GWERD QUALITY ASSURANCE PROJECT PLAN

TITLE: DATA COLLECTION/MINING FOR HYDRAULIC FRACTURING CASE STUDIES

WORK ASSIGNMENT NO.: WA-HF-2-10

TECHNICAL DIRECTIVE: 7HF101HF

QA ID NO.: G-15952 QA Category: I

CONTRACTOR: Shaw Environmental & Infrastructure, Inc., #EP-C-08-034

REVISION: 1

DATE: January 10, 2012

/s/ 1/18/2012
EPA Technical Research Lead Date

/s/ 1/12/2012
Shaw Technical Lead Date

EPA APPROVALS:

/s/ 1/11/2012
Project Officer Date

/s/ 1/16/2012
GWERD QA Manager Date

SHAW APPROVALS:

/s/ 1/13/2012
Shaw Program Manager Date

/s/ 1/10/2012
Shaw QA Officer Date

Table of Contents

LIST OF FIGURES

LIST OF TABLES

DISTRIBUTION LIST

Cynthia Paul – EPA GWERD, Ada, OK
Susan Mravik - EPA GWERD, Ada, OK
Steve Vandegrift- EPA GWERD, Ada, OK
Dr. David Jewett - EPA GWERD, Ada, OK
Dr. Stephen Kraemer - EPA ERD, Athens, GA
Sujith Kumar – Shaw, EPA Contract, Ada, OK
Dr. Duane Root – Shaw, Knoxville, TN
Jonathan Shireman – Shaw, Knoxville, TN
Victor Murray – Shaw, Ada, OK
Chaitanya Nellutla – Shaw, Denver, CO
Rob Earle – Shaw, Ada, OK
Vickie Grissom – Shaw File Administrator, Ada, OK

1.0 PROJECT MANAGEMENT

1.1 PROJECT / TASK ORGANIZATION

Project Organization is as shown below in Figure 1.

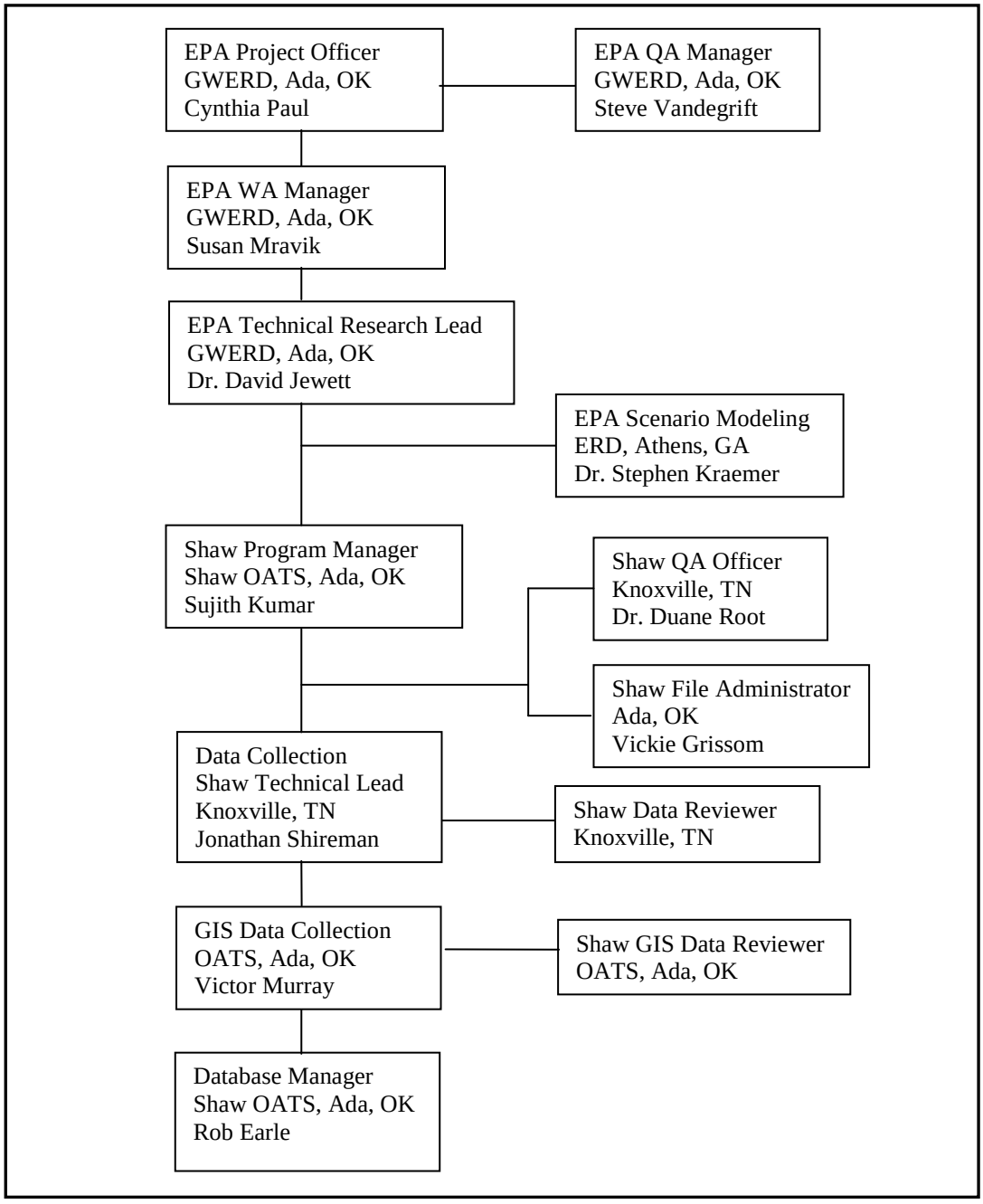

Figure 1. Project Organization Chart

Dr. David Jewett, EPA Technical Research Lead, U.S. Environmental Protection Agency, Ground Water and Ecosystems Restoration Division, Ada, OK. Responsible for technical oversight, technical review of the Quality Assurance Project Plan (QAPP), ensuring project goals are achieved, and review/approval of project deliverables.

Dr. Stephen Kraemer, EPA Scenario Modeling, U.S. Environmental Protection Agency, Ecosystems Research Division, Athens, GA. Responsible for hydraulic fracturing scenario modeling (to be addressed in a separate QAPP) and for conforming to approved QAPP requirements.

Ms. Cynthia Paul, EPA Project Officer, U.S. Environmental Protection Agency, Ground Water and Ecosystems Restoration Division, Ada, OK. Responsible for contract oversight, review/approval of the QAPP,

Ms. Susan Mravik, EPA Work Assignment Manager (WAM), Environmental Protection Agency, Ground Water and Ecosystems Restoration Division, Ada, OK. Responsible for providing technical direction, review of the QAPP and review/approval of project deliverables, management of project records, QA, and resolution of QA issues.

Mr. Steve Vandegrift, EPA QA Manager, Environmental Protection Agency, Ground Water and Ecosystems Restoration Division, Ada, OK. Responsible for QA review/approval of the QAPP, conducting audits, and QA review/approval of the final product.

Mr. Sujith Kumar, Shaw Program Manager, Shaw Environmental & Infrastructure, Inc., Ada, OK. Responsible for management of Shaw project activities, review and approval of the QAPP, assuring implementation of the approved QAPP, and review and approval of Shaw project deliverables.

Mr. Jonathan Shireman, Shaw Technical Lead and Data Provider, Shaw Environmental & Infrastructure, Inc., Knoxville, TN. Responsible for data collection, evaluation of data against acceptance criteria, preparation of compiled data files, and reviewing and conforming to approved QAPP requirements.

Mr. Victor Murray, Shaw GIS Analyst, Shaw Environmental & Infrastructure, Inc., Ada, OK. Responsible for GIS data collection, evaluation of GIS data against acceptance criteria, preparation of compiled GIS data files, and reviewing and conforming to approved QAPP requirements.

Mr. Rob Earle, Shaw Database Manager, Shaw Environmental & Infrastructure, Inc., Ada, OK. Responsible for database development and management and for conforming to approved QAPP requirements. Mr. Earle will be assisted as needed by Mr. Chaitanya Nellutla, Shaw, Denver, CO.

Dr. Duane Root, Shaw QA Officer, Shaw Environmental & Infrastructure, Inc., Knoxville, TN. Responsible for preparation of the QAPP, implementation of approved QAPP, conducting project TSAs, reviewing corrective actions and preparing quarterly reports to management on the status of the QMS with respect to project goals.

Ms Vickie Grissom, Shaw File Administrator, Ada, OK. Responsible for maintaining project files as described in the QAPP.

Shaw Data Reviewer, Shaw Environmental & Infrastructure, Inc., Ada, OK/Knoxville, TN. The Data Reviewer shall have experience and training as described in Section 1.5, Special Training and Certification. The Data Reviewer will be responsible for reviewing collected data against acceptance criteria, reviewing data and data file manipulations for accuracy, and conforming to approved QAPP requirements.

1.2 PROBLEM DEFINITION / BACKGROUND

As natural gas production has increased, so have concerns about the potential environmental and human health impacts of hydraulic fracturing in the United States. Hydraulic fracturing, which involves the pressurized injection of water, chemical additives, and proppants into a geologic formation, induces fractures in the formation that stimulate the flow of natural gas or oil, thus increasing the volume of gas or oil that can be recovered from coalbeds, shales, and tight sands—the so-called "unconventional" reservoirs. Many concerns about hydraulic fracturing center on potential risks to drinking water resources, although other issues have been raised. In response to public concern, Congress directed the United States Environmental Protection Agency (EPA) to conduct research to examine the relationship between hydraulic fracturing and drinking water resources (USEPA, 2011a).

EPA will compile data on hydraulic fracturing water use and the hydrology of selected study areas, case studies or scenario evaluations. These data will include precipitation data, ground water levels, surface water flows, and water quality as well as data on hydraulic fracturing operations, such as the location of wells and the recorded water used/handled during fracturing activities. The EPA study approach will include specific case studies, both retrospective and prospective as well as data collection from selected regional study areas and scenario evaluations as described in the *Plan to Study the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources* (USEPA, 2011a).

Retrospective case studies are focused on investigating reported instances of drinking water resource contamination in areas where hydraulic fracturing events have already occurred. The goal is to determine whether or not the reported impacts are due to hydraulic fracturing activities. These studies will use existing data and may include environmental field sampling, modeling, and/or parallel laboratory investigations.

Prospective case studies involve sites where hydraulic fracturing will be implemented after the research is initiated. These cases allow sampling and characterization of the site prior to, during,

and after drilling, water extraction, injection of the fracturing fluid, flowback, and production. At each step in the process, data will be collected to characterize both the pre- and post-fracturing conditions at the site.

Scenario evaluations explore realistic, hypothetical scenarios across the hydraulic fracturing water lifecycle that may result in adverse impacts to drinking water resources based on current understanding and available data. The scenarios will include a reference case involving typical management and engineering practices in representative geologic settings. Typical management and engineering practices will be based on what EPA learns from case studies as well as the minimum requirements imposed by state regulatory agencies. Potential modes of failure, both in terms of engineering controls and geologic characteristics, will be introduced and modeled to represent various states of system vulnerability. The scenario evaluations will produce insights into site-specific and regional vulnerabilities.

Simple water balance analysis will be conducted using available data. The collected data will be compiled in conjunction with hydrological trends over the same period of time. Control areas that have similar baseline water demands and have no oil and gas development will be compared to areas with intense hydraulic fracturing activity to isolate and identify the impacts of hydraulic fracturing on water availability. Control areas will be within the study area if possible or in the same vicinity but will be remote from oil and gas development activities to be devoid of hydraulic fracturing impacts (see Section 2.9.1). A critical analysis of trends in water flows (including "environmental flows") and water usage patterns in areas impacted by hydraulic fracturing activities will be conducted to determine whether water withdrawals for hydraulic fracturing activities alter ground and surface water flows. Data collection will support the assessment of the impacts of hydraulic fracturing on water availability at various spatial scales (e.g., site, watershed, basin, and play) and temporal scales (e.g., days, months, and years).

The Ground Water and Ecosystems Restoration Division (GWERD) is the U.S. Environmental Protection Agency's center for risk management research on the subsurface environment and the interface of that environment with other environmental compartments such as surface water and the atmosphere. The GWERD is involved with the EPA Office of Research and Development (ORD) Hydraulic Fracturing research efforts and is leading activities for targeted data collection, federal partner data collection, study data collection and web access for the public information. The Shaw Work Assignment (WA) #WA-HF-2-10 from GWERD through the Onsite Analytical and Technical Support (OATS) contract provides for field, technical, and database support to the EPA for investigations of hydraulic fracturing impacts to drinking water resources. (USEPA 2011b). Shaw's support will be directed by EPA as projects or tasks in the form of documented Technical Directives (TD).

1.3 PROJECT/TASK DESCRIPTION

This Quality Assurance Project Plan (QAPP) will address data collection activities as described below.

The data collection activities will be focused on existing water supply and water quantity data, but may also include water quality data, if available, for hydraulic fracturing activities related to natural gas production. Data will be collected from sources, such as USGS, state Departments of Environmental Quality (DEQ) or Departments of Environmental Protection (DEP), private water management groups, Army Corps of Engineers, USDA-NRCS and state GIS Departments, etc., that have collected and/or published data. Shaw will begin identification and survey of existing sources of water supply, quantity, and quality data as well as hydraulic fracturing activity information for the following two selected scenario study area/regional locations:

- Susquehanna River Basin/Marcellus Shale in Pennsylvania
- Garfield County/Piceance Basin in Colorado.

These locations represent humid and arid areas of the country, for which sufficient data are available for study. Due to the large volumes of water required for fracturing operations significant impacts may vary from one part of the country to another and from one time of the year to another. Humid areas with greater precipitation and surface water volume will likely be less affected by large volume water use in fracturing activities than arid areas with generally less water availability. Including study areas that represent both arid and humid conditions will provide a contrast of impacts for the different conditions and the range of impacts is likely to be inclusive of most areas of the country.

Case study data collection may be from any number of retrospective and prospective case study nominations as discussed in Section 2.9.1, Scope of Data Collection. Case studies may be defined or amended by EPA based on results from data collection activities.

Initial project activities will involve three phases of data collection and organization as defined below:

1. Initial discovery, survey and inventory of applicable data from known, recommended or otherwise discovered sources,
2. Collection of data in electronic format - Data from applicable information sources will be evaluated against acceptance criteria (Section 2.9 Non-Direct Measurements) by the Data Collector and if deemed acceptable, the data will be saved in electronic data files in an organized manner. The Data Collector will also prepare an inventory of data sources, data type and amount of data collected.
3. Compilation of the data into usable formats for entry into Geographic Information System (GIS), database software and use in modeling programs - It may be possible to directly compile collected data into user software programs; however, data entry (or use by programs) may require preparation of compiled data files.

In addition to these specific data collection activities, other activities including GIS map preparation as well as database and modeling scoping and requirement specification may be performed. GIS activities will include preparation of national and regional map projections or overlays with addition of specific geographic locations, identifications and details as defined by

EPA. Database and site modeling scoping will be ongoing based on development of the source information inventory as source survey and data collection activities progress.

1.3.1 Initial Discovery, Survey and Inventory of Source Data

The Data Collector will search for sources that have collected and/or published applicable data and inventory data that is available. EPA's suggested sources of initial data for the two scenario study locations are:

- Susquehanna River Basin (PA): Susquehanna River Commission, USGS, Pennsylvania Dept. of Environmental Protection, Pennsylvania Department of Natural Resources (PADNR), RAIN Water Quality Network, US Army Corps of Engineers
- Garfield County (CO): Colorado Oil and Gas Conservation Commission, (COGCC), Colorado Department of Natural Resources (CONDR), USGS.

Shaw has also been directed to additional data sources (references) that can be found in the Hydraulic Fracturing Study Plan, in particular Chapter 6 (USEPA, 2011a). Other potential sources of existing data are listed in EPA's guidance document for QAPPs, Chapter 3 (USEPA, 2002c). These include state and local monitoring programs, state and federal agencies which have quality management for databases of spatial, environmental and natural resources data, EPA's Environmental Information Management System (EIMS) and Environmental Data Registry, as well as published literature and research in trade journals and professional periodical publications.

During the initial discovery and survey of information sources phase an inventory of different types of data available from sources will be developed. The inventory will be a text file matrix, or a spreadsheet prepared and maintained by the Data Collector containing a list of information sources with an accounting of the specific types of data (see data types in Section 2.9.1) provided from each source. In addition the inventory will include the number of data records, range of data or other appropriate indication of the amount of data available from each source. The inventory will be used as a tool to communicate results of data collection activities, assess data needs or gaps, and determine GIS and database requirements.

Similar approaches will be used for retrospective and prospective case study data, but specific direction will be provided in technical directives from the EPA.

1.3.2 Data Collection

Data will be assessed against acceptance criteria (Section 2.9 Non-Direct Measurements) and collected as electronic data files by the Data Collector onto a computer. The data files will be segregated from non-project files and organized by data type, source or source type and by format. Some file manipulation may be necessary to consolidate data, files or file types to simplify file organization and data handling. The activities performed by the Data Collector will be reviewed by the Data Reviewer for concurrence and accuracy.

1.3.3 Data Compilation

Based on the inventory from data survey and collection activities requirements for GIS maps, information database and scenario modeling will be formulated to aid in software selection and development. In addition, record fields, data elements and standard units for data will be defined and overall structure of information organization can be formulated for use of the information. Once a usable form of information format is defined and approved, then the collected data can be compiled. Ideally the data will be compiled directly from source data files into a database or user software, but it may be necessary to manipulate source data and prepare compiled data files in specified format for uploading into or use by selected software. Compiled data files are distinguished from the information database that will be developed for end user use. The Data Collector will inform EPA of the processes that will be used to prepare these compiled data files and will prepare the data files with assistance from the Database Manager. These activities will be reviewed by the Data Reviewer for accuracy.

1.4 QUALITY OBJECTIVES AND CRITERIA

Research activities associated with this study will be conducted in accordance with EPA's Quality System for environmental data and technology (USEPA, 2002a) and Shaw's on-site EPA contract Quality Management Plan (QMP) (Shaw, 2009).

EPA's policy is based on the national consensus standard ANSI/ASQ E4-1994, *Specifications and Guidelines for Environmental Data Collection and Environmental Technology programs (USEPA, 2002)*. This policy recommends applying a graded approach to quality systems according to the specific objectives and needs of the organization and the intended use of the information being collected. Because the information collected for this project will have significant national interest and importance, this project requires the most detailed and rigorous QA and QC for legal and scientific defensibility (EPA Category I).

Current directed project activities involve locating, identifying, surveying and collecting existing data for the identified sites. Other directed activities include preparation of GIS maps and database development with the collected data. Quality objectives and requirements for these activities are described in Section 2.9, Non-Direct Measurements.

Case study activities that may involve collection of direct measurement data as well as modeling efforts for scenario evaluations will be addressed in task-specific QAPPs or QAPP revisions.

All project deliverables will meet EPA's standards of transparency, objectivity, integrity, and utility (USEPA, 2002b). This will be done by providing sources of data, limitations on the data, assumptions, and manipulations or calculations performed so that the work can be reproduced by qualified third parties.

1.5 SPECIAL TRAINING AND CERTIFICATION

The Data Collector shall have an advanced degree in Geology, Environmental Science or

Engineering, and have sufficient experience to understand and evaluate the information collected. This person will also be trained in the use of GIS maps/software and will have sufficient experience in the use of GIS computer software to prepare GIS maps to meet project objectives.

The Data Reviewer must have similar experience as the Data Collector, i.e., an advanced degree in Geology, Environmental Science or Engineering, and have sufficient experience to understand and evaluate the information collected. This person will also be trained in the use of GIS maps/software and will have sufficient experience in the use of GIS information and GIS computer software to evaluate prepared GIS maps against project objectives.

The Database Manager should have an advanced degree in computer science and have sufficient experience to define database requirements to assist in software selection and develop a database that is consistent with project objectives.

1.6 DOCUMENTS AND RECORDS

Documents and records will be managed in accordance with Shaw's EPA contract (OATS) Quality Management Plan (Shaw, 2009).

Typically, the maintenance and eventual turnover of contractual and technical records will be as specified by the EPA. Records prepared and maintained by Shaw which are pertinent only to Shaw will be maintained as specified by Shaw records management procedures.

A filing system is established and implemented for Work Assignments (Was) undertaken within the OATS program. Applicable items included in these files may include:

1) WAs
2) Work Plans
3) Organizational conflict of interest/conflict of business checks
4) Work Plan approvals, WA amendments, and Technical Directives
5) Project-related correspondence
6) WA closeout documents
7) Quality records and documents
8) Procurement Documentation

Hardcopy documents and electronic files generated off-site on behalf of OATS for this WA will be forwarded to the OATS (Shaw, Ada, OK) File Administrator for incorporation into the project file. Ongoing working electronic files being generated off-site as an eventual project product, such as compiled data files, spreadsheets or database, will be forwarded to the EPA or saved to an electronic portal site available to the EPA, or at a minimum, forwarded to the OATS File Administrator for backup within the EPA computer network system on a weekly basis, at a minimum, while being generated, edited or modified.

Hardcopies of source data that will be manually entered into data compilations or database shall be maintained to facilitate checking and documenting checks of entries. This documentation must be retained in the OATS project file.

Shaw will control the review, revision, and distribution of the most recent version of the QAPP. Document control format (Sec. No., Rev. No., and Date) shall appear in the upper right-hand corner of each page of the QAPP. A signed approval form will accompany the QAPP. Any revision to the QAPP shall be circulated to Shaw and EPA project staff for review and approval. Documentation of approval is evidenced by signatures. Final approved version of the QAPP will be distributed by Shaw to all project staff.

Project staff conducting data collection, database development and implementation, and/or GIS activities shall document their work in notebooks or other means approved by the EPA QA Manager per requirements in ORD PPM 13.2, *Paper Laboratory Records* (USEPA ORD, 2006).

As this is a QA Category 1 project, permanent retention of project records is required per Agency Records Schedules 501. All planning documents (QAPPs), data, databases, GIS files, maps, project deliverables, notebooks, correspondence, etc., generated during the course of this project shall be transferred to Susan Mravik, the EPA WAM. They shall be stored in her office at RSKERC until they are transferred to RSKERC's Records Storage Room. At an as yet to be determined time in the future the records will be transferred to a National Archive facility.

Management of project data is described in Sec. 2.10, Data Management.

2.0 DATA GENERATION AND ACQUISITION

2.1 SAMPLING PROCESS DESIGN (EXPERIMENTAL DESIGN)

There are no physical sampling activities identified as part of current project activities.

2.2 SAMPLING METHODS

There are no physical sampling activities identified as part of current project activities.

2.3 SAMPLE HANDLING AND CUSTODY

There are no physical sampling activities identified as part of current project activities.

2.4 ANALYTICAL METHODS

There are no sampling or analysis activities identified as part of current project activities.

2.5 QUALITY CONTROL

There are no sampling or analysis activities identified as part of current project activities, so no

measurement QC activities are currently planned.

2.6 INSTRUMENT/EQUIPMENT TESTING INSPECTION AND MAINTENANCE

Use of instruments or equipment is not planned as part of current project activities.

2.7 INSTRUMENT/EQUIPMENT CALIBRATION AND FREQUENCY

Use of instruments or equipment is not planned as part of current project activities.

2.8 INSPECTION/ACCEPTANCE OF SUPPLIES AND CONSUMABLES

Inspection/acceptance of supplies and consumables on the project is currently anticipated to be minimal and essentially administrative or information handling related. Inspection and acceptance will be indicated and documented by noting approval along with approver's initials and dates on the packing list, receipt, or invoice.

2.9 NON-DIRECT MEASUREMENTS

Non-direct measurements are data collected from existing sources, not directly measured or generated in this project. These data are also referred to as secondary data. Secondary data will be prepared for inclusion in the database following the procedure outlined in OATS Standard Operating Procedure (SOP), SOP-03, *Hydraulic Fracturing Data Handling and Database Management* (Shaw, 2011), which is attached.

2.9.1 Scope of Data Collection

The initial activity for this project is identifying sources and surveying the type of non-direct measurement information available from hydraulic fracturing case studies and other data concerning potentially impacted surface water and groundwater. Shaw will begin with the two currently selected regional study area locations:

- Susquehanna River Basin/Marcellus Shale in Pennsylvania
- Garfield County/Piceance Basin in Colorado.

The scope of this study is further defined below for each location:

- Data from Susquehanna River Basin (PA) area will be from within the entire basin down to the Chesapeake Bay and will be from the year 2005 or later. Data from natural gas/oil well hydraulic fracturing activities will be from wells that have been put into operation in 2005 or later.
- Data from Garfield County (CO) area will be from within the county boundaries, or from locations within watersheds that impinge on the Garfield County booundary, especially in the vicinity of the towns of Silt, Rifle and Battlement Mesa, and will be from the year 2000 or later. Data from natural gas/oil well hydraulic fracturing activities will be from wells that have been put into operation in 2000 or later.

The specific regional study data targeted for collection will be of two basic types: 1) data specific to hydraulic fracturing activities at wells; and 2) data concerning natural or engineered surface water or groundwater in the vicinity of hydraulic fracturing activities, and in the appropriate time frame to register impacts and trends from hydraulic fracturing activities:

Hydraulic fracturing well data will include specific well identification and location information. Data will also include the date of record or activity as well as available specifics such as the following:

- elevation at location
- quantity of water used
- site geologic cross-sections
- water source
- water quality (such as pH, temperature, conductivity)
- water storage, i.e., natural (stream), tanks, pit, engineered impoundment
- well type, i.e., vertical, horizontal or directional
- depth fractured
- fracture directionality
- fracturing pressures
- geologic zone fractured
- site or geologic zone hydrologic information, e.g., hydraulic gradient and conductivity
- quantity of recovered water
- recovered water quality
- fate of recovered water, i.e., treated and discharged (receptor), discharged (receptor), injected (depth), etc

Groundwater and surface water data will include measurement identification information with date and location of measurements as well as available data such as the following:

- water type, i.e., groundwater, stream, lake
- location ID, i.e., well or reading/gauge station ID, etc.
- precipitation data
- location elevation
- surface water level- elevation/depth for lake, reservoir or impoundment
- groundwater level- elevation/depth
- stream stage
- stream flow or discharge rate
- water well – fractured or not fractured
- water well discharge rate
- water quality (such as pH, temperature, conductivity)

The main objective of collecting ground and surface water data is to capture behavioral and temporal trends related to hydraulic fracturing activities.

Additional data targets and additions or changes to the data collection scope may be made by the EPA based on initial survey of available information from sources. These modifications must be documented.

Data will be identified and collected from acceptable sources during information source survey activities. Sources that provide metadata or information that describes the data and their quality criteria will be assessed for suitability and the metadata will be captured or documented. The inventory of different types of data available that is developed during survey activities will be used to define database requirements as well as plan database development and GIS mapping activities.

Case study data collection may be from any number of retrospective and prospective case study nominations as listed below. These studies will typically be more focused and will involve a smaller geospatial scale than the regional studies. Case studies may be defined or amended by EPA based on results from data collection activities and specific direction will be provided on a site specific basis in technical directives.

Nominated Retrospective Case Studies:

- Washington County Pennsylvania (Marcellus Shale)
- Bradford/Susquehanna counties Pennsylvania (Marcellus Shale)
- Wise County Texas (Barnett Shale)
- Kildeer North Dakota (Bakken Shale)
- Las Animas and Huerfano Counties Colorado (Raton Basin)

Nominated Prospective Case Studies:

- Washington County Pennsylvania (Marcellus Shale)
- DeSoto Parish Louisiana (Haynesville Shale)

2.9.2 Geospatial (Locational, Elevational and Temporal) Information Requirements

Locational data should adhere to EPA National Geospatial Data Policy (USEPA, 2005), and any deviations shall be identified. Specific requirements include the following:

- Geo-Referenced Point Data – EPA policy requires geo-referenced coordinates be collected or derived, and appropriately documented in accordance to the adopted EPA/EDSC Latitude / Longitude Data Standard (USEPA, 2006). Locational data will span national to regional to site-specific scales and geospatial references and coordinate systems must be defined for each scale.
- Geospatial Data Accuracy (Locational) – In the absence of program-specific procedures addressing required minimum accuracy for geospatial data, EPA policy requires a minimum accuracy of Tier 5 (USEPA, 2005), which is described in Table 1. However, locational data will span national to regional to site-specific scales and acceptable geospatial accuracy must be defined for each scale or on a site specific basis. Data that

does not meet the established minimum accuracy requirements or data for which the accuracy is not defined and is otherwise deemed acceptable must be qualified and the qualification must be a part of the data set.

Table 1. Geospatial Accuracy Tiers

Tier Level	Accuracy and Precision	Examples of Horizontal Collection Method	Example Program Application
Tier 1	<1 m	Classical Surveying Techniques; plus GPS Carrier Phase Static Relative Position	Definition of contamination boundaries of site
Tier 2	1 – 5 m	GPS Carrier Phase Kinematic Relative Position	Definition of contamination boundaries of site
Tier 3	6 -25 m	GPS Code (Pseudo Range) Standard Position	Stack location; drinking water intake location
Tier 4	26 – 100 m	GPS Unspecified; Photo/GIS Interpolation	Site centroid; large area facility boundary
Tier 5	101 - 200 m	Urban Style Address Matching	Preliminary Site Location
Tier 6	201 - 999 m	Public Land Survey – Sixteenth Section	Prediction of Local Air Dispersion
Tier 7	1000 - 2000 m	Address Matching – Block Face	Batch Geo-coding
Tier 8	2001 - 5000 m	Census Block Centroid	State-level Population Statistics
Tier 9	> 5000 m	Zip Code Centroid	Generalized National Mapping
Tier 10	Unknown	N/A	Relative contextual data

- Geospatial Metadata – EPA policy requires metadata describing geospatial data in accordance with FGDC Content Standard for Digital Geospatial Metadata (FGDC, 1998). This includes temporal and elevation data. Data that does not meet the standard but for which documented information exists providing equivalent metadata information may be used, but this information must be documented and shall become a part of the data set in conformance to the standard. Methods used to construct metadata information must be documented.
- Elevation Data Accuracy – Elevation data for a location is needed at a minimum to recognize impacts on hydrologic conditions in relation to other locations. This data is obtained from standard Global Positioning System (GPS) readings. Minimum accuracy requirements must be established in a manner similar to that for locational (latitude and longitude) information and EPA's Tier levels may be applicable to elevation data as well.
- Map Projections and Scale - The projection associated with maps will be matched to scale and purpose. Regional and national scale map projections will be seamless across the conterminous USA and preserve area. Local scale map projections will be Cartesian

and thus preserve angle and distance.

- For data processing purposes, the Albers 2396 projection will be used for regional and national scale maps, with the following attributes:

  USGS US PROJECTION DATA
  Albers Conic Equal Area
  GRS 1980
  NAD 83

  CENTER PARALLEL 23.0°
  CENTER MERIDIAN -96.0°
  1ST PARALLEL 29.5°
  2ND PARALLEL 45.5°
  FALSE N 0
  FALSE E 0
  UNITS METERS

- For display and visualization (e.g. Google Earth), the web Mercator projection (WGW 1984 Web Mercator) is desirable.
- The UTM projection will be used for local mapping. In the event that the study involves more than one UTM zone, the State Plane coordinate system shall be considered.
- Custom maps may be necessary, and when used, a metadata file will be supplied describing in detail the associated projection.

2.9.3 Acceptance and QA Requirements for Water Quality Data

Water quality data may include a number of standard physical and analytical parameters, such as temperature, pH and conductivity. This data may be obtained from both hydraulic fracturing well information as well as groundwater well and surface water information. Acceptance and QA requirements for this data are described below and summarized in Table 2.

- Water quality data must have been acquired by methods approved by the federal government such as EPA, USGS, DOE or DoD methods, or universally-accepted methods such as ASTM or Standard Methods for Water and Wastewater (SMWW). These methods have defined data quality objectives, i.e., accuracy, precision, detection limits and quantitation limits.
- Data collected by other methods may be used, with EPA approval, but data quality objectives required by the methods must be documented and results must meet the data quality objectives required for the methods used to acquire the data. In addition the data must be qualified as "alternate method." If the methods used and data quality objectives of the methods are notably different than those from methods either approved by the federal government or universally-accepted, than the data may be used with EPA approval, but must be qualified by specifying the differences.
- In some cases it may be difficult to obtain sufficient information about methods used or their data quality objectives to evaluate data with respect to acceptance criteria. In these

cases the data may be used with EPA approval, but the data must be appropriately qualified to indicate the unknown quality information.

- Data for which the methods used to obtain the data and the data quality objectives or data quality indicator results are not defined or otherwise indicated may be used with EPA approval, but results must be qualified as "quality unknown."
- Any data that has been identified as not meeting the quality objectives of the methods used to collect the data will be rejected.

Table 2. Water Quality Data - Data Quality Objectives and Data Qualifications

Measurement	Method	Data Quality Objectives (DQO)	Acceptance	Data Qualifier
Water quality parameters, e.g., pH, temperature, conductivity	EPA, USGS, ASTM, SMWW, federal government approved or universally accepted	Method requirements	Meets DQO	None
	Other methods that are consistent with government approved or universally accepted methods	DQO consistent with government approved or universally accepted methods	Meets DQO and EPA approval	Alternate Method
	Methods notably different from government approved or universally accepted methods	Method requirements	Meets method requirements and EPA approval	Difference(s) defined
	Method or DQO Unavailable or undefined	Method or DQO Unavailable or undefined	EPA Approval	Missing Information Identified
	Unavailable or undefined	Unavailable or undefined and no indication that data did not meet method requirements	EPA Approval	Quality Unknown

2.9.4 Acceptance and QA Requirements for Other Measurement Data

Other measurement data include hydraulic fracturing operation data, groundwater level measurements, precipitation data, stream stage measurements, stream and water well discharge rates, etc. Methods for acquiring or generating this data may include those approved by the federal government such as EPA, and USGS, universally-accepted methods such as ASTM or may be standard operating procedures (SOPs) used by the source entity or even instrument/device manufacturers operating instructions (IMOI). The level to which the methods used and data quality objectives, i.e., accuracy, precision, etc., are defined or even available for this data may be variable. Acceptance and QA requirements for this data are described below and summarized in Table 3.

- Data must have been acquired by documented and approved methods with defined data quality objectives, i.e., accuracy and precision, etc., that are consistent with methods either approved by the federal government or universally-accepted, if such methods exist.
- If the methods used and data quality objectives of the methods are notably different than those from methods either approved by the federal government or universally-accepted, if such methods exist, then the data may be used with EPA approval, but must be qualified by specifying the differences.
- If sufficient information about the methods used or the quality objectives of the methods are not available to evaluate data with respect to the acceptance criteria, the data may be used with EPA approval, but the data must be appropriately qualified to indicate the unknown information.
- Data for which the methods used to obtain the data and the data quality objectives or data quality indicator results are not defined or otherwise indicated may be used with EPA approval, but results must be qualified as "quality unknown."
- Any data that has been identified as not meeting the quality objectives of the methods used to collect the data will be rejected.

Table 3. Other Measurements - Data Quality Objectives and Data Qualifications

Measurement	Method	Data Quality Objectives (DQO)	Acceptance	Data Qualifier
HF operating parameters, quantities	ASTM, USGS, SOP, IMOI	Method requirements or Procedure requirements with DQO consistent with government approved or universally accepted method(s)	Meets DQO	None
Groundwater level	EPA, SOP, IMOI			
Stream discharge rate	USGS, SOP, IMOI			
Water well discharge rate	ASTM, USGS, SOP			
Stream stage	USGS, SOP, IMOI			
HF operating parameters, quantities, Groundwater level, Stream discharge rate, Water well discharge rate, Stream stage	Methods notably different from government approved or universally accepted methods, if any exist	Method requirements	Meets method requirements and EPA approval	Difference(s) defined
	Method or DQO unavailable or undefined	Method or DQO unavailable or undefined	EPA approval	Missing information identified
	Method unavailable or undefined	DQO unavailable or undefined and no indication that data did not meet method requirements	EPA Approval	Quality Unknown

2.9.5 Other Acceptance Requirements for Data

Non-Direct measurement data requirements are described below:

- Methods used to discover and collect non-direct measurement data must be documented and generally accepted from both a technical and QA standpoint. The quality objective is to assure that discovery, identification and selection efforts are thorough, comprehensive and unbiased. Current methods are to investigate known sources that have been recommended by EPA and documented in this QAPP as well as secondary sources derived from these sources and evaluate available data for acceptance as outlined in Section, 2.9. Other methods may be searching for information in specific professional society journals, which shall be documented as a note identifying the journal, how it was selected and the time frame covered by the journal dates searched. This documentation shall be retained in the OATS project file.
- Data shall meet the spatial and temporal requirements as described in Section 2.9.2.
- In general, data must be from a known or recognized and accepted source, such as; a state or federal government agency (e.g., USGS, COGCC, PADNR); a peer reviewed publication; from a source with a documented Quality Management System (QMS) including a QMP; or from a source with documented evidence that generally accepted methods were used in generation of data. See requirements for specific data types in Sections 2.9.2, 2.9.3 and 2.9.4.
- The source of non-direct measurement data used in the study must be identifiable and recorded along with the data as a reference. Copies of reference source data documentation must be retained in the project file and be readily retrievable.
- The reference source of data will be an element of the data record and will be included with the compiled data and database, at a minimum.
- Data qualifiers from either the source or from the Data Collector as to its quality or applicability will be an element of the data record and will be included with the compiled data and database, at a minimum.
- Data collected must be independently reviewed and approved by the Data Reviewer. This review must be documented.

2.10 DATA MANAGEMENT

Information from data collection activities will be collected and saved on computers from network websites and other sources using standard commercially available internet, internet search and data handling/manipulation software, such as Word, Excel, or other EPA approved software.

Data collection computer system requirements are specified below.

- The computers must be Shaw corporate computers with appropriate and active security system software installed including firewall and web protection with regular scans for viruses and spyware.
- The computers shall not have any software installed that is not approved by Shaw Information Technology (IT) Services.

- The data collected will be secure, either by encryption of computer hard-drives when used off-site or on a secure computer network such as the ShawNet intranet system, which has additional server based security and scheduled backup systems.
- The computers will also be rebooted daily to ensure system updates are installed in a timely manner.
- Computer connections to the internet will be made using Shaw computers equipped with active and Shaw-approved anti-virus and anti-phishing software and that are compliant with Shaw corporate policy regarding computer safety.

The data collected will subsequently be used to prepare GIS maps and assemble a database of information for the study. Ideally data can be compiled directly from collected source data files into selected database software, but it may be necessary to manipulate source data and prepare compiled data files in specified format for uploading into or use by selected GIS and database software. Compiled data files are distinguished from the information database that will be developed for end user use.

Data management requirements are the following:

- Project files will be segregated from non-project files on an electronic portal site or storage location by a specified storage drive or drive subdirectory location.
- Source data files will be stored on an electronic portal site or in a designated location on a computer and organized by data type, source or source type and by format. File manipulations performed to consolidate data, files or file types to simplify file organization and data handling will be reviewed by the Data Reviewer to assure the integrity of the data file or data set is intact. This review must be documented. Approved data files will be designated as such in the file name and will be saved to a segregated location. The approved designation will be consistently applied to source data files and will be readily identifiable. Data manipulation such as that for conversion of geospatial reference coordinates or measurement data units for consistency must be approved by EPA, documented, and the process must be reviewed and checked for errors by the Data Reviewer.
- The EPA will be informed by the Data Collector of the format and method for preparing compiled data files. This applies to the process of preparing compiled data files from data set files (source data files) for use by or uploadable into GIS or database software. Data manipulation steps as well as prepared (final) compiled data files must be reviewed and approved by the Data Reviewer. This review must be documented. Approved compiled data files will be designated as such in the file name and will be saved to a segregated location. A file naming convention that is consistent and renders file types and QA approval status readily identifiable will be used to identify compiled data files.
- Compiled or constructed data files must have 100% of manual data entries checked independently against source documentation for correct transcription. This check must be documented. If errors are discovered, they will be presented to the Data Collector for

concurrence and resolution or correction. Corrections will be reviewed by the Data Reviewer for approval. If concurrence between the Data Collector and Data Reviewer cannot be achieved on potential errors, the Shaw QAO will be consulted for resolution. Data files that have been checked and approved for manual entries will be designated as such in the file name using a consistent and readily identifiable naming convention.

- Compiled data files must have ten percent (10%) of electronic data transfers checked against source data for correct transcription. This check must be documented. If errors are discovered, they will be presented to the Data Collector for concurrence and resolution or correction. Corrections will be reviewed by the Data Reviewer for approval. If concurrence between the Data Collector and Data Reviewer cannot be achieved on potential errors, the Shaw QAO will be consulted for resolution. The discovery of a transcription error will trigger an increase in the percent data transfer checks to twenty-five percent (25%) for the file. Data files that have been checked and approved for electronic data transfers will be designated as such in the file name using a consistent and readily identifiable naming convention.
- Source documentation used for manual entries must be maintained in a project file and managed as prescribed in Section 1.6, Documents and Records.
- Collected and compiled data files must be backed up on an independent computer, network server or independent data storage device daily if changes, additions or modifications have been made.
- Products from data collection, such as the source data inventory and compiled data files will be forwarded to the EPA or saved to an electronic portal site available to the EPA, or at a minimum, forwarded to the OATS File Administrator for backup within the EPA computer network system on a weekly basis if changes, additions or modifications have been made.
- Data collected off-site of the OATS contract intended for input into an onsite OATS/EPA database will be transferred electronically to the Database Manager by email or saved to an electronic portal site or common computer drive available to the developer.

The process for preparing the raw data for upload to the database is given in OATS SOP-03 (Shaw, 2011). The general process is briefly outlined as follows.

- Raw Data is provided to the Database Manager and stored on a local computer file folder, folders are named by month and date data are received;
- A sample format file or other translation matrix, which indicates how the required fields (USEPA, 2003, 2008a) are to be populated, if the data in the source files are absent or inconsistent with the EPA Valid Values (EPA, 2008b) is developed by the Technical Lead working with the Database Manager.
- If necessary, a script is developed to extract or transform the data to a form compatible with the required EDD format. The script developed will be available for review;
- Output is loaded into an EDD for upload into the database. ;
- The EDD is reviewed to assure data handling process(es) performed as intended.

- An electronic data processor (EDP) tool is used to check the data against the EDD format requirements and load the data into the database;
- The final database entries are spot-checked for correctness.
- Data files reside on local network drive, the network drives are periodically backed up on tape drives.

3.0 ASSESSMENT AND OVERSIGHT

3.1 ASSESSMENTS AND RESPONSE ACTIONS

Assessors do not have stop work authority, however, they can advise the EPA Work Assignment Manager (WAM) if a stop work order is needed in situations where data quality may be significantly impacted. The EPA WAM makes the final determination as to whether or not to issue a stop work order. The EPA WAM may consult with the EPA technical lead to assist with this determination. . This does not preclude the WAM from working through the proper contract channels to accomplish this activity.

3.1.1 Assessments by Shaw

Technical Systems Assessments (TSA) will be used to monitor project activities for implementation and conformance to the requirements of this QAPP and related Quality documents. The TSA will include assessment of data collection activities, documentation, quality checks, record management and reporting.

A TSA will be performed quarterly by the Shaw QAO. Results of the TSA will be reported to the OATS Program Manager (PM), EPA Quality Assurance Manager (QAM), and EPA WAM. Nonconformances will be identified as findings, recommendations or observations. Corrective actions for findings will be developed, documented and proposed by the OATS PM within 15 days of TSA report issuance to the Shaw QAO, EPA QAM, and EPA WAM for concurrence and approval. The Shaw QAO is responsible for ensuring resolution of findings. The Shaw QAO will monitor implementation and completion of corrective actions. After all corrective actions have been implemented and confirmed to be completed, the Shaw QAO shall send documentation to the OATS PM, EPA QA Manager and EPA WAM that the TSA is closed. TSA reports and responses shall be maintained by the OATS File Administrator in the project file.

3.1.2 Assessments by EPA

Technical Systems Audits (TSAs) and Audits of Data Quality (ADQs) will be conducted early in the project by the EPA to allow for identification and correction of any issues that may affect data quality. Detailed checklists, based on the procedures and requirements specified in this QAPP, will be prepared and used during these TSAs. ADQs will be done on a representative sample of the data. These audits will be conducted with contract support from Neptune and Co.

Audit reports will be prepared by the QA support contractor, which will be reviewed and

approved by the EPA QA Manager prior to release. Audit reports shall be sent and copied to the EPA PO and supervisor, EPA Technical Lead, EPA WAM, GWERD Division Director, and OATS PM. Specific actions will be identified in the reports. For assessments that identify deficiencies requiring corrective action, the OATS PM must provide a written response, within 10 working days, to each finding and observation to the EPA QA Manager and WAM, which shall include a plan for corrective action and schedule. The EPA QA Manager will review the written response to determine its appropriateness and provide feedback to the OATS PM, if necessary. The OATS PM is responsible for ensuring resolution of audit findings. The EPA QA Manager will monitor implementation and completion of corrective actions. After all corrective actions have been implemented and confirmed to be completed, the EPA QA Manager shall send documentation to the same parties that received the audit report that the audit is closed. Audit reports and responses shall be maintained by the OATS File Administrator in the project file and the EPA QA Manager in the QA files, including the QLOG database.

3.2 REPORTS TO MANAGEMENT

Meetings with the EPA will be held as scheduled or directed by the EPA in the Technical Directive (TD). Meeting invitees will include the EPA Project Officer, EPA WAM, EPA Technical Lead, EPA QA Manager, Shaw PM and Shaw Database Manager with the Shaw Data Collector, Shaw Data Reviewer and Shaw QAO participating by teleconference. If the Shaw PM does not participate in these meetings he shall be supplied the information provided in these meetings in an update document, notes or verbal briefing by the Shaw QAO.

In addition Monthly Status Reports will be made by the Shaw PM to the EPA Project Officer, EPA WAM and EPA QA Manager.

A Quality report to management (Shaw PM) will be made quarterly in conjunction with the TSA report. The Quality/TSA report will discuss the results of the TSA, status of the quality management program and other related issues.

4.0 DATA VALIDATION AND USABILITY

4.1 DATA REVIEW, VERIFICATION AND VALIDATION

As stated in Sections 2.9 and 2.10, products of project data collection activities will be reviewed and approved from a technical and QA standpoint by the Data Reviewer using criteria in 2.9.2, 2.9.3, 2.9.4 and 2.10. This is a documented verification process in which the existing data mined from sources will be reviewed for the following:

- Applicability of the data for the project objectives will be reviewed by examining the source documentation for concurrence with the collector's interpretation and use.
- Usability of the data will be assessed based on the documented or perceived quality of the source, metadata, available data quality indicators, measurement performance or

methods of generation.

- Completeness in collection of data and quality related information available from the source
- Accuracy in interpreting and assembling the information in a usable format
- Format in which the data is organized and compiled with respect to EPA direction, user understanding, ease of use, and achieving project goals.

In addition, Shaw will use procedures described in 2.10 to verify that data transcribed from existing data sources have been transcribed accurately. Compiled or constructed data files must have manual data entries checked independently against source documentation for correct transcription. Compiled or constructed data files must have ten percent (10%) of electronic data transfers checked against source data for correct transcription. These checks must be documented.

4.2 VERIFICATION AND VALIDATION METHODS

The Shaw Data Collector will request a review be performed of compiled data files that have undergone his review and determination as to their acceptability. The review request directed to the Data Reviewer will be documented and will include a description of the data files to be reviewed. The Data Reviewer will have access to the project data files in a restricted electronic portal location and will review them against the criteria in 2.9.2; 2.9.3, 2.9.4 and 2.10 to determine their acceptability, and document the review. The results will be reported to the Shaw Data Collector who shall then identify this data as validated as described in Section 2.10, if it was found acceptable. Only verified and validated data will be included in data compiled for database entry.

4.3 RECONCILIATION WITH USER REQUIREMENTS

The development of this information database will facilitate use of the information to achieve overall study goals, however, this is an early stage of the study process. The information database is intended to provide a basis for additional activities to continue the study. To that end, the product of this work should provide a thorough and open framework to support further work. Limitations on the use of the collected information will be based on the final outcomes of the information development process.

Project deliverables will be reviewed internally by Shaw and externally by the EPA Technical Lead and WAM to ensure they meet EPA's requirements. Shaw shall describe data quality and data limitations in its project deliverables so that later data users may determine if the data are of sufficient quality for their use.

5.0 REFERENCES

ANSI/ASQC E4-1994, *Specifications and Guidelines for Environmental Data Collection and Environmental Technology Programs,* ASQ Quality Press, Milwaukee, WI, www.asq.org.

FGDC, 1998, *Content Standard for Digital Geospatial Metadata,* FGDC-STD-001-1998, Federal Geographic Data Committee, June 1998.

Shaw, 2009, *On-Site Analytical and Technical Support Services Quality Management Plan for the NRMRL/GWERD, QA-001, March 01,2009, Edited October 10, 2010.*

Shaw, 2011, *Hydraulic Fracturing Data Handling and Database Management, OATS SOP-03, December 13, 2011.*

USEPA, 2002a, *Overview of the EPA Quality System for Environmental Data and Technology,* U.S. Environmental Protection Agency, EPA/240/R-02/003, November 2002.

USEPA, 2002b, *EPA's Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency (Information Quality Guidelines),* U.S. Environmental Protection Agency, EPA/260R-02-008, October 2002.

USEPA, 2002c, *Guidance for Quality Assurance Project Plans,* U.S. Environmental Protection Agency, EPA QA/G-5, EPA/240/R-02/009, December 2002.

USEPA, 2003, Electronic Data Deliverable (EDD) Specification Manual, Version 1.1, USEPA, Region 5; June

USEPA ORD, 2006, *Paper Laboratory Records,* Office of Research and Development, Policies and Procedures Manual, Chapter 13.2, December 2006.

USEPA, 2005, *National Geospatial Data Policy,* U.S. Environmental Protection Agency, CIO Policy Transmittal 05-002, August 2005.

USEPA, 2006, *Latitude/Longitude Data Standard,* Standard No.: EX000017.2, EPA/EDSC Data Standard, January 6, 2006.

USEPA, 2008a, Comprehensive Electronic Data Deliverable (EDD) Specification Manual, Version 2.0, USEPA, Region 5; August, 2008

USEPA, 2008b, Comprehensive Electronic Data Deliverable (EDD) Valid Value Appendix, Specification Manual, Version 2.0, USEPA, Region 5; August, 2008

USEPA, 2011a, *Plan to Study the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources*, Office of Research and Development, U.S. Environmental Protection Agency, EPA/600/11/001, November, 2011. Available online at: http://water.epa.gov/type/groundwater/uic/class2/hydraulicfracturing/index.cfm

USEPA, 2011b, *Hydraulic Fracturing Support*, Work Assignment No. 2-10 (HF), EPA Contract No. EP-C-08-034, February 7, 2011.

ATTACHMENT

STANDARD OPERATING PROCEDURE
HYDRAULIC FRACTURING DATA HANDLING AND DATABASE MANAGEMENT
OATS SOP-03

1.0 SCOPE/OVERVIEW

This SOP details the required elements for receiving, manipulating, loading and storing of hydraulic fracturing data for GWERD Hydraulic Fracturing Research case studies. The QAPP that covers these activities is entitled "Data Collection/Mining for Hydraulic Fracturing Case Studies." The intent of the SOP is to document the approved process including QA/QC requirements for preparing data that has been received from various sources, and importing it into an EQuIS-built database.

2.0 RESPONSIBILITIES

2.1 **Database Manager** – Receive raw data intended for use in Hydraulic Fracturing Research case studies and process the data for loading into a project database. Notify the Shaw Technical Lead of problems associated with the raw data package. Move the raw data through the process and document the steps involved from data receipt all the way to importing the data to the EQuIS database. Respond to reviewer's comments, resolve issues and perform corrective actions in a timely manner to expedite the data receiving / loading process. Assure that processed data progress through the review process in a timely manner. Assist with the review process by performing raw data QC and transcription checks as needed or requested.

2.2 **Shaw Technical Lead / Data Collector -** Communicate instructions to the Database Manager regarding handling of source data files and processing of source data to be compliant with EPA Electronic Data Deliverable (EDD) requirements prior to loading into the EQuIS database. Perform final review of uploaded data if needed in a timely manner and document the review and results. Notify the Program Manager, as appropriate, of any problems associated with the data or the data receiving / loading process.

2.3 **Data Reviewer** – Perform review of processed data and the data handling process steps from the time the data is received all the way through importation into the data warehouse. Document reviews and resolve review issues with the Database Manager.

3.0 PROCEDURE

The data loading process is initiated upon receipt of a data file to be loaded into a database. A data handling process for the data file is developed based on decisions regarding the extraction of data from complex source fields, how to translate codes from the source files to EPA Valid Values (EPA, 2008b) and what values to use where required fields of the electronic data deliverable (EDD) are not present in the source file. The specific transformations/translations are performed on a case by case basis and will be documented as part of the discovery and review process. as described in the QAPP (USEPA, 2011).

The **Database Manager** plays a central role in guiding the data through the data handling process. The work flow from receipt of data to upload to the EQuIS database is illustrated in the flowchart shown in Figure 1 and described below.

1. Raw Data are provided to the **Database Manager** either through a secure FTP Site, or in an email, or as a package from a Shaw data sharing site (Shaw XNet);
2. On receipt, the data are downloaded into a local computer file folder (e.g. HF Meeting 05.11), named by month and date data are received;
3. The data is reviewed and compared to EPA Region 5 (R5) EDD requirements by the **Technical Lead** working with the **Database Manager** to determine if data needs to be extracted or transformed, if data is absent for required fields (EPA, 2003, 2008a), or if data in the source files are inconsistent with the EPA Valid Values (EPA, 2008b)**.**
4. A sample format file or other translation matrix, which indicates how the required fields are to be populated, if data in the source file(s) are absent or inconsistent with the EPA Valid Values is developed by the **Technical Lead** working with the **Database Manager**.
5. If necessary, the **Database Manager** assures that a script is developed to extract or transform the data to a form compatible with the required EPA R5 EDD format (EPA, 2003, 2008). Needed scripts will be developed by the **Database Manager** in consultation with Shaw personnel familiar with writing database scripts;
6. The script developed will be available for review;
7. The EDD script product is reviewed to assure the script process(es) perform as intended. The review is documented.
8. The source file code definitions in the EDD product are compared and translated to EPA Valid Values (EPA, 2008b) according to the sample format file or translation matrix.
9. Constant data for required fields that are absent in the EDD are added to conform to EPA R5 EDD requirements per the sample format file or translation matrix.
10. A QA review of the EDD is performed to assure code definitions are translated and required fields are populated accurately as described in the QAPP. The review is documented.
11. The EDD file is given an appropriate name based on the format file and source file, and saved in a local computer folder;
12. The EQuIS 5 Professional Electronic Data Processor (EDP) tool is used to load the saved text files into the EQuIS database;
13. The final database entries are spot-checked for correctness as determined by decisions made in steps 3 and 4 above. The spot-check process including details about which data was checked and how it was checked is documented.
14. Data files reside on local shared **L Drive** under (L:\Lab\CSMOS\Hydraulic Fracturing 2011\Chem), which will be backed up every day and also periodically backed up on tape drives and stored in three different locations. The EQuIS database resides in Oracle, also on the network drive.
15. All folders mentioned above, which are on the **Database Manager's** local computer, are synchronized with the **L Drive** under (L:\Lab\CSMOS\Hydraulic Fracturing 2011).

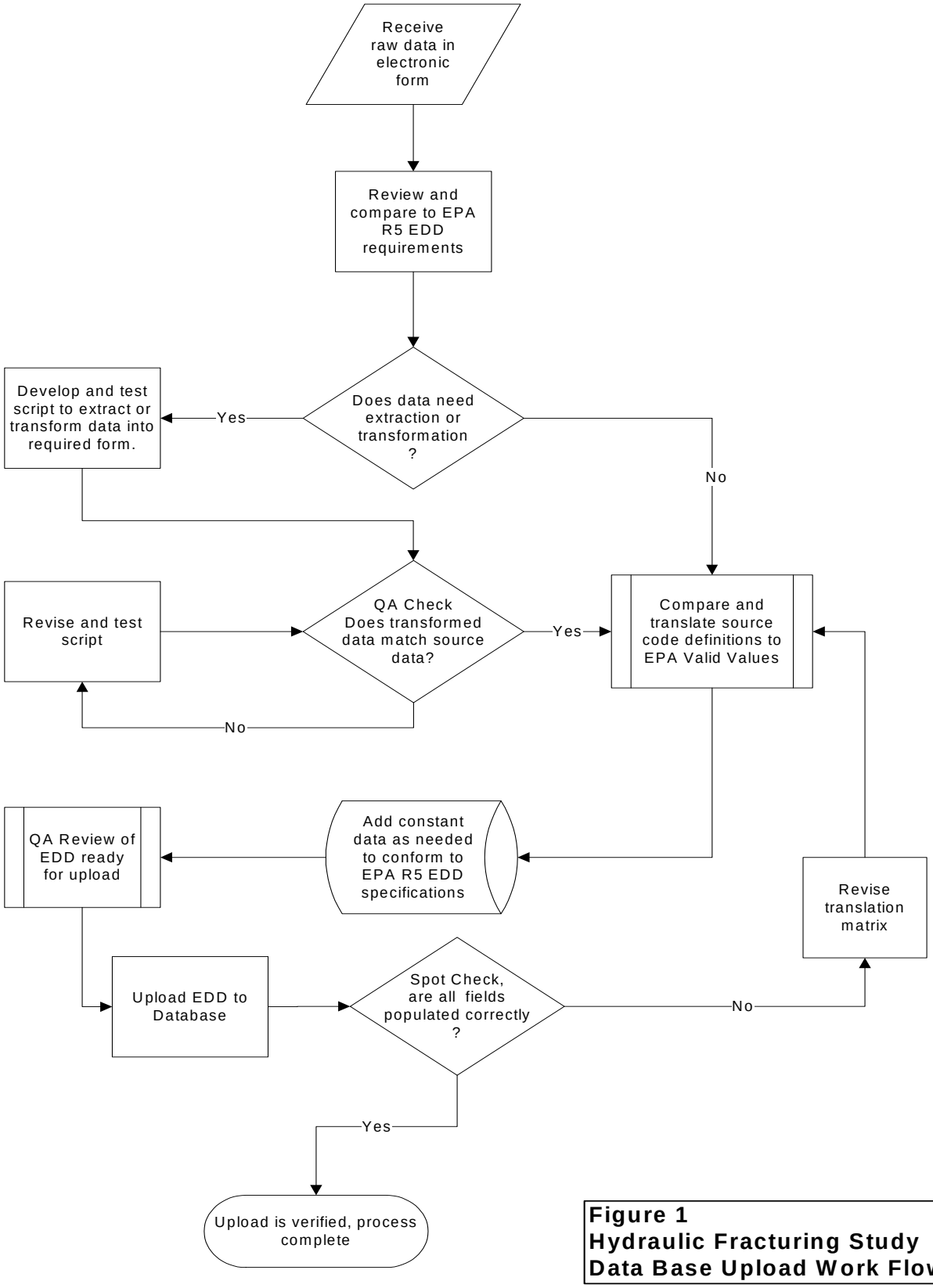

Figure 1
Hydraulic Fracturing Study
Data Base Upload Work Flow

4.0 QUALITY ASSURANCE

Throughout the process, all steps will be carefully documented in dedicated notebooks according to the procedures in the GWERD QAPP for Data Collection/Mining for Hydraulic Fracturing Case Studies, 2011.

For each data file that is loaded into a database the steps in the data handling process (Section 3.0) must be independently verified by the **Data Reviewer** and this verification must be documented. Because of the large number of records that may be involved in this process, the verification will be performed by randomly spot checking end database entries against received data to assure the data handling process has performed as intended. The spot-check verification should include data that represents all data processes performed and span across data fields or types and the entire data set. The review should be such that all data handling process steps across the entire data set are verified on as many different data types as is reasonably achievable.

5.0 REFERENCES

U.S. Environmental Protection Agency , 2011, GWERD Quality Assurance Project Plan for: Data Collection/Mining For Hydraulic Fracturing Case Studies, May 2011.

U.S. Environmental Protection Agency, 2003, Electronic Data Deliverable (EDD) Specification Manual, Version 1.1, USEPA, Region 5; June

U.S. Environmental Protection Agency, 2008a, Comprehensive Electronic Data Deliverable (EDD) Specification Manual, Version 2.0, USEPA, Region 5; August

U.S. Environmental Protection Agency, 2008b, Comprehensive Electronic Data Deliverable (EDD) Valid Value Appendix, Specification Manual, Version 2.0, USEPA, Region 5; August

Memorandum

304 DIRECTORS DRIVE, KNOXVILLE, TN 37923 • 865.690.3211 • FAX: 865.694.9573

To: Steve Vandegrift – EPA QAM, Ada, OK
Susan Mravik – EPA WAM, Ada, OK
Sujith Kumar – Shaw OATS Program Manager – EPA Contract, RSKERC, Ada, OK

From: Duane Root – Shaw Project QAO, Knoxville, TN

CC: Shauna Bennett – OATS QC Coordinator, Ada, OK
Jonathan Shireman – Shaw Technical Lead, Knoxville, TN

Date: December 21, 2011

Re: Changes to the Data Collection/Mining for Hydraulic Fracturing Case Studies QAPP

Below is a list of changes made to the QAPP in the attached version (Revision 1). The proposed QAPP revision is provided for your review, comment and acceptance.

1. Revisions from corrective actions (CAs) that resulted from the internal TSA conducted in August and were proposed in Shaw CA report dated 9/16/11.
    a. Revised section 2.10 (3rd bullet) – to remove requirement that the Shaw data collection computer remain in a secure Shaw facility
    b. Revised section 2.10 (5th bullet) – to remove requirement that internet connection be made through Shaw intranet
2. Revisions from CAs that resulted from the EPA audits conducted in Knoxville and Ada in September/October and were proposed in Shaw CA report dated 10/27/22
    a. Revised section 2.9.1 – to remove discussion of "Control Area" data collection (3rd paragraph)
    b. Revised section 2.9 – to add a reference to the approved Shaw SOP (OATS-SOP-03) for Data Handling
3. Revisions that should be made to the QAPP based on the approved Shaw Data Handling SOP
    a. Changed QAPP project organization chart (section 1.1) "Database Developer" to "Database Manager" and changed Chaitanya to Rob Earle
    b. Revised section 2.10, Data Management – to add a description of data handling for data upload into database based on SOP (Shaw OATS-SOP-03) at the end of the section.
    c. Revised section 5.0, References – to add the OATS-SOP-03 citation to the list as well as USEPA references from text addition in 3b above.

4. Revision that resulted from the internal TSA conducted in November and recommended in the Shaw TSA report dated 12/15/11
    a. Changed the QAPP in Sections 1.3.3 (second to last sentence) and 2.10 (8th bullet, first sentence) to state that the EPA will be informed of the required data handling processes rather than requiring EPA approval of data handling processes.
5. Revision based on changes in the nominated retrospective and prospective case studies that have occurred since the QAPP Revision 0 was issued.
    a. Revised section 1.3 – to remove the former nominated retrospective and prospective case study locations and replace with a reference to discussion in section 2.9.1.
    b. Revised section 2.9.1 – to update the nominated retrospective and prospective case study areas to the current nominations.
6. Revisions based on the finalization of EPA's Draft Plan to Study the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources.
    a. Revised text references as needed and revised the citation in the reference list to reflect the final plan title, designation and date.

GWERD QUALITY ASSURANCE PROJECT PLAN

TITLE: DATA COLLECTION/MINING FOR HYDRAULIC FRACTURING CASE STUDIES

WORK ASSIGNMENT NO.: WA-HF-2-10

TECHNICAL DIRECTIVE: 7HF101HF

QA ID NO.: G-15952 QA Category: I

CONTRACTOR: Shaw Environmental & Infrastructure, Inc., #EP-C-08-034

REVISION: 1

DATE: January 10, 2012

_____ _____
EPA Technical Research Lead Date

_____ _____
Shaw Technical Lead Date

EPA APPROVALS:

_____ _____
Project Officer Date

_____ _____
GWERD QA Manager Date

SHAW APPROVALS:

_____ _____
Shaw Program Manager Date

_____ _____
Shaw QA Officer Date

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 quality assurance/quality control activities and technical requirements that will be used during the research study. EPA plans to publish the research study results in a draft report, which will be reviewed by the EPA Science Advisory Board. The final research report would be considered the official Agency dissemination. Mention of trade names or commercial products in this planning document does not constitute endorsement or recommendation for use.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

DISTRIBUTION LIST

Cynthia Paul – EPA GWERD, Ada, OK
Susan Mravik - EPA GWERD, Ada, OK
Steve Vandegrift- EPA GWERD, Ada, OK
Dr. David Jewett - EPA GWERD, Ada, OK
Dr. Stephen Kraemer - EPA ERD, Athens, GA
Sujith Kumar – Shaw, EPA Contract, Ada, OK
Dr. Duane Root – Shaw, Knoxville, TN
Jonathan Shireman – Shaw, Knoxville, TN
Victor Murray – Shaw, Ada, OK
Chaitanya Nellutla – Shaw, Denver, CO
Rob Earle – Shaw, Ada, OK
Vickie Grissom – Shaw File Administrator, Ada, OK

1.0 PROJECT MANAGEMENT

1.1 PROJECT / TASK ORGANIZATION

Project Organization is as shown below in Figure 1.

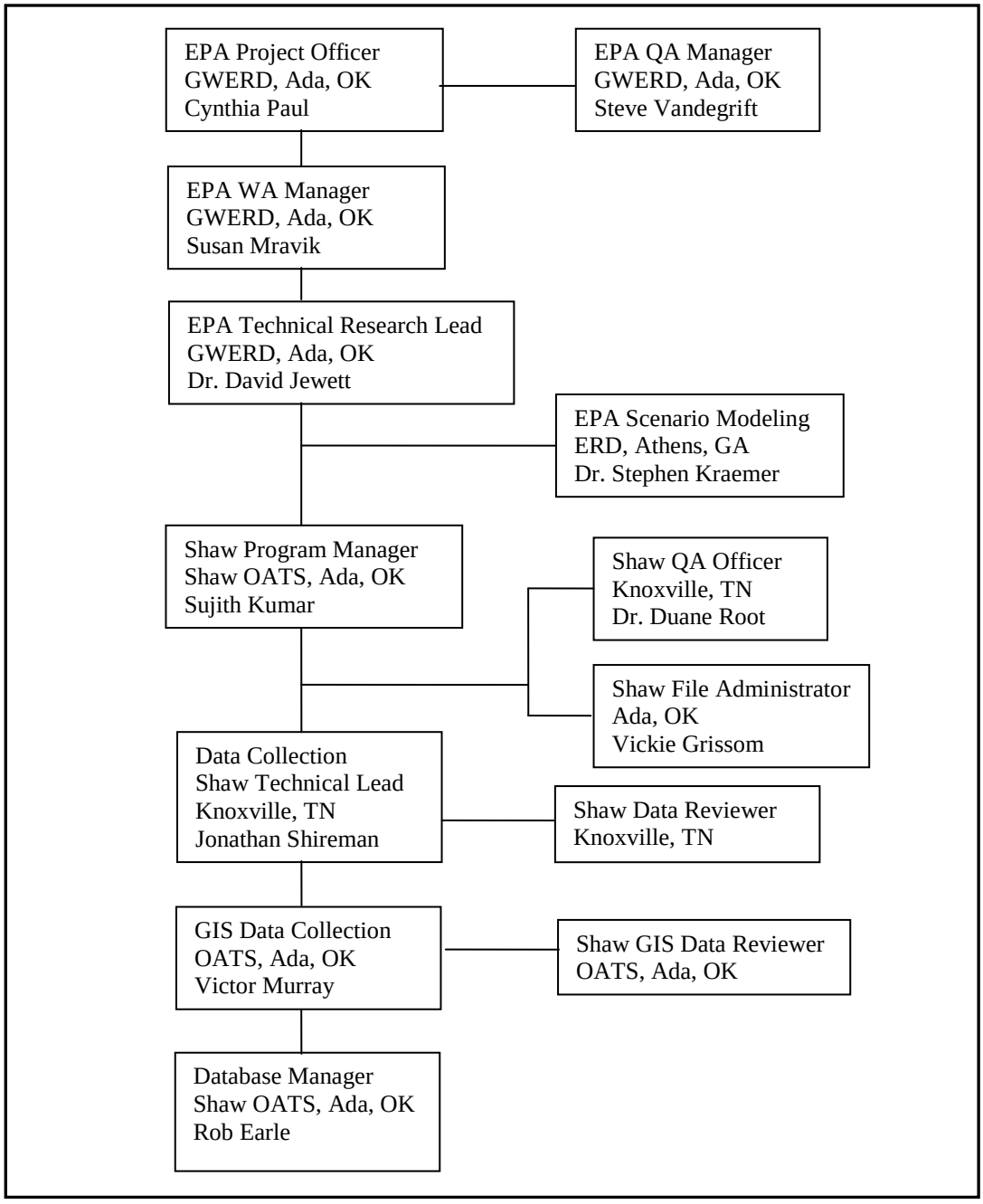

Figure 1. Project Organization Chart

Dr. David Jewett, EPA Technical Research Lead, U.S. Environmental Protection Agency, Ground Water and Ecosystems Restoration Division, Ada, OK. Responsible for technical oversight, technical review of the Quality Assurance Project Plan (QAPP), ensuring project goals are achieved, and review/approval of project deliverables.

Dr. Stephen Kraemer, EPA Scenario Modeling, U.S. Environmental Protection Agency, Ecosystems Research Division, Athens, GA. Responsible for hydraulic fracturing scenario modeling (to be addressed in a separate QAPP) and for conforming to approved QAPP requirements.

Ms. Cynthia Paul, EPA Project Officer, U.S. Environmental Protection Agency, Ground Water and Ecosystems Restoration Division, Ada, OK. Responsible for contract oversight, review/approval of the QAPP,

Ms. Susan Mravik, EPA Work Assignment Manager (WAM), Environmental Protection Agency, Ground Water and Ecosystems Restoration Division, Ada, OK. Responsible for providing technical direction, review of the QAPP and review/approval of project deliverables, management of project records, QA, and resolution of QA issues.

Mr. Steve Vandegrift, EPA QA Manager, Environmental Protection Agency, Ground Water and Ecosystems Restoration Division, Ada, OK. Responsible for QA review/approval of the QAPP, conducting audits, and QA review/approval of the final product.

Mr. Sujith Kumar, Shaw Program Manager, Shaw Environmental & Infrastructure, Inc., Ada, OK. Responsible for management of Shaw project activities, review and approval of the QAPP, assuring implementation of the approved QAPP, and review and approval of Shaw project deliverables.

Mr. Jonathan Shireman, Shaw Technical Lead and Data Provider, Shaw Environmental & Infrastructure, Inc., Knoxville, TN. Responsible for data collection, evaluation of data against acceptance criteria, preparation of compiled data files, and reviewing and conforming to approved QAPP requirements.

Mr. Victor Murray, Shaw GIS Analyst, Shaw Environmental & Infrastructure, Inc., Ada, OK. Responsible for GIS data collection, evaluation of GIS data against acceptance criteria, preparation of compiled GIS data files, and reviewing and conforming to approved QAPP requirements.

Mr. Rob Earle, Shaw Database Manager, Shaw Environmental & Infrastructure, Inc., Ada, OK. Responsible for database development and management and for conforming to approved QAPP requirements. Mr. Earle will be assisted as needed by Mr. Chaitanya Nellutla, Shaw, Denver, CO.

Dr. Duane Root, Shaw QA Officer, Shaw Environmental & Infrastructure, Inc., Knoxville, TN. Responsible for preparation of the QAPP, implementation of approved QAPP, conducting project TSAs, reviewing corrective actions and preparing quarterly reports to management on the status of the QMS with respect to project goals.

Ms Vickie Grissom, Shaw File Administrator, Ada, OK. Responsible for maintaining project files as described in the QAPP.

Shaw Data Reviewer, Shaw Environmental & Infrastructure, Inc., Ada, OK/Knoxville, TN. The Data Reviewer shall have experience and training as described in Section 1.5, Special Training and Certification. The Data Reviewer will be responsible for reviewing collected data against acceptance criteria, reviewing data and data file manipulations for accuracy, and conforming to approved QAPP requirements.

1.2 PROBLEM DEFINITION / BACKGROUND

As natural gas production has increased, so have concerns about the potential environmental and human health impacts of hydraulic fracturing in the United States. Hydraulic fracturing, which involves the pressurized injection of water, chemical additives, and proppants into a geologic formation, induces fractures in the formation that stimulate the flow of natural gas or oil, thus increasing the volume of gas or oil that can be recovered from coalbeds, shales, and tight sands—the so-called "unconventional" reservoirs. Many concerns about hydraulic fracturing center on potential risks to drinking water resources, although other issues have been raised. In response to public concern, Congress directed the United States Environmental Protection Agency (EPA) to conduct research to examine the relationship between hydraulic fracturing and drinking water resources (USEPA, 2011a).

EPA will compile data on hydraulic fracturing water use and the hydrology of selected study areas, case studies or scenario evaluations. These data will include precipitation data, ground water levels, surface water flows, and water quality as well as data on hydraulic fracturing operations, such as the location of wells and the recorded water used/handled during fracturing activities. The EPA study approach will include specific case studies, both retrospective and prospective as well as data collection from selected regional study areas and scenario evaluations as described in the *Plan to Study the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources* (USEPA, 2011a).

Retrospective case studies are focused on investigating reported instances of drinking water resource contamination in areas where hydraulic fracturing events have already occurred. The goal is to determine whether or not the reported impacts are due to hydraulic fracturing activities. These studies will use existing data and may include environmental field sampling, modeling, and/or parallel laboratory investigations.

Prospective case studies involve sites where hydraulic fracturing will be implemented after the research is initiated. These cases allow sampling and characterization of the site prior to, during,

and after drilling, water extraction, injection of the fracturing fluid, flowback, and production. At each step in the process, data will be collected to characterize both the pre- and post-fracturing conditions at the site.

Scenario evaluations explore realistic, hypothetical scenarios across the hydraulic fracturing water lifecycle that may result in adverse impacts to drinking water resources based on current understanding and available data. The scenarios will include a reference case involving typical management and engineering practices in representative geologic settings. Typical management and engineering practices will be based on what EPA learns from case studies as well as the minimum requirements imposed by state regulatory agencies. Potential modes of failure, both in terms of engineering controls and geologic characteristics, will be introduced and modeled to represent various states of system vulnerability. The scenario evaluations will produce insights into site-specific and regional vulnerabilities.

Simple water balance analysis will be conducted using available data. The collected data will be compiled in conjunction with hydrological trends over the same period of time. Control areas that have similar baseline water demands and have no oil and gas development will be compared to areas with intense hydraulic fracturing activity to isolate and identify the impacts of hydraulic fracturing on water availability. Control areas will be within the study area if possible or in the same vicinity but will be remote from oil and gas development activities to be devoid of hydraulic fracturing impacts (see Section 2.9.1). A critical analysis of trends in water flows (including "environmental flows") and water usage patterns in areas impacted by hydraulic fracturing activities will be conducted to determine whether water withdrawals for hydraulic fracturing activities alter ground and surface water flows. Data collection will support the assessment of the impacts of hydraulic fracturing on water availability at various spatial scales (e.g., site, watershed, basin, and play) and temporal scales (e.g., days, months, and years).

The Ground Water and Ecosystems Restoration Division (GWERD) is the U.S. Environmental Protection Agency's center for risk management research on the subsurface environment and the interface of that environment with other environmental compartments such as surface water and the atmosphere. The GWERD is involved with the EPA Office of Research and Development (ORD) Hydraulic Fracturing research efforts and is leading activities for targeted data collection, federal partner data collection, study data collection and web access for the public information. The Shaw Work Assignment (WA) #WA-HF-2-10 from GWERD through the Onsite Analytical and Technical Support (OATS) contract provides for field, technical, and database support to the EPA for investigations of hydraulic fracturing impacts to drinking water resources. (USEPA 2011b). Shaw's support will be directed by EPA as projects or tasks in the form of documented Technical Directives (TD).

1.3 PROJECT/TASK DESCRIPTION

This Quality Assurance Project Plan (QAPP) will address data collection activities as described below.

The data collection activities will be focused on existing water supply and water quantity data, but may also include water quality data, if available, for hydraulic fracturing activities related to natural gas production. Data will be collected from sources, such as USGS, state Departments of Environmental Quality (DEQ) or Departments of Environmental Protection (DEP), private water management groups, Army Corps of Engineers, USDA-NRCS and state GIS Departments, etc., that have collected and/or published data. Shaw will begin identification and survey of existing sources of water supply, quantity, and quality data as well as hydraulic fracturing activity information for the following two selected scenario study area/regional locations:

- Susquehanna River Basin/Marcellus Shale in Pennsylvania
- Garfield County/Piceance Basin in Colorado.

These locations represent humid and arid areas of the country, for which sufficient data are available for study. Due to the large volumes of water required for fracturing operations significant impacts may vary from one part of the country to another and from one time of the year to another. Humid areas with greater precipitation and surface water volume will likely be less affected by large volume water use in fracturing activities than arid areas with generally less water availability. Including study areas that represent both arid and humid conditions will provide a contrast of impacts for the different conditions and the range of impacts is likely to be inclusive of most areas of the country.

Case study data collection may be from any number of retrospective and prospective case study nominations as discussed in Section 2.9.1, Scope of Data Collection. Case studies may be defined or amended by EPA based on results from data collection activities.

Initial project activities will involve three phases of data collection and organization as defined below:

1. Initial discovery, survey and inventory of applicable data from known, recommended or otherwise discovered sources,
2. Collection of data in electronic format - Data from applicable information sources will be evaluated against acceptance criteria (Section 2.9 Non-Direct Measurements) by the Data Collector and if deemed acceptable, the data will be saved in electronic data files in an organized manner. The Data Collector will also prepare an inventory of data sources, data type and amount of data collected.
3. Compilation of the data into usable formats for entry into Geographic Information System (GIS), database software and use in modeling programs - It may be possible to directly compile collected data into user software programs; however, data entry (or use by programs) may require preparation of compiled data files.

In addition to these specific data collection activities, other activities including GIS map preparation as well as database and modeling scoping and requirement specification may be performed. GIS activities will include preparation of national and regional map projections or overlays with addition of specific geographic locations, identifications and details as defined by

The data collection activities will be focused on existing water supply and water quantity data, but may also include water quality data, if available, for hydraulic fracturing activities related to natural gas production. Data will be collected from sources, such as USGS, state Departments of Environmental Quality (DEQ) or Departments of Environmental Protection (DEP), private water management groups, Army Corps of Engineers, USDA-NRCS and state GIS Departments, etc., that have collected and/or published data. Shaw will begin identification and survey of existing sources of water supply, quantity, and quality data as well as hydraulic fracturing activity information for the following two selected scenario study area/regional locations:

- Susquehanna River Basin/Marcellus Shale in Pennsylvania
- Garfield County/Piceance Basin in Colorado.

These locations represent humid and arid areas of the country, for which sufficient data are available for study. Due to the large volumes of water required for fracturing operations significant impacts may vary from one part of the country to another and from one time of the year to another. Humid areas with greater precipitation and surface water volume will likely be less affected by large volume water use in fracturing activities than arid areas with generally less water availability. Including study areas that represent both arid and humid conditions will provide a contrast of impacts for the different conditions and the range of impacts is likely to be inclusive of most areas of the country.

Case study data collection may be from any number of retrospective and prospective case study nominations as discussed in Section 2.9.1, Scope of Data Collection. Case studies may be defined or amended by EPA based on results from data collection activities.

Initial project activities will involve three phases of data collection and organization as defined below:

1. Initial discovery, survey and inventory of applicable data from known, recommended or otherwise discovered sources,
2. Collection of data in electronic format - Data from applicable information sources will be evaluated against acceptance criteria (Section 2.9 Non-Direct Measurements) by the Data Collector and if deemed acceptable, the data will be saved in electronic data files in an organized manner. The Data Collector will also prepare an inventory of data sources, data type and amount of data collected.
3. Compilation of the data into usable formats for entry into Geographic Information System (GIS), database software and use in modeling programs - It may be possible to directly compile collected data into user software programs; however, data entry (or use by programs) may require preparation of compiled data files.

In addition to these specific data collection activities, other activities including GIS map preparation as well as database and modeling scoping and requirement specification may be performed. GIS activities will include preparation of national and regional map projections or overlays with addition of specific geographic locations, identifications and details as defined by

EPA. Database and site modeling scoping will be ongoing based on development of the source information inventory as source survey and data collection activities progress.

1.3.1 Initial Discovery, Survey and Inventory of Source Data

The Data Collector will search for sources that have collected and/or published applicable data and inventory data that is available. EPA's suggested sources of initial data for the two scenario study locations are:

- Susquehanna River Basin (PA): Susquehanna River Commission, USGS, Pennsylvania Dept. of Environmental Protection, Pennsylvania Department of Natural Resources (PADNR), RAIN Water Quality Network, US Army Corps of Engineers
- Garfield County (CO): Colorado Oil and Gas Conservation Commission, (COGCC), Colorado Department of Natural Resources (CONDR), USGS.

Shaw has also been directed to additional data sources (references) that can be found in the Hydraulic Fracturing Study Plan, in particular Chapter 6 (USEPA, 2011a). Other potential sources of existing data are listed in EPA's guidance document for QAPPs, Chapter 3 (USEPA, 2002c). These include state and local monitoring programs, state and federal agencies which have quality management for databases of spatial, environmental and natural resources data, EPA's Environmental Information Management System (EIMS) and Environmental Data Registry, as well as published literature and research in trade journals and professional periodical publications.

During the initial discovery and survey of information sources phase an inventory of different types of data available from sources will be developed. The inventory will be a text file matrix, or a spreadsheet prepared and maintained by the Data Collector containing a list of information sources with an accounting of the specific types of data (see data types in Section 2.9.1) provided from each source. In addition the inventory will include the number of data records, range of data or other appropriate indication of the amount of data available from each source. The inventory will be used as a tool to communicate results of data collection activities, assess data needs or gaps, and determine GIS and database requirements.

Similar approaches will be used for retrospective and prospective case study data, but specific direction will be provided in technical directives from the EPA.

1.3.2 Data Collection

Data will be assessed against acceptance criteria (Section 2.9 Non-Direct Measurements) and collected as electronic data files by the Data Collector onto a computer. The data files will be segregated from non-project files and organized by data type, source or source type and by format. Some file manipulation may be necessary to consolidate data, files or file types to simplify file organization and data handling. The activities performed by the Data Collector will be reviewed by the Data Reviewer for concurrence and accuracy.

1.3.3 Data Compilation

Based on the inventory from data survey and collection activities requirements for GIS maps, information database and scenario modeling will be formulated to aid in software selection and development. In addition, record fields, data elements and standard units for data will be defined and overall structure of information organization can be formulated for use of the information. Once a usable form of information format is defined and approved, then the collected data can be compiled. Ideally the data will be compiled directly from source data files into a database or user software, but it may be necessary to manipulate source data and prepare compiled data files in specified format for uploading into or use by selected software. Compiled data files are distinguished from the information database that will be developed for end user use. The Data Collector will inform EPA of the processes that will be used to prepare these compiled data files and will prepare the data files with assistance from the Database Manager. These activities will be reviewed by the Data Reviewer for accuracy.

1.4 QUALITY OBJECTIVES AND CRITERIA

Research activities associated with this study will be conducted in accordance with EPA's Quality System for environmental data and technology (USEPA, 2002a) and Shaw's on-site EPA contract Quality Management Plan (QMP) (Shaw, 2009).

EPA's policy is based on the national consensus standard ANSI/ASQ E4-1994, *Specifications and Guidelines for Environmental Data Collection and Environmental Technology programs (USEPA, 2002)*. This policy recommends applying a graded approach to quality systems according to the specific objectives and needs of the organization and the intended use of the information being collected. Because the information collected for this project will have significant national interest and importance, this project requires the most detailed and rigorous QA and QC for legal and scientific defensibility (EPA Category I).

Current directed project activities involve locating, identifying, surveying and collecting existing data for the identified sites. Other directed activities include preparation of GIS maps and database development with the collected data. Quality objectives and requirements for these activities are described in Section 2.9, Non-Direct Measurements.

Case study activities that may involve collection of direct measurement data as well as modeling efforts for scenario evaluations will be addressed in task-specific QAPPs or QAPP revisions.

All project deliverables will meet EPA's standards of transparency, objectivity, integrity, and utility (USEPA, 2002b). This will be done by providing sources of data, limitations on the data, assumptions, and manipulations or calculations performed so that the work can be reproduced by qualified third parties.

1.5 SPECIAL TRAINING AND CERTIFICATION

The Data Collector shall have an advanced degree in Geology, Environmental Science or

EPA. Database and site modeling scoping will be ongoing based on development of the source information inventory as source survey and data collection activities progress.

1.3.1 Initial Discovery, Survey and Inventory of Source Data

The Data Collector will search for sources that have collected and/or published applicable data and inventory data that is available. EPA's suggested sources of initial data for the two scenario study locations are:

- Susquehanna River Basin (PA): Susquehanna River Commission, USGS, Pennsylvania Dept. of Environmental Protection, Pennsylvania Department of Natural Resources (PADNR), RAIN Water Quality Network, US Army Corps of Engineers
- Garfield County (CO): Colorado Oil and Gas Conservation Commission, (COGCC), Colorado Department of Natural Resources (CONDR), USGS.

Shaw has also been directed to additional data sources (references) that can be found in the Hydraulic Fracturing Study Plan, in particular Chapter 6 (USEPA, 2011a). Other potential sources of existing data are listed in EPA's guidance document for QAPPs, Chapter 3 (USEPA, 2002c). These include state and local monitoring programs, state and federal agencies which have quality management for databases of spatial, environmental and natural resources data, EPA's Environmental Information Management System (EIMS) and Environmental Data Registry, as well as published literature and research in trade journals and professional periodical publications.

During the initial discovery and survey of information sources phase an inventory of different types of data available from sources will be developed. The inventory will be a text file matrix, or a spreadsheet prepared and maintained by the Data Collector containing a list of information sources with an accounting of the specific types of data (see data types in Section 2.9.1) provided from each source. In addition the inventory will include the number of data records, range of data or other appropriate indication of the amount of data available from each source. The inventory will be used as a tool to communicate results of data collection activities, assess data needs or gaps, and determine GIS and database requirements.

Similar approaches will be used for retrospective and prospective case study data, but specific direction will be provided in technical directives from the EPA.

1.3.2 Data Collection

Data will be assessed against acceptance criteria (Section 2.9 Non-Direct Measurements) and collected as electronic data files by the Data Collector onto a computer. The data files will be segregated from non-project files and organized by data type, source or source type and by format. Some file manipulation may be necessary to consolidate data, files or file types to simplify file organization and data handling. The activities performed by the Data Collector will be reviewed by the Data Reviewer for concurrence and accuracy.

The data collection activities will be focused on existing water supply and water quantity data, but may also include water quality data, if available, for hydraulic fracturing activities related to natural gas production. Data will be collected from sources, such as USGS, state Departments of Environmental Quality (DEQ) or Departments of Environmental Protection (DEP), private water management groups, Army Corps of Engineers, USDA-NRCS and state GIS Departments, etc., that have collected and/or published data. Shaw will begin identification and survey of existing sources of water supply, quantity, and quality data as well as hydraulic fracturing activity information for the following two selected scenario study area/regional locations:

- Susquehanna River Basin/Marcellus Shale in Pennsylvania
- Garfield County/Piceance Basin in Colorado.

These locations represent humid and arid areas of the country, for which sufficient data are available for study. Due to the large volumes of water required for fracturing operations significant impacts may vary from one part of the country to another and from one time of the year to another. Humid areas with greater precipitation and surface water volume will likely be less affected by large volume water use in fracturing activities than arid areas with generally less water availability. Including study areas that represent both arid and humid conditions will provide a contrast of impacts for the different conditions and the range of impacts is likely to be inclusive of most areas of the country.

Case study data collection may be from any number of retrospective and prospective case study nominations as discussed in Section 2.9.1, Scope of Data Collection. Case studies may be defined or amended by EPA based on results from data collection activities.

Initial project activities will involve three phases of data collection and organization as defined below:

1. Initial discovery, survey and inventory of applicable data from known, recommended or otherwise discovered sources,
2. Collection of data in electronic format - Data from applicable information sources will be evaluated against acceptance criteria (Section 2.9 Non-Direct Measurements) by the Data Collector and if deemed acceptable, the data will be saved in electronic data files in an organized manner. The Data Collector will also prepare an inventory of data sources, data type and amount of data collected.
3. Compilation of the data into usable formats for entry into Geographic Information System (GIS), database software and use in modeling programs - It may be possible to directly compile collected data into user software programs; however, data entry (or use by programs) may require preparation of compiled data files.

In addition to these specific data collection activities, other activities including GIS map preparation as well as database and modeling scoping and requirement specification may be performed. GIS activities will include preparation of national and regional map projections or overlays with addition of specific geographic locations, identifications and details as defined by

The data collection activities will be focused on existing water supply and water quantity data, but may also include water quality data, if available, for hydraulic fracturing activities related to natural gas production. Data will be collected from sources, such as USGS, state Departments of Environmental Quality (DEQ) or Departments of Environmental Protection (DEP), private water management groups, Army Corps of Engineers, USDA-NRCS and state GIS Departments, etc., that have collected and/or published data. Shaw will begin identification and survey of existing sources of water supply, quantity, and quality data as well as hydraulic fracturing activity information for the following two selected scenario study area/regional locations:

- Susquehanna River Basin/Marcellus Shale in Pennsylvania
- Garfield County/Piceance Basin in Colorado.

These locations represent humid and arid areas of the country, for which sufficient data are available for study. Due to the large volumes of water required for fracturing operations significant impacts may vary from one part of the country to another and from one time of the year to another. Humid areas with greater precipitation and surface water volume will likely be less affected by large volume water use in fracturing activities than arid areas with generally less water availability. Including study areas that represent both arid and humid conditions will provide a contrast of impacts for the different conditions and the range of impacts is likely to be inclusive of most areas of the country.

Case study data collection may be from any number of retrospective and prospective case study nominations as discussed in Section 2.9.1, Scope of Data Collection. Case studies may be defined or amended by EPA based on results from data collection activities.

Initial project activities will involve three phases of data collection and organization as defined below:

1. Initial discovery, survey and inventory of applicable data from known, recommended or otherwise discovered sources,
2. Collection of data in electronic format - Data from applicable information sources will be evaluated against acceptance criteria (Section 2.9 Non-Direct Measurements) by the Data Collector and if deemed acceptable, the data will be saved in electronic data files in an organized manner. The Data Collector will also prepare an inventory of data sources, data type and amount of data collected.
3. Compilation of the data into usable formats for entry into Geographic Information System (GIS), database software and use in modeling programs - It may be possible to directly compile collected data into user software programs; however, data entry (or use by programs) may require preparation of compiled data files.

In addition to these specific data collection activities, other activities including GIS map preparation as well as database and modeling scoping and requirement specification may be performed. GIS activities will include preparation of national and regional map projections or overlays with addition of specific geographic locations, identifications and details as defined by

EPA. Database and site modeling scoping will be ongoing based on development of the source information inventory as source survey and data collection activities progress.

1.3.1 Initial Discovery, Survey and Inventory of Source Data

The Data Collector will search for sources that have collected and/or published applicable data and inventory data that is available. EPA's suggested sources of initial data for the two scenario study locations are:

- Susquehanna River Basin (PA): Susquehanna River Commission, USGS, Pennsylvania Dept. of Environmental Protection, Pennsylvania Department of Natural Resources (PADNR), RAIN Water Quality Network, US Army Corps of Engineers
- Garfield County (CO): Colorado Oil and Gas Conservation Commission, (COGCC), Colorado Department of Natural Resources (CONDR), USGS.

Shaw has also been directed to additional data sources (references) that can be found in the Hydraulic Fracturing Study Plan, in particular Chapter 6 (USEPA, 2011a). Other potential sources of existing data are listed in EPA's guidance document for QAPPs, Chapter 3 (USEPA, 2002c). These include state and local monitoring programs, state and federal agencies which have quality management for databases of spatial, environmental and natural resources data, EPA's Environmental Information Management System (EIMS) and Environmental Data Registry, as well as published literature and research in trade journals and professional periodical publications.

During the initial discovery and survey of information sources phase an inventory of different types of data available from sources will be developed. The inventory will be a text file matrix, or a spreadsheet prepared and maintained by the Data Collector containing a list of information sources with an accounting of the specific types of data (see data types in Section 2.9.1) provided from each source. In addition the inventory will include the number of data records, range of data or other appropriate indication of the amount of data available from each source. The inventory will be used as a tool to communicate results of data collection activities, assess data needs or gaps, and determine GIS and database requirements.

Similar approaches will be used for retrospective and prospective case study data, but specific direction will be provided in technical directives from the EPA.

1.3.2 Data Collection

Data will be assessed against acceptance criteria (Section 2.9 Non-Direct Measurements) and collected as electronic data files by the Data Collector onto a computer. The data files will be segregated from non-project files and organized by data type, source or source type and by format. Some file manipulation may be necessary to consolidate data, files or file types to simplify file organization and data handling. The activities performed by the Data Collector will be reviewed by the Data Reviewer for concurrence and accuracy.

1.3.3 Data Compilation

Based on the inventory from data survey and collection activities requirements for GIS maps, information database and scenario modeling will be formulated to aid in software selection and development. In addition, record fields, data elements and standard units for data will be defined and overall structure of information organization can be formulated for use of the information. Once a usable form of information format is defined and approved, then the collected data can be compiled. Ideally the data will be compiled directly from source data files into a database or user software, but it may be necessary to manipulate source data and prepare compiled data files in specified format for uploading into or use by selected software. Compiled data files are distinguished from the information database that will be developed for end user use. The Data Collector will inform EPA of the processes that will be used to prepare these compiled data files and will prepare the data files with assistance from the Database Manager. These activities will be reviewed by the Data Reviewer for accuracy.

1.4 QUALITY OBJECTIVES AND CRITERIA

Research activities associated with this study will be conducted in accordance with EPA's Quality System for environmental data and technology (USEPA, 2002a) and Shaw's on-site EPA contract Quality Management Plan (QMP) (Shaw, 2009).

EPA's policy is based on the national consensus standard ANSI/ASQ E4-1994, *Specifications and Guidelines for Environmental Data Collection and Environmental Technology programs (USEPA, 2002)*. This policy recommends applying a graded approach to quality systems according to the specific objectives and needs of the organization and the intended use of the information being collected. Because the information collected for this project will have significant national interest and importance, this project requires the most detailed and rigorous QA and QC for legal and scientific defensibility (EPA Category I).

Current directed project activities involve locating, identifying, surveying and collecting existing data for the identified sites. Other directed activities include preparation of GIS maps and database development with the collected data. Quality objectives and requirements for these activities are described in Section 2.9, Non-Direct Measurements.

Case study activities that may involve collection of direct measurement data as well as modeling efforts for scenario evaluations will be addressed in task-specific QAPPs or QAPP revisions.

All project deliverables will meet EPA's standards of transparency, objectivity, integrity, and utility (USEPA, 2002b). This will be done by providing sources of data, limitations on the data, assumptions, and manipulations or calculations performed so that the work can be reproduced by qualified third parties.

1.5 SPECIAL TRAINING AND CERTIFICATION

The Data Collector shall have an advanced degree in Geology, Environmental Science or

EPA. Database and site modeling scoping will be ongoing based on development of the source information inventory as source survey and data collection activities progress.

1.3.1 Initial Discovery, Survey and Inventory of Source Data

The Data Collector will search for sources that have collected and/or published applicable data and inventory data that is available. EPA's suggested sources of initial data for the two scenario study locations are:

- Susquehanna River Basin (PA): Susquehanna River Commission, USGS, Pennsylvania Dept. of Environmental Protection, Pennsylvania Department of Natural Resources (PADNR), RAIN Water Quality Network, US Army Corps of Engineers
- Garfield County (CO): Colorado Oil and Gas Conservation Commission, (COGCC), Colorado Department of Natural Resources (CONDR), USGS.

Shaw has also been directed to additional data sources (references) that can be found in the Hydraulic Fracturing Study Plan, in particular Chapter 6 (USEPA, 2011a). Other potential sources of existing data are listed in EPA's guidance document for QAPPs, Chapter 3 (USEPA, 2002c). These include state and local monitoring programs, state and federal agencies which have quality management for databases of spatial, environmental and natural resources data, EPA's Environmental Information Management System (EIMS) and Environmental Data Registry, as well as published literature and research in trade journals and professional periodical publications.

During the initial discovery and survey of information sources phase an inventory of different types of data available from sources will be developed. The inventory will be a text file matrix, or a spreadsheet prepared and maintained by the Data Collector containing a list of information sources with an accounting of the specific types of data (see data types in Section 2.9.1) provided from each source. In addition the inventory will include the number of data records, range of data or other appropriate indication of the amount of data available from each source. The inventory will be used as a tool to communicate results of data collection activities, assess data needs or gaps, and determine GIS and database requirements.

Similar approaches will be used for retrospective and prospective case study data, but specific direction will be provided in technical directives from the EPA.

1.3.2 Data Collection

Data will be assessed against acceptance criteria (Section 2.9 Non-Direct Measurements) and collected as electronic data files by the Data Collector onto a computer. The data files will be segregated from non-project files and organized by data type, source or source type and by format. Some file manipulation may be necessary to consolidate data, files or file types to simplify file organization and data handling. The activities performed by the Data Collector will be reviewed by the Data Reviewer for concurrence and accuracy.

1.3.3 Data Compilation

Based on the inventory from data survey and collection activities requirements for GIS maps, information database and scenario modeling will be formulated to aid in software selection and development. In addition, record fields, data elements and standard units for data will be defined and overall structure of information organization can be formulated for use of the information. Once a usable form of information format is defined and approved, then the collected data can be compiled. Ideally the data will be compiled directly from source data files into a database or user software, but it may be necessary to manipulate source data and prepare compiled data files in specified format for uploading into or use by selected software. Compiled data files are distinguished from the information database that will be developed for end user use. The Data Collector will inform EPA of the processes that will be used to prepare these compiled data files and will prepare the data files with assistance from the Database Manager. These activities will be reviewed by the Data Reviewer for accuracy.

1.4 QUALITY OBJECTIVES AND CRITERIA

Research activities associated with this study will be conducted in accordance with EPA's Quality System for environmental data and technology (USEPA, 2002a) and Shaw's on-site EPA contract Quality Management Plan (QMP) (Shaw, 2009).

EPA's policy is based on the national consensus standard ANSI/ASQ E4-1994, *Specifications and Guidelines for Environmental Data Collection and Environmental Technology programs (USEPA, 2002)*. This policy recommends applying a graded approach to quality systems according to the specific objectives and needs of the organization and the intended use of the information being collected. Because the information collected for this project will have significant national interest and importance, this project requires the most detailed and rigorous QA and QC for legal and scientific defensibility (EPA Category I).

Current directed project activities involve locating, identifying, surveying and collecting existing data for the identified sites. Other directed activities include preparation of GIS maps and database development with the collected data. Quality objectives and requirements for these activities are described in Section 2.9, Non-Direct Measurements.

Case study activities that may involve collection of direct measurement data as well as modeling efforts for scenario evaluations will be addressed in task-specific QAPPs or QAPP revisions.

All project deliverables will meet EPA's standards of transparency, objectivity, integrity, and utility (USEPA, 2002b). This will be done by providing sources of data, limitations on the data, assumptions, and manipulations or calculations performed so that the work can be reproduced by qualified third parties.

1.5 SPECIAL TRAINING AND CERTIFICATION

The Data Collector shall have an advanced degree in Geology, Environmental Science or

Engineering, and have sufficient experience to understand and evaluate the information collected. This person will also be trained in the use of GIS maps/software and will have sufficient experience in the use of GIS computer software to prepare GIS maps to meet project objectives.

The Data Reviewer must have similar experience as the Data Collector, i.e., an advanced degree in Geology, Environmental Science or Engineering, and have sufficient experience to understand and evaluate the information collected. This person will also be trained in the use of GIS maps/software and will have sufficient experience in the use of GIS information and GIS computer software to evaluate prepared GIS maps against project objectives.

The Database Manager should have an advanced degree in computer science and have sufficient experience to define database requirements to assist in software selection and develop a database that is consistent with project objectives.

1.6 DOCUMENTS AND RECORDS

Documents and records will be managed in accordance with Shaw's EPA contract (OATS) Quality Management Plan (Shaw, 2009).

Typically, the maintenance and eventual turnover of contractual and technical records will be as specified by the EPA. Records prepared and maintained by Shaw which are pertinent only to Shaw will be maintained as specified by Shaw records management procedures.

A filing system is established and implemented for Work Assignments (Was) undertaken within the OATS program. Applicable items included in these files may include:

1) WAs
2) Work Plans
3) Organizational conflict of interest/conflict of business checks
4) Work Plan approvals, WA amendments, and Technical Directives
5) Project-related correspondence
6) WA closeout documents
7) Quality records and documents
8) Procurement Documentation

Hardcopy documents and electronic files generated off-site on behalf of OATS for this WA will be forwarded to the OATS (Shaw, Ada, OK) File Administrator for incorporation into the project file. Ongoing working electronic files being generated off-site as an eventual project product, such as compiled data files, spreadsheets or database, will be forwarded to the EPA or saved to an electronic portal site available to the EPA, or at a minimum, forwarded to the OATS File Administrator for backup within the EPA computer network system on a weekly basis, at a minimum, while being generated, edited or modified.

Hardcopies of source data that will be manually entered into data compilations or database shall be maintained to facilitate checking and documenting checks of entries. This documentation must be retained in the OATS project file.

Shaw will control the review, revision, and distribution of the most recent version of the QAPP. Document control format (Sec. No., Rev. No., and Date) shall appear in the upper right-hand corner of each page of the QAPP. A signed approval form will accompany the QAPP. Any revision to the QAPP shall be circulated to Shaw and EPA project staff for review and approval. Documentation of approval is evidenced by signatures. Final approved version of the QAPP will be distributed by Shaw to all project staff.

Project staff conducting data collection, database development and implementation, and/or GIS activities shall document their work in notebooks or other means approved by the EPA QA Manager per requirements in ORD PPM 13.2, *Paper Laboratory Records* (USEPA ORD, 2006).

As this is a QA Category 1 project, permanent retention of project records is required per Agency Records Schedules 501. All planning documents (QAPPs), data, databases, GIS files, maps, project deliverables, notebooks, correspondence, etc., generated during the course of this project shall be transferred to Susan Mravik, the EPA WAM. They shall be stored in her office at RSKERC until they are transferred to RSKERC's Records Storage Room. At an as yet to be determined time in the future the records will be transferred to a National Archive facility.

Management of project data is described in Sec. 2.10, Data Management.

2.0 DATA GENERATION AND ACQUISITION

2.1 SAMPLING PROCESS DESIGN (EXPERIMENTAL DESIGN)

There are no physical sampling activities identified as part of current project activities.

2.2 SAMPLING METHODS

There are no physical sampling activities identified as part of current project activities.

2.3 SAMPLE HANDLING AND CUSTODY

There are no physical sampling activities identified as part of current project activities.

2.4 ANALYTICAL METHODS

There are no sampling or analysis activities identified as part of current project activities.

2.5 QUALITY CONTROL

There are no sampling or analysis activities identified as part of current project activities, so no

Hardcopies of source data that will be manually entered into data compilations or database shall be maintained to facilitate checking and documenting checks of entries. This documentation must be retained in the OATS project file.

Shaw will control the review, revision, and distribution of the most recent version of the QAPP. Document control format (Sec. No., Rev. No., and Date) shall appear in the upper right-hand corner of each page of the QAPP. A signed approval form will accompany the QAPP. Any revision to the QAPP shall be circulated to Shaw and EPA project staff for review and approval. Documentation of approval is evidenced by signatures. Final approved version of the QAPP will be distributed by Shaw to all project staff.

Project staff conducting data collection, database development and implementation, and/or GIS activities shall document their work in notebooks or other means approved by the EPA QA Manager per requirements in ORD PPM 13.2, *Paper Laboratory Records* (USEPA ORD, 2006).

As this is a QA Category 1 project, permanent retention of project records is required per Agency Records Schedules 501. All planning documents (QAPPs), data, databases, GIS files, maps, project deliverables, notebooks, correspondence, etc., generated during the course of this project shall be transferred to Susan Mravik, the EPA WAM. They shall be stored in her office at RSKERC until they are transferred to RSKERC's Records Storage Room. At an as yet to be determined time in the future the records will be transferred to a National Archive facility.

Management of project data is described in Sec. 2.10, Data Management.

2.0 DATA GENERATION AND ACQUISITION

2.1 SAMPLING PROCESS DESIGN (EXPERIMENTAL DESIGN)

There are no physical sampling activities identified as part of current project activities.

2.2 SAMPLING METHODS

There are no physical sampling activities identified as part of current project activities.

2.3 SAMPLE HANDLING AND CUSTODY

There are no physical sampling activities identified as part of current project activities.

2.4 ANALYTICAL METHODS

There are no sampling or analysis activities identified as part of current project activities.

2.5 QUALITY CONTROL

There are no sampling or analysis activities identified as part of current project activities, so no

measurement QC activities are currently planned.

2.6 INSTRUMENT/EQUIPMENT TESTING INSPECTION AND MAINTENANCE

Use of instruments or equipment is not planned as part of current project activities.

2.7 INSTRUMENT/EQUIPMENT CALIBRATION AND FREQUENCY

Use of instruments or equipment is not planned as part of current project activities.

2.8 INSPECTION/ACCEPTANCE OF SUPPLIES AND CONSUMABLES

Inspection/acceptance of supplies and consumables on the project is currently anticipated to be minimal and essentially administrative or information handling related. Inspection and acceptance will be indicated and documented by noting approval along with approver's initials and dates on the packing list, receipt, or invoice.

2.9 NON-DIRECT MEASUREMENTS

Non-direct measurements are data collected from existing sources, not directly measured or generated in this project. These data are also referred to as secondary data. Secondary data will be prepared for inclusion in the database following the procedure outlined in OATS Standard Operating Procedure (SOP), SOP-03, *Hydraulic Fracturing Data Handling and Database Management* (Shaw, 2011), which is attached.

2.9.1 Scope of Data Collection

The initial activity for this project is identifying sources and surveying the type of non-direct measurement information available from hydraulic fracturing case studies and other data concerning potentially impacted surface water and groundwater. Shaw will begin with the two currently selected regional study area locations:

- Susquehanna River Basin/Marcellus Shale in Pennsylvania
- Garfield County/Piceance Basin in Colorado.

The scope of this study is further defined below for each location:

- Data from Susquehanna River Basin (PA) area will be from within the entire basin down to the Chesapeake Bay and will be from the year 2005 or later. Data from natural gas/oil well hydraulic fracturing activities will be from wells that have been put into operation in 2005 or later.
- Data from Garfield County (CO) area will be from within the county boundaries, or from locations within watersheds that impinge on the Garfield County booundary, especially in the vicinity of the towns of Silt, Rifle and Battlement Mesa, and will be from the year 2000 or later. Data from natural gas/oil well hydraulic fracturing activities will be from wells that have been put into operation in 2000 or later.

The specific regional study data targeted for collection will be of two basic types: 1) data specific to hydraulic fracturing activities at wells; and 2) data concerning natural or engineered surface water or groundwater in the vicinity of hydraulic fracturing activities, and in the appropriate time frame to register impacts and trends from hydraulic fracturing activities:

Hydraulic fracturing well data will include specific well identification and location information. Data will also include the date of record or activity as well as available specifics such as the following:

- elevation at location
- quantity of water used
- site geologic cross-sections
- water source
- water quality (such as pH, temperature, conductivity)
- water storage, i.e., natural (stream), tanks, pit, engineered impoundment
- well type, i.e., vertical, horizontal or directional
- depth fractured
- fracture directionality
- fracturing pressures
- geologic zone fractured
- site or geologic zone hydrologic information, e.g., hydraulic gradient and conductivity
- quantity of recovered water
- recovered water quality
- fate of recovered water, i.e., treated and discharged (receptor), discharged (receptor), injected (depth), etc

Groundwater and surface water data will include measurement identification information with date and location of measurements as well as available data such as the following:

- water type, i.e., groundwater, stream, lake
- location ID, i.e., well or reading/gauge station ID, etc.
- precipitation data
- location elevation
- surface water level- elevation/depth for lake, reservoir or impoundment
- groundwater level- elevation/depth
- stream stage
- stream flow or discharge rate
- water well – fractured or not fractured
- water well discharge rate
- water quality (such as pH, temperature, conductivity)

The main objective of collecting ground and surface water data is to capture behavioral and temporal trends related to hydraulic fracturing activities.

Additional data targets and additions or changes to the data collection scope may be made by the EPA based on initial survey of available information from sources. These modifications must be documented.

Data will be identified and collected from acceptable sources during information source survey activities. Sources that provide metadata or information that describes the data and their quality criteria will be assessed for suitability and the metadata will be captured or documented. The inventory of different types of data available that is developed during survey activities will be used to define database requirements as well as plan database development and GIS mapping activities.

Case study data collection may be from any number of retrospective and prospective case study nominations as listed below. These studies will typically be more focused and will involve a smaller geospatial scale than the regional studies. Case studies may be defined or amended by EPA based on results from data collection activities and specific direction will be provided on a site specific basis in technical directives.

Nominated Retrospective Case Studies:

- Washington County Pennsylvania (Marcellus Shale)
- Bradford/Susquehanna counties Pennsylvania (Marcellus Shale)
- Wise County Texas (Barnett Shale)
- Kildeer North Dakota (Bakken Shale)
- Las Animas and Huerfano Counties Colorado (Raton Basin)

Nominated Prospective Case Studies:

- Washington County Pennsylvania (Marcellus Shale)
- DeSoto Parish Louisiana (Haynesville Shale)

2.9.2 Geospatial (Locational, Elevational and Temporal) Information Requirements

Locational data should adhere to EPA National Geospatial Data Policy (USEPA, 2005), and any deviations shall be identified. Specific requirements include the following:

- Geo-Referenced Point Data – EPA policy requires geo-referenced coordinates be collected or derived, and appropriately documented in accordance to the adopted EPA/EDSC Latitude / Longitude Data Standard (USEPA, 2006). Locational data will span national to regional to site-specific scales and geospatial references and coordinate systems must be defined for each scale.
- Geospatial Data Accuracy (Locational) – In the absence of program-specific procedures addressing required minimum accuracy for geospatial data, EPA policy requires a minimum accuracy of Tier 5 (USEPA, 2005), which is described in Table 1. However, locational data will span national to regional to site-specific scales and acceptable geospatial accuracy must be defined for each scale or on a site specific basis. Data that

does not meet the established minimum accuracy requirements or data for which the accuracy is not defined and is otherwise deemed acceptable must be qualified and the qualification must be a part of the data set.

Table 1. Geospatial Accuracy Tiers

Tier Level	Accuracy and Precision	Examples of Horizontal Collection Method	Example Program Application
Tier 1	<1 m	Classical Surveying Techniques; plus GPS Carrier Phase Static Relative Position	Definition of contamination boundaries of site
Tier 2	1 – 5 m	GPS Carrier Phase Kinematic Relative Position	Definition of contamination boundaries of site
Tier 3	6 -25 m	GPS Code (Pseudo Range) Standard Position	Stack location; drinking water intake location
Tier 4	26 – 100 m	GPS Unspecified; Photo/GIS Interpolation	Site centroid; large area facility boundary
Tier 5	101 - 200 m	Urban Style Address Matching	Preliminary Site Location
Tier 6	201 - 999 m	Public Land Survey – Sixteenth Section	Prediction of Local Air Dispersion
Tier 7	1000 - 2000 m	Address Matching – Block Face	Batch Geo-coding
Tier 8	2001 - 5000 m	Census Block Centroid	State-level Population Statistics
Tier 9	> 5000 m	Zip Code Centroid	Generalized National Mapping
Tier 10	Unknown	N/A	Relative contextual data

- Geospatial Metadata – EPA policy requires metadata describing geospatial data in accordance with FGDC Content Standard for Digital Geospatial Metadata (FGDC, 1998). This includes temporal and elevation data. Data that does not meet the standard but for which documented information exists providing equivalent metadata information may be used, but this information must be documented and shall become a part of the data set in conformance to the standard. Methods used to construct metadata information must be documented.
- Elevation Data Accuracy – Elevation data for a location is needed at a minimum to recognize impacts on hydrologic conditions in relation to other locations. This data is obtained from standard Global Positioning System (GPS) readings. Minimum accuracy requirements must be established in a manner similar to that for locational (latitude and longitude) information and EPA's Tier levels may be applicable to elevation data as well.
- Map Projections and Scale - The projection associated with maps will be matched to scale and purpose. Regional and national scale map projections will be seamless across the conterminous USA and preserve area. Local scale map projections will be Cartesian

and thus preserve angle and distance.

- For data processing purposes, the Albers 2396 projection will be used for regional and national scale maps, with the following attributes:

  USGS US PROJECTION DATA
  Albers Conic Equal Area
  GRS 1980
  NAD 83

  CENTER PARALLEL 23.0°
  CENTER MERIDIAN -96.0°
  1ST PARALLEL 29.5°
  2ND PARALLEL 45.5°
  FALSE N 0
  FALSE E 0
  UNITS METERS

- For display and visualization (e.g. Google Earth), the web Mercator projection (WGW 1984 Web Mercator) is desirable.
- The UTM projection will be used for local mapping. In the event that the study involves more than one UTM zone, the State Plane coordinate system shall be considered.
- Custom maps may be necessary, and when used, a metadata file will be supplied describing in detail the associated projection.

2.9.3 Acceptance and QA Requirements for Water Quality Data

Water quality data may include a number of standard physical and analytical parameters, such as temperature, pH and conductivity. This data may be obtained from both hydraulic fracturing well information as well as groundwater well and surface water information. Acceptance and QA requirements for this data are described below and summarized in Table 2.

- Water quality data must have been acquired by methods approved by the federal government such as EPA, USGS, DOE or DoD methods, or universally-accepted methods such as ASTM or Standard Methods for Water and Wastewater (SMWW). These methods have defined data quality objectives, i.e., accuracy, precision, detection limits and quantitation limits.
- Data collected by other methods may be used, with EPA approval, but data quality objectives required by the methods must be documented and results must meet the data quality objectives required for the methods used to acquire the data. In addition the data must be qualified as "alternate method." If the methods used and data quality objectives of the methods are notably different than those from methods either approved by the federal government or universally-accepted, than the data may be used with EPA approval, but must be qualified by specifying the differences.
- In some cases it may be difficult to obtain sufficient information about methods used or their data quality objectives to evaluate data with respect to acceptance criteria. In these

cases the data may be used with EPA approval, but the data must be appropriately qualified to indicate the unknown quality information.

- Data for which the methods used to obtain the data and the data quality objectives or data quality indicator results are not defined or otherwise indicated may be used with EPA approval, but results must be qualified as "quality unknown."
- Any data that has been identified as not meeting the quality objectives of the methods used to collect the data will be rejected.

Table 2. Water Quality Data - Data Quality Objectives and Data Qualifications

Measurement	Method	Data Quality Objectives (DQO)	Acceptance	Data Qualifier
Water quality parameters, e.g., pH, temperature, conductivity	EPA, USGS, ASTM, SMWW, federal government approved or universally accepted	Method requirements	Meets DQO	None
	Other methods that are consistent with government approved or universally accepted methods	DQO consistent with government approved or universally accepted methods	Meets DQO and EPA approval	Alternate Method
	Methods notably different from government approved or universally accepted methods	Method requirements	Meets method requirements and EPA approval	Difference(s) defined
	Method or DQO Unavailable or undefined	Method or DQO Unavailable or undefined	EPA Approval	Missing Information Identified
	Unavailable or undefined	Unavailable or undefined and no indication that data did not meet method requirements	EPA Approval	Quality Unknown

2.9.4 Acceptance and QA Requirements for Other Measurement Data

Other measurement data include hydraulic fracturing operation data, groundwater level measurements, precipitation data, stream stage measurements, stream and water well discharge rates, etc. Methods for acquiring or generating this data may include those approved by the federal government such as EPA, and USGS, universally-accepted methods such as ASTM or may be standard operating procedures (SOPs) used by the source entity or even instrument/device manufacturers operating instructions (IMOI). The level to which the methods used and data quality objectives, i.e., accuracy, precision, etc., are defined or even available for this data may be variable. Acceptance and QA requirements for this data are described below and summarized in Table 3.

- Data must have been acquired by documented and approved methods with defined data quality objectives, i.e., accuracy and precision, etc., that are consistent with methods either approved by the federal government or universally-accepted, if such methods exist.
- If the methods used and data quality objectives of the methods are notably different than those from methods either approved by the federal government or universally-accepted, if such methods exist, then the data may be used with EPA approval, but must be qualified by specifying the differences.
- If sufficient information about the methods used or the quality objectives of the methods are not available to evaluate data with respect to the acceptance criteria, the data may be used with EPA approval, but the data must be appropriately qualified to indicate the unknown information.
- Data for which the methods used to obtain the data and the data quality objectives or data quality indicator results are not defined or otherwise indicated may be used with EPA approval, but results must be qualified as "quality unknown."
- Any data that has been identified as not meeting the quality objectives of the methods used to collect the data will be rejected.

Table 3. Other Measurements - Data Quality Objectives and Data Qualifications

Measurement	Method	Data Quality Objectives (DQO)	Acceptance	Data Qualifier
HF operating parameters, quantities	ASTM, USGS, SOP, IMOI	Method requirements or Procedure requirements with DQO consistent with government approved or universally accepted method(s)	Meets DQO	None
Groundwater level	EPA, SOP, IMOI			
Stream discharge rate	USGS, SOP, IMOI			
Water well discharge rate	ASTM, USGS, SOP			
Stream stage	USGS, SOP, IMOI			
HF operating parameters, quantities, Groundwater level, Stream discharge rate, Water well discharge rate, Stream stage	Methods notably different from government approved or universally accepted methods, if any exist	Method requirements	Meets method requirements and EPA approval	Difference(s) defined
	Method or DQO unavailable or undefined	Method or DQO unavailable or undefined	EPA approval	Missing information identified
	Method unavailable or undefined	DQO unavailable or undefined and no indication that data did not meet method requirements	EPA Approval	Quality Unknown

2.9.5 Other Acceptance Requirements for Data

Non-Direct measurement data requirements are described below:

- Methods used to discover and collect non-direct measurement data must be documented and generally accepted from both a technical and QA standpoint. The quality objective is to assure that discovery, identification and selection efforts are thorough, comprehensive and unbiased. Current methods are to investigate known sources that have been recommended by EPA and documented in this QAPP as well as secondary sources derived from these sources and evaluate available data for acceptance as outlined in Section, 2.9. Other methods may be searching for information in specific professional society journals, which shall be documented as a note identifying the journal, how it was selected and the time frame covered by the journal dates searched. This documentation shall be retained in the OATS project file.
- Data shall meet the spatial and temporal requirements as described in Section 2.9.2.
- In general, data must be from a known or recognized and accepted source, such as; a state or federal government agency (e.g., USGS, COGCC, PADNR); a peer reviewed publication; from a source with a documented Quality Management System (QMS) including a QMP; or from a source with documented evidence that generally accepted methods were used in generation of data. See requirements for specific data types in Sections 2.9.2, 2.9.3 and 2.9.4.
- The source of non-direct measurement data used in the study must be identifiable and recorded along with the data as a reference. Copies of reference source data documentation must be retained in the project file and be readily retrievable.
- The reference source of data will be an element of the data record and will be included with the compiled data and database, at a minimum.
- Data qualifiers from either the source or from the Data Collector as to its quality or applicability will be an element of the data record and will be included with the compiled data and database, at a minimum.
- Data collected must be independently reviewed and approved by the Data Reviewer. This review must be documented.

2.10 DATA MANAGEMENT

Information from data collection activities will be collected and saved on computers from network websites and other sources using standard commercially available internet, internet search and data handling/manipulation software, such as Word, Excel, or other EPA approved software.

Data collection computer system requirements are specified below.

- The computers must be Shaw corporate computers with appropriate and active security system software installed including firewall and web protection with regular scans for viruses and spyware.
- The computers shall not have any software installed that is not approved by Shaw Information Technology (IT) Services.

Hardcopies of source data that will be manually entered into data compilations or database shall be maintained to facilitate checking and documenting checks of entries. This documentation must be retained in the OATS project file.

Shaw will control the review, revision, and distribution of the most recent version of the QAPP. Document control format (Sec. No., Rev. No., and Date) shall appear in the upper right-hand corner of each page of the QAPP. A signed approval form will accompany the QAPP. Any revision to the QAPP shall be circulated to Shaw and EPA project staff for review and approval. Documentation of approval is evidenced by signatures. Final approved version of the QAPP will be distributed by Shaw to all project staff.

Project staff conducting data collection, database development and implementation, and/or GIS activities shall document their work in notebooks or other means approved by the EPA QA Manager per requirements in ORD PPM 13.2, *Paper Laboratory Records* (USEPA ORD, 2006).

As this is a QA Category 1 project, permanent retention of project records is required per Agency Records Schedules 501. All planning documents (QAPPs), data, databases, GIS files, maps, project deliverables, notebooks, correspondence, etc., generated during the course of this project shall be transferred to Susan Mravik, the EPA WAM. They shall be stored in her office at RSKERC until they are transferred to RSKERC's Records Storage Room. At an as yet to be determined time in the future the records will be transferred to a National Archive facility.

Management of project data is described in Sec. 2.10, Data Management.

2.0 DATA GENERATION AND ACQUISITION

2.1 SAMPLING PROCESS DESIGN (EXPERIMENTAL DESIGN)

There are no physical sampling activities identified as part of current project activities.

2.2 SAMPLING METHODS

There are no physical sampling activities identified as part of current project activities.

2.3 SAMPLE HANDLING AND CUSTODY

There are no physical sampling activities identified as part of current project activities.

2.4 ANALYTICAL METHODS

There are no sampling or analysis activities identified as part of current project activities.

2.5 QUALITY CONTROL

There are no sampling or analysis activities identified as part of current project activities, so no

measurement QC activities are currently planned.

2.6 INSTRUMENT/EQUIPMENT TESTING INSPECTION AND MAINTENANCE

Use of instruments or equipment is not planned as part of current project activities.

2.7 INSTRUMENT/EQUIPMENT CALIBRATION AND FREQUENCY

Use of instruments or equipment is not planned as part of current project activities.

2.8 INSPECTION/ACCEPTANCE OF SUPPLIES AND CONSUMABLES

Inspection/acceptance of supplies and consumables on the project is currently anticipated to be minimal and essentially administrative or information handling related. Inspection and acceptance will be indicated and documented by noting approval along with approver's initials and dates on the packing list, receipt, or invoice.

2.9 NON-DIRECT MEASUREMENTS

Non-direct measurements are data collected from existing sources, not directly measured or generated in this project. These data are also referred to as secondary data. Secondary data will be prepared for inclusion in the database following the procedure outlined in OATS Standard Operating Procedure (SOP), SOP-03, *Hydraulic Fracturing Data Handling and Database Management* (Shaw, 2011), which is attached.

2.9.1 Scope of Data Collection

The initial activity for this project is identifying sources and surveying the type of non-direct measurement information available from hydraulic fracturing case studies and other data concerning potentially impacted surface water and groundwater. Shaw will begin with the two currently selected regional study area locations:

- Susquehanna River Basin/Marcellus Shale in Pennsylvania
- Garfield County/Piceance Basin in Colorado.

The scope of this study is further defined below for each location:

- Data from Susquehanna River Basin (PA) area will be from within the entire basin down to the Chesapeake Bay and will be from the year 2005 or later. Data from natural gas/oil well hydraulic fracturing activities will be from wells that have been put into operation in 2005 or later.
- Data from Garfield County (CO) area will be from within the county boundaries, or from locations within watersheds that impinge on the Garfield County booundary, especially in the vicinity of the towns of Silt, Rifle and Battlement Mesa, and will be from the year 2000 or later. Data from natural gas/oil well hydraulic fracturing activities will be from wells that have been put into operation in 2000 or later.

The specific regional study data targeted for collection will be of two basic types: 1) data specific to hydraulic fracturing activities at wells; and 2) data concerning natural or engineered surface water or groundwater in the vicinity of hydraulic fracturing activities, and in the appropriate time frame to register impacts and trends from hydraulic fracturing activities:

Hydraulic fracturing well data will include specific well identification and location information. Data will also include the date of record or activity as well as available specifics such as the following:

- elevation at location
- quantity of water used
- site geologic cross-sections
- water source
- water quality (such as pH, temperature, conductivity)
- water storage, i.e., natural (stream), tanks, pit, engineered impoundment
- well type, i.e., vertical, horizontal or directional
- depth fractured
- fracture directionality
- fracturing pressures
- geologic zone fractured
- site or geologic zone hydrologic information, e.g., hydraulic gradient and conductivity
- quantity of recovered water
- recovered water quality
- fate of recovered water, i.e., treated and discharged (receptor), discharged (receptor), injected (depth), etc

Groundwater and surface water data will include measurement identification information with date and location of measurements as well as available data such as the following:

- water type, i.e., groundwater, stream, lake
- location ID, i.e., well or reading/gauge station ID, etc.
- precipitation data
- location elevation
- surface water level- elevation/depth for lake, reservoir or impoundment
- groundwater level- elevation/depth
- stream stage
- stream flow or discharge rate
- water well – fractured or not fractured
- water well discharge rate
- water quality (such as pH, temperature, conductivity)

The main objective of collecting ground and surface water data is to capture behavioral and temporal trends related to hydraulic fracturing activities.

Additional data targets and additions or changes to the data collection scope may be made by the EPA based on initial survey of available information from sources. These modifications must be documented.

Data will be identified and collected from acceptable sources during information source survey activities. Sources that provide metadata or information that describes the data and their quality criteria will be assessed for suitability and the metadata will be captured or documented. The inventory of different types of data available that is developed during survey activities will be used to define database requirements as well as plan database development and GIS mapping activities.

Case study data collection may be from any number of retrospective and prospective case study nominations as listed below. These studies will typically be more focused and will involve a smaller geospatial scale than the regional studies. Case studies may be defined or amended by EPA based on results from data collection activities and specific direction will be provided on a site specific basis in technical directives.

Nominated Retrospective Case Studies:

- Washington County Pennsylvania (Marcellus Shale)
- Bradford/Susquehanna counties Pennsylvania (Marcellus Shale)
- Wise County Texas (Barnett Shale)
- Kildeer North Dakota (Bakken Shale)
- Las Animas and Huerfano Counties Colorado (Raton Basin)

Nominated Prospective Case Studies:

- Washington County Pennsylvania (Marcellus Shale)
- DeSoto Parish Louisiana (Haynesville Shale)

2.9.2 Geospatial (Locational, Elevational and Temporal) Information Requirements

Locational data should adhere to EPA National Geospatial Data Policy (USEPA, 2005), and any deviations shall be identified. Specific requirements include the following:

- Geo-Referenced Point Data – EPA policy requires geo-referenced coordinates be collected or derived, and appropriately documented in accordance to the adopted EPA/EDSC Latitude / Longitude Data Standard (USEPA, 2006). Locational data will span national to regional to site-specific scales and geospatial references and coordinate systems must be defined for each scale.
- Geospatial Data Accuracy (Locational) – In the absence of program-specific procedures addressing required minimum accuracy for geospatial data, EPA policy requires a minimum accuracy of Tier 5 (USEPA, 2005), which is described in Table 1. However, locational data will span national to regional to site-specific scales and acceptable geospatial accuracy must be defined for each scale or on a site specific basis. Data that

does not meet the established minimum accuracy requirements or data for which the accuracy is not defined and is otherwise deemed acceptable must be qualified and the qualification must be a part of the data set.

Table 1. Geospatial Accuracy Tiers

Tier Level	Accuracy and Precision	Examples of Horizontal Collection Method	Example Program Application
Tier 1	<1 m	Classical Surveying Techniques; plus GPS Carrier Phase Static Relative Position	Definition of contamination boundaries of site
Tier 2	1 – 5 m	GPS Carrier Phase Kinematic Relative Position	Definition of contamination boundaries of site
Tier 3	6 -25 m	GPS Code (Pseudo Range) Standard Position	Stack location; drinking water intake location
Tier 4	26 – 100 m	GPS Unspecified; Photo/GIS Interpolation	Site centroid; large area facility boundary
Tier 5	101 - 200 m	Urban Style Address Matching	Preliminary Site Location
Tier 6	201 - 999 m	Public Land Survey – Sixteenth Section	Prediction of Local Air Dispersion
Tier 7	1000 - 2000 m	Address Matching – Block Face	Batch Geo-coding
Tier 8	2001 - 5000 m	Census Block Centroid	State-level Population Statistics
Tier 9	> 5000 m	Zip Code Centroid	Generalized National Mapping
Tier 10	Unknown	N/A	Relative contextual data

- Geospatial Metadata – EPA policy requires metadata describing geospatial data in accordance with FGDC Content Standard for Digital Geospatial Metadata (FGDC, 1998). This includes temporal and elevation data. Data that does not meet the standard but for which documented information exists providing equivalent metadata information may be used, but this information must be documented and shall become a part of the data set in conformance to the standard. Methods used to construct metadata information must be documented.
- Elevation Data Accuracy – Elevation data for a location is needed at a minimum to recognize impacts on hydrologic conditions in relation to other locations. This data is obtained from standard Global Positioning System (GPS) readings. Minimum accuracy requirements must be established in a manner similar to that for locational (latitude and longitude) information and EPA's Tier levels may be applicable to elevation data as well.
- Map Projections and Scale - The projection associated with maps will be matched to scale and purpose. Regional and national scale map projections will be seamless across the conterminous USA and preserve area. Local scale map projections will be Cartesian

and thus preserve angle and distance.

- For data processing purposes, the Albers 2396 projection will be used for regional and national scale maps, with the following attributes:

  USGS US PROJECTION DATA
  Albers Conic Equal Area
  GRS 1980
  NAD 83

  CENTER PARALLEL 23.0°
  CENTER MERIDIAN -96.0°
  1ST PARALLEL 29.5°
  2ND PARALLEL 45.5°
  FALSE N 0
  FALSE E 0
  UNITS METERS

- For display and visualization (e.g. Google Earth), the web Mercator projection (WGW 1984 Web Mercator) is desirable.
- The UTM projection will be used for local mapping. In the event that the study involves more than one UTM zone, the State Plane coordinate system shall be considered.
- Custom maps may be necessary, and when used, a metadata file will be supplied describing in detail the associated projection.

2.9.3 Acceptance and QA Requirements for Water Quality Data

Water quality data may include a number of standard physical and analytical parameters, such as temperature, pH and conductivity. This data may be obtained from both hydraulic fracturing well information as well as groundwater well and surface water information. Acceptance and QA requirements for this data are described below and summarized in Table 2.

- Water quality data must have been acquired by methods approved by the federal government such as EPA, USGS, DOE or DoD methods, or universally-accepted methods such as ASTM or Standard Methods for Water and Wastewater (SMWW). These methods have defined data quality objectives, i.e., accuracy, precision, detection limits and quantitation limits.
- Data collected by other methods may be used, with EPA approval, but data quality objectives required by the methods must be documented and results must meet the data quality objectives required for the methods used to acquire the data. In addition the data must be qualified as "alternate method." If the methods used and data quality objectives of the methods are notably different than those from methods either approved by the federal government or universally-accepted, than the data may be used with EPA approval, but must be qualified by specifying the differences.
- In some cases it may be difficult to obtain sufficient information about methods used or their data quality objectives to evaluate data with respect to acceptance criteria. In these

cases the data may be used with EPA approval, but the data must be appropriately qualified to indicate the unknown quality information.

- Data for which the methods used to obtain the data and the data quality objectives or data quality indicator results are not defined or otherwise indicated may be used with EPA approval, but results must be qualified as "quality unknown."
- Any data that has been identified as not meeting the quality objectives of the methods used to collect the data will be rejected.

Table 2. Water Quality Data - Data Quality Objectives and Data Qualifications

Measurement	Method	Data Quality Objectives (DQO)	Acceptance	Data Qualifier
Water quality parameters, e.g., pH, temperature, conductivity	EPA, USGS, ASTM, SMWW, federal government approved or universally accepted	Method requirements	Meets DQO	None
	Other methods that are consistent with government approved or universally accepted methods	DQO consistent with government approved or universally accepted methods	Meets DQO and EPA approval	Alternate Method
	Methods notably different from government approved or universally accepted methods	Method requirements	Meets method requirements and EPA approval	Difference(s) defined
	Method or DQO Unavailable or undefined	Method or DQO Unavailable or undefined	EPA Approval	Missing Information Identified
	Unavailable or undefined	Unavailable or undefined and no indication that data did not meet method requirements	EPA Approval	Quality Unknown

2.9.4 Acceptance and QA Requirements for Other Measurement Data

Other measurement data include hydraulic fracturing operation data, groundwater level measurements, precipitation data, stream stage measurements, stream and water well discharge rates, etc. Methods for acquiring or generating this data may include those approved by the federal government such as EPA, and USGS, universally-accepted methods such as ASTM or may be standard operating procedures (SOPs) used by the source entity or even instrument/device manufacturers operating instructions (IMOI). The level to which the methods used and data quality objectives, i.e., accuracy, precision, etc., are defined or even available for this data may be variable. Acceptance and QA requirements for this data are described below and summarized in Table 3.

- Data must have been acquired by documented and approved methods with defined data quality objectives, i.e., accuracy and precision, etc., that are consistent with methods either approved by the federal government or universally-accepted, if such methods exist.
- If the methods used and data quality objectives of the methods are notably different than those from methods either approved by the federal government or universally-accepted, if such methods exist, then the data may be used with EPA approval, but must be qualified by specifying the differences.
- If sufficient information about the methods used or the quality objectives of the methods are not available to evaluate data with respect to the acceptance criteria, the data may be used with EPA approval, but the data must be appropriately qualified to indicate the unknown information.
- Data for which the methods used to obtain the data and the data quality objectives or data quality indicator results are not defined or otherwise indicated may be used with EPA approval, but results must be qualified as "quality unknown."
- Any data that has been identified as not meeting the quality objectives of the methods used to collect the data will be rejected.

Table 3. Other Measurements - Data Quality Objectives and Data Qualifications

Measurement	Method	Data Quality Objectives (DQO)	Acceptance	Data Qualifier
HF operating parameters, quantities	ASTM, USGS, SOP, IMOI	Method requirements or Procedure requirements with DQO consistent with government approved or universally accepted method(s)	Meets DQO	None
Groundwater level	EPA, SOP, IMOