseflow separation. This recharge is handed to the single layer groundwater model as a boundary condition, and a manual calibration will be conducted that maintains the analyzed baseflow to the rivers and fits a predicted water table to observed the discrete point measures of hydraulic heads at wells, resulting in a ratio of the effective total recharge (shallow infiltration recharge minus deep leakage) over the shallow aquifer hydraulic conductivity.

At this point in the workflow it makes sense to check for consistency with the watershed modeling and calibration. An initial point of connection is the predicted perennial stream network. Any major discrepancies between the watershed modeling approach and the groundwater modeling approach (e.g., a change in total network stream length greater than 10%) would need to be explained and evaluated. The accepted watershed model would provide predictions of average shallow infiltration recharge and deep leakage as a function of time and space, and the accepted single-layer ground water model would provide an average effective watershed scale hydraulic conductivity. Future ground water model refinements should be checked to maintain this watershed scale average hydraulic conductivity.

The results from the single-layer ground water model are expected to provide initial and boundary conditions to the multi-layer groundwater model. The single-layer groundwater model suggests the lateral boundaries of the no-flow boundary condition for the groundwatershed.

The selection of specific models for purpose is shown in Table A13.

Table A 13. Ground water modeling tool selection and detailing of workflow.

BASEFLOW SEPARATION	USGS PART/RECESS/RORA (Rutledge,1998); recession curve displacement method that estimates average baseflow at pour point of watershed using recession method, and assuming no deep aquifer leakage and assuming the groundwatershed equals the surface watershed, estimates average recharge over the watershed.
SINGLE LAYER GROUND WATER MODEL	Haitjema Software GFLOW (Haitjema,1995b). Vector-based analytic element model, steady state flow, open boundary conditions, dynamic stream network generation to satisfy baseflow at pour point and observed heads, optimal values for average hydraulic conductivity and streambed resistance, leakage.
MULTI-LAYER GROUND WATER MODEL	*We will make a decision to use MODFLOW if conceptual need is indicated and data is available. Steps will be specified if that event occurs.* USGS MODFLOW (Harbaugh, 2005) Cell-based finite difference model, transient flow, options for parameter estimation and uncertainty analysis using PEST (Watermark Numerical Computing, 2004).

PART

The computer program PART (Rutledge, 1998) uses streamflow partitioning to estimate a daily record of groundwater discharge under the streamflow record, or baseflow (http://water.usgs.gov/ogw/part/). The method designates groundwater discharge to be equal to streamflow on days that fit a requirement of antecedent recession, linearly interpolates groundwater discharge for other days, and is applied to a long period of record to obtain an estimate of the mean rate of groundwater discharge.

Basic Steps

1. Download the historical daily discharge records from the USGS stream gages in standard ASCII text format.
2. Preview the records for completeness.
3. Run baseflow separation and report the results.

GFLOW

GFLOW is a regional groundwater models based on the analytic element solution methods (www.analyticelements.org). The analytic element is a technique for regional groundwater modeling that in practice captures the accuracy of exact analytical solutions with the computational capabilities (Strack, 1989; Haitjema, 1995b; Hunt, 2006; Kraemer, 2007). The solutions are based on the superposition of point sinks representing wells, line sinks representing rivers, and area elements representing inhomogeneities in aquifer properties. The vector basis of the analytic element method interfaces logically with the vector basis of the chosen watershed models HSPF and SWAT. The conceptual model is single layer steady flow (GFLOW, www.haitjema.com).

Basic steps

1. Build the basemap.
2. Create linesink strings representing rivers.
3. Create area element polygon representing recharge over the groundwatershed.
4. Create point coverage for known public and self supplied wells.
5. Create test point coverage of observed heads.
6. Run the GFLOW model and conduct manual calibration.

The analytic element model GFLOW will be used for modeling the "groundwatershed" (Haitjema, 1995a) draining to the pour point of the topographically-defined catchment under investigation and the relevant time period. The GFLOW model will be calibrated to: (1) observed/estimated average annual baseflow in the streams, either observed at the USGS gages or estimated based on regression at selected river points; and (2) observed annual averaged shallow aquifer water levels in wells; and (3) water elevations in perennial streambeds as inferred from USGS topographic maps.

We will use the conjunctive groundwater and surface water analytic element modeling technique of Mitchell-Bruker and Haitjema (1996) to define the perennial stream network needed to support the long term average baseflow from the PART baseflow separation analysis (Figure A30).

Three parameters are known to control the model predicted piezometric surface (heads) and baseflow in the streams: recharge, hydraulic conductivity or transmissivity, and streambed resistance. Annual average recharge will be constrained by the basin-scale watershed models per sub-basin. Streambed resistance will be estimated based on soil materials and assumed thickness. Hydraulic conductivity will be varied to minimize residual error between model predictions and observations.

The project team will consider the use of MODFLOW based on conceptual need (e.g. deep leakage is found to be important) and there is data available to support its use. In this event, the use of the model will be defined at that time.

The decision on need for MODFLOW will be assisted by a rule of thumb and the equation for dimensionless tau and the categories shown in Table 14:

$$\tau = \frac{SL^2}{4kbP}$$

where *S* is the aquifer storativity or specific yield [-], *L* is the average distance between surface water features [L], *k* is the hydraulic conductivity [L/T], *b* is the average saturated thickness [L], and *P* is the periodic forcing [T].

Table A 14. Guidance criteria for selecting groundwater model.	
$\tau > 1$	Use a steady-state model with time-averaged boundary conditions and recharge rates (e.g., GFLOW)
$0.1 < \tau < 1$	Use a transient model with transient boundary conditions and recharge rates (e.g., MODFLOW or TTim)
$\tau < 0.1$	Use a steady-state model with instantaneous boundary conditions and recharge rates, for instance, representing summer or winter conditions (e.g., GFLOW)

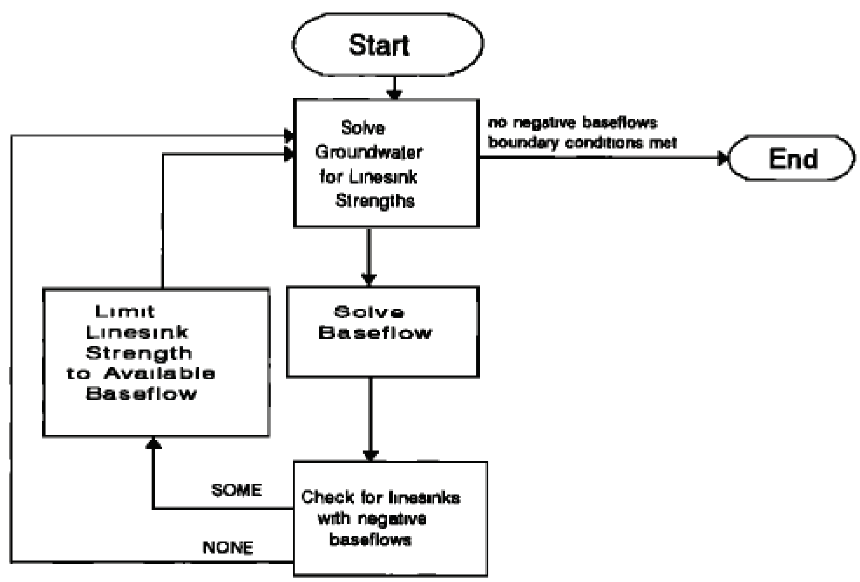
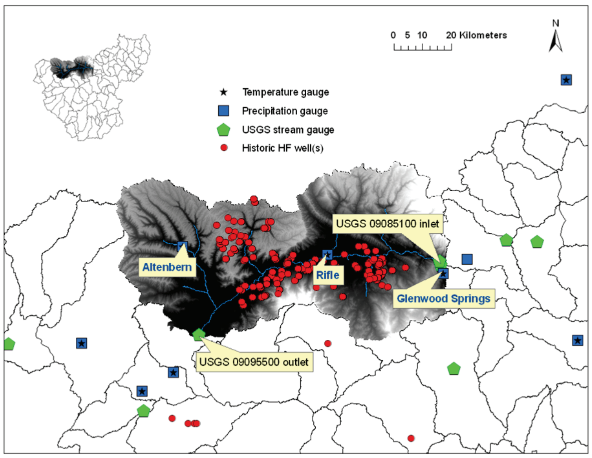
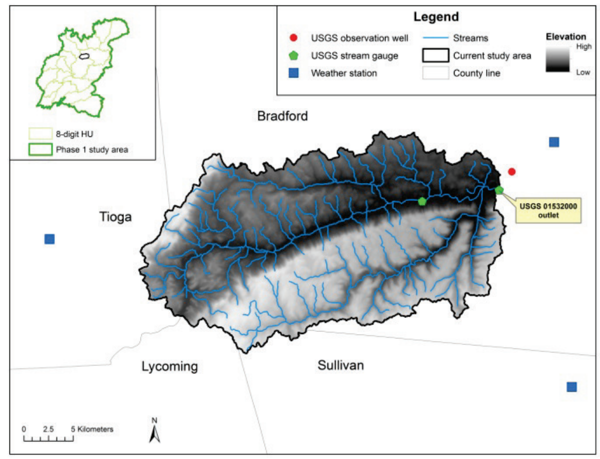


Figure A30. Flowchart for coupling base flow and groundwater flow after Mitchell-Bruker (1996) as implemented in GFLOW.

Table A15. Groundwater modeling data requirements (in addition to USGS data)

Garfield	**Bradford**
	
Hydrography map based on NHD+	**Hydrography map based on NHD+**
DEM catchment based on USGS 10 m	**DEM catchment based on USGS 10m**
Major public and self-supplied wells based on EPA records	**Major public and self-supplied wells based on EPA records**
Observed heads based on State records	**Observed heads based on State records**

Section 2.5 Model Calibration

Following initial setup and parameter selection, the watershed and groundwater models will be calibrated independently but cross-referencing key parameters simultaneously. This process of model calibration, cross referencing, and validation was described briefly in Section A5. Details of this process that will be conducted for both watershed models are described in this section.

Surface Water Models

1. Overview

Watershed model calibration will be conducted:

- After watershed model pre-processing and sensitivity analyses have been completed (Appendix A sections 2.3 (SWAT) and 2.4 (HSPF)
- Prior to scenario modeling (Project Task 3)

Our key calibration objective is to develop parameterizations that can be reasonably applied to other regional watersheds to estimate water stress vulnerability resulting from HF. Mechanistic watershed model calibration is performed to optimize a number of parameters that influence the routing, timing, and relative importance of various watershed processes. Our calibration will focus on median and lower flows, when water stresses from HF are potentially greatest. We will use a maximum likelihood approach, over physically realistic ranges of a minimized number of parameters, which includes multiple objective functions and returns sets of the most likely inputs and associated model outputs. The multi-objective method will allow us to customize our calibration target to the most important flow volumes for HF impact assessment. The simulation set approach provides quantitative uncertainty estimates to supplement the simulated streamflows and optimized parameter values. While our calibration method and post-processing are computationally intensive, we have recently performed these types of calibrations using available resources (Price *et al.*, 2012; Price *et al.*, 2013). For any parts of this process that exceed our desktop computing capabilities within reasonable time frames, we have an onsite parallel computing facility at ERD.

In addition to seeking an optimized parameterization via these model calibration methods, we have set additional acceptance criteria to ensure our findings are regionally representative. We will compare our simulated flow duration statistics to published regional values, and we will compare subsurface outputs from our watershed models with benchmark groundwater models applied in the same study watersheds. If necessary, we will reinitialize and re-parameterize the calibrations to meet these criteria. We will report all final calibration and validation statistics, quantitative uncertainty estimates for parameter ranges and simulated streamflows, and the results from regional cross-validation on other gauged watersheds.

2. Calibration Procedure – Watershed Models

The calibration procedure will be identical for SWAT and HSPF, following standard sensitivity analyses specific to each model (see Section 2.2 and 2.3), and our two study watersheds will be calibrated independently with both models. Our calibration process will utilize a likelihood-weighted Monte Carlo scheme to explore parameter combinations, in order to optimize agreement between simulated and USGS-observed streamflow at the outlets of the modeled watersheds. We will use a standard split-sample approach to calibration and validation (Klemes, 1986), comprising a 27-year observed streamflow record for each watershed. Because our ultimate goal is to explore future scenarios, we will use the latter part of the record (1996-2012) for calibration, and test the resultant optimized parameter set on the earlier part of the record (1986-1996). The USGS streamflow record for the study watersheds confirms that both time series are reasonably stationary and contain a fully representative range of high and low flows, and NLCD landcover data from 1993 and 2006 indicate reasonably consistent land use during this period.

2.1 Multiobjective function – Weighted Nash Sutcliffe (WNS)

An "objective function" is used in mechanistic model calibration to quantify agreement between simulated and observed flows. There is no single objective function that emphasizes all stages and dynamic characteristics of streamflow. One of the most commonly used objective functions in watershed modeling is the Nash-Sutcliffe Efficiency (NS), which is a standardized form of mean squared error (MSE) (Nash and Stufcliffe, 1970; Efstratiadis and Koutsoyiannis, 2010):

$$\mathrm{NS} = 1 - \frac{\sum_{i=1}^{n}(O_i - S_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2}$$

Equation A 1

Where *i* represents each timestep in the series, O = observed streamflow, and S = simulated streamflow. NS and other common objective functions, such as MSE and r^2, bias the optimization to fit flood peaks, at the expense of fitting median and low flows. As in the NS equation, these three common metrics all contain a squared discrepancy between simulated and observed flows, which gives strong bias to large flow volumes and shows less influence on lower flow volumes (Price *et al.*, 2012). Because we are primarily concerned with median and low flows, we require a calibration target that emphasizes these flow characteristics. A successful approach to reducing the high-flow bias of NS is to calculate the fit using log-transformed flow values (Wöhling *et al.*, 2013), which is known as the NS_{log}:

$$\mathrm{NS_{log}} = 1 - \frac{\sum_{i=1}^{n}[\ln(O_i) - \ln(S_i)]^2}{\sum_{i=1}^{n}[\ln(O_i) - \ln(\bar{O})]^2}$$

Equation A 2

While our emphasis will be on median and low flows, we recognize the importance of achieving a reasonable fit to observed flood peaks, in terms of the overall water balance and the routing times of water. It has become common practice to simultaneously use more than one calibration target (a multiobjective function), with the component functions selected on the basis of research goals (Gupta *et al.*, 2009; Price *et al.*, 2012). To achieve a balanced calibration on low and high flows, many previous researchers have used a multiobjective function that aggregates NS (or other MSE derivative) and NS_{log} into a single calibration target (Bekele and Nicklow, 2007; Schaefli *et al.*, 2005; Efstratiadis and Koutsoyiannis, 2010), which will be our approach in this study. We will develop a weighting scheme for these two metrics that will prioritize good fits to observed low and median flows, while maintaining a reasonable representation of flood magnitudes. We will refer to this aggregated objective function as a "weighted Nash-Sutcliffe" score, or WNS. *For context, we will also present r^2 and raw NS values for* *these calibrations.* However, we anticipate these values will be lower than modelers might expect, as the calibration was not designed around these metrics.

2.2 Base WNS scores

Because the WNS is a customized objective function, there is no standard scale for development of acceptance criteria. Both of the component functions produce values ranging from $-\infty$ to unity, which indicates a perfect fit in both metrics. However, we cannot use standard interpretations of the values produced, because the WNS seeks a compromise between the two components, which will reduce the score, while a perfect fit remains unity. To evaluate our success in simulating streamflows, we will develop a set of baseline WNS (Base WNS) scores from observed streamflow and precipitation records (Schaefli and Gupta, 2007).

In the first approach, we will use the mean streamflow of each calendar day, calculated across the entire calibration period. This series of mean daily values will be repeated 17 times, in order to represent the entire calibration period. The WNS will be calculated using the mean daily streamflow as the "simulated" streamflow, using the same weighting scheme of NSE and NSE_{log} that will be applied during the mechanistic model calibration. In the second approach, we will first use basic time series analysis to identify the optimum lag between adjusted precipitation and streamflow (accumulation time). The adjusted precipitation time series will then be shifted to accommodate this lag and treated as the "simulated" streamflow, for which WNS values will be calculated (Schaefli and Gupta, 2007) .

At a minimum, our mechanistic model simulations should exceed the WNS scores calculated for these baseline scenarios. Treating the average daily streamflow and precipitation forcing data as simulated data allows us to estimate the predictive improvement gained from using the mechanistic model, as opposed to simpler empirical methods. The NS and its derivations produce a score < 0 if the observed mean offers more predictive power than the simulated time series. While we expect our benchmark values to exceed 0, in the event that they are negative, we will use the mean model Base WNS of 0 as our baseline.

2.3 Monte Carlo optimization

Our calibrations will be based on 10,000 uniquely parameterized simulations generated by SWAT and HSPF. A Monte Carlo method will be used to parameterize the model runs, and maximum likelihood methods will be used during post-processing to identify optimal ranges for each parameter. The Monte Carlo parameterization scheme will be established via Latin hypercube sampling (LHS). In LHS, the parameter space is conceived as multidimensional, with each parameter as a dimension. The allowable range for each parameter is subdivided into sections of equal size, and unique parameter combinations are derived from all possible combinations across the dimensions. This reduces the number of simulations required to explore the full parameter space, compared to truly random Monte Carlo approaches (Uhlenbrook and Sieber, 2005). This is similar to commonly applied PEST implementations (Doherty and Johnston, 2003), and is largely the same as the Sufi2 process in SWAT-CUP, with the exception that Sufi2 only allows for 2001 parameter sets at a time (Abbaspour, 2009). The choice of a 10,000 run simulation set was based on prior research showing this is generally more than sufficient to approximate the maximum possible fit score that can be achieved within a given modeling setup (Schaefli *et al.*, 2005; Thorndahl *et al.*, 2008).

We will develop a program in R to wrap the models, parameterizing and executing 10,000 simulations (hereafter, "simulation set"), using each of the parameter sets developed in the LHS. We will first run a simulation set for the calibration period (1996-2012). This will generate 10,000 simulated streamflow time series, each associated with a unique parameter set. From this output, we will calculate the WNS for each simulation, maintaining linkages between these fit statistics and the unique parameter sets associated with them. At this point, we will verify that 10,000 runs was sufficient by creating a convergence graph, indicating improvements in fit scores with increasing numbers of simulation. We will initially define convergence as the point when improvements drop below 0.1% over 100 simulations. If this condition is not met within 10,000 runs, we will redo the LHS and/or revisit the convergence criterion. Once the simulation sets are complete, we will retain the 1,000 parameter sets with the highest WNS scores, provided all surpass the Base WNS. If fewer than 1,000 simulation sets exceed the Base WNS, we will revisit the LHS and re-run the calibration period.

2.4 Verification

Before proceeding to formal model validation, we will compare calibration characteristics to other models and to observed data not used in the calibration process. There are two qualitative verification steps we will employ, the first to confirm reasonable surface water variability, and the second to confirm reasonable subsurface recharge rates. From our calibration simulation set, we will calculate flow duration statistics and compare these to published regional ranges for these values, ensuring our values fall within regional variability. In addition, we will extract recharge rates from the surface water model output and compare these to recharge rates generated by a groundwater model, ensuring the benchmark recharge values are within the uncertainty bounds of our simulated recharge. If our first iteration fails to meet either of these conditions, we will adjust initial parameter ranges and reinitialize the calibration process until the acceptance criteria are met. We will accept reductions in WNS scores to accommodate these empirical criteria.

2.5 Model validation

Formal model validation will be performed to evaluate whether the calibrated parameterization is applicable outside the calibration time period. The validation simulation set will consist of the parameter values sets the 1,000 runs calibration runs generating the highest WNS values. The models will be run with these parameter sets for the independent validation time period of 1986-1996. WNS values will be reported for the validation time series. We will again use the best 1000 runs (determined by WNS) for uncertainty assessment. If we find that SWAT and HSPF fail to surpass the Base WNS models, we will rely upon one of the Base WNS models for scenario development.

2.6 Calibration outputs, uncertainty estimation, and cross-validation

Calibration will strive to obtain the best fit possible, with the requirements of 1) exceeding the Base WNS scores described in 2.2, 2) demonstrating agreement with regional flow duration curves, and 3) demonstrating agreement with recharge estimates from an independent groundwater model. Formal validation will be performed on a time series independent of the calibration period. Once the conditions of the verification and validation processes are met, the calibration phase of watershed modeling is complete. The calibrated and validated parameter sets will then be used to generate simulation sets for each scenario to be modeled, which are described in sections 2.2 and 2.3. The WNS score will be calculated for each simulation. We will identify and report the best individual simulation as well as quantified uncertainty on the simulated streamflows, incorporating the full range of streamflow values generated for each timestep. The distributions of parameter values selected via the calibration process will provide uncertainty estimates for the parameter values themselves. We will report the quantified uncertainty for the both the optimized parameters and the simulated streamflow, to aid decision makers in interpretation of results.

The feasibility of using our optimized parameterizations in ungauged regional watersheds will be examined using a cross-validation approach. This approach will determine the loss of model accuracy that occurs when applying calibrated parameters from a similar watershed. As an example, if we have 10 sets of maximum likelihood parameter calibration for 10 different (but hydrologically similar) watersheds, we also have 10 different estimates of a fit statistic (e.g., WNS). In each watershed, we can rerun the surface water model 9 times using the calibration parameter sets from the other watersheds and recalculate the fit statistic. This allows us to compare the fit score distribution for the calibrated watersheds to fit score distribution of "transporting" calibrated parameters across watersheds. If the loss of prediction accuracy is acceptably small, then this approach can be used to produce surface water flow estimates at watersheds where we do not have calibration data available.

3. Groundwater Models

Groundwater model calibration will be conducted:

- Independently from surface water model calibration
- After groundwater model initialization
- Iteratively, for optimized alignment with leakage values from surface water model output

The groundwater calibration workflow for this project is shown in Figure A 31. In this project we will use field observations of stream discharge and piezometric head (or water table elevation for unconfined aquifers) to test the performance of the ground water simulation model. The process of calibration involves the systematic selection of model parameters in order to minimize the difference between the observed groundwater property and simulated groundwater property, often referred to as the "goodness of fit". The quantitative comparison between the simulated and the observed is captured in an objective function, which will be described below. The calibration process is also open to modification of the hydrogeologic conceptual model. For example, an initial single layer ground water model that assumes no deep aquifer leakage may not result in a satisfactory calibration, while a more complex multi-layer ground water model with leakage represented as a parameter may improve the goodness of fit. The iterative workflow accommodates the improvement in both the conceptual model and the parameterization until a satisfactory or optimum value of the objective function is achieved. The calibrated model is then evaluated for its appropriateness for the intended use, and his involves the assessment of the uncertainty of predictions.

We will be predicting changes in groundwater storage in space and time. The groundwater modeling workflow allows for the return to the adjustment of the conceptual model and parameterization if the prediction uncertainty is not acceptable. The goodness of fit acceptance criteria is discussed below. The intended use in this project is to characterize the impact of hydraulic fracturing on drinking water, and the groundwater model is expected to predict ground water availability for various spatial and temporal scales, in order to characterize a ground water index of impact as previously described (Section A5).

We will follow the broad guidelines for ground water model calibration as documented by the U.S. Geological Survey (Hill, 1998; Figure A31). The initial phase of ground water model calibration will be done manually using simple models and focus on the evaluation of the conceptual model. The next phase of ground water model calibration will be automated and optimize on parameterization to minimize the objective function.

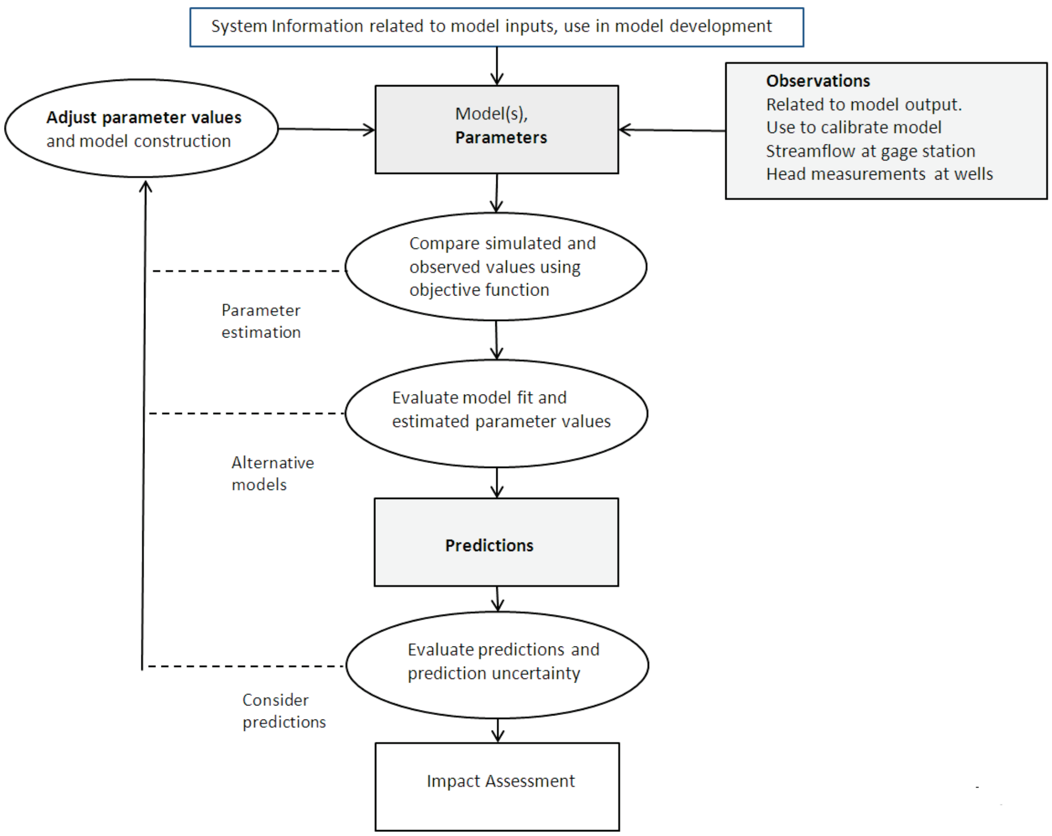


Figure A31. The general ground water modeling workflow for this project will include the model parameterization phase and the model prediction phase with the objective to evaluate impact (after Hill and Tiedeman, 2007).

3.1 Objective function

The objective function is calculated as the squared sum of the weighted residuals (including prior information).

$$\Phi = \sum_{i=1}^{m}(w_i r_i)^2$$ **Equation A 3**

where,

r_i is the *i*'th residual expressing the difference between the model outcome and the actual field measurement for the *i*'th observation, and

w_i is the weight associated with the *i*'th observation.

Calibration Guidelines

Model development

1. Start simple, add complexity carefully
2. Use a broad range of information
3. Be well-posed & be comprehensive
4. Include diverse observation data for 'best fit'
5. Use prior information carefully
6. Assign weights that reflect 'observation' error
7. Encourage convergence by making the model more accurate
8. Consider alternative models

Model testing

9. Evaluate model fit
10. Evaluate optimal parameter values

Potential new data

11. Identify new data to improve parameter estimates
12. Identify new data to improve predictions

Prediction uncertainty

13. Use deterministic methods
14. Use statistical methods

Figure A32. The guidelines for ground water model calibration as recommended by the US Geological Survey (after Hill and Tiedeman, 2007).

We have multiple options available for implementing this groundwater model calibration strategy. GFLOW and MODFLOW both provide onboard linkages to the PEST model calibration program (Watermark Numerical Computing, 2004), which may facilitate our parameter optimization approach using the above objective function (Equation A.3). If this strategy proves to be inordinately time-consuming or require specific expertise for PEST customization, we will instead use the model-independent likelihood estimation Monte Carlo approach described for surface water model calibration (Section 2.5).

The proposed step-wise and progressive groundwater modeling strategy will progress in three steps: (1) baseflow separation; (2) single-layer ground water model; and, if necessary, (3) multi-layer ground water modeling. The baseflow separation will provide an estimate of the averaged infiltration recharge over the entire watershed draining to the pour point or outlet. Deep leakage is assumed negligible during baseflow separation. This recharge is handed to the single layer groundwater model as a boundary condition, and a manual calibration will be conducted that maintains the analyzed baseflow to the rivers and fits a predicted water table to observed the discrete point measures of hydraulic heads at wells, resulting in a ratio of the effective total recharge (shallow infiltration recharge minus deep leakage) over the shallow aquifer hydraulic conductivity.

3.2 Verification

At various points in the workflow we will check for consistency with the watershed modeling and calibration. An initial point of connection is the predicted perennial stream network. Any major discrepancies between the watershed modeling approach and the groundwater modeling approach (e.g., a change in total network stream length greater than 10%) would need to be explained and evaluated. The accepted watershed model would provide predictions of average shallow infiltration recharge and deep leakage as a function of time and space, and the accepted single-layer ground water model would provide an average effective watershed scale hydraulic conductivity. Future ground water model refinements should be checked to maintain this watershed scale average hydraulic conductivity.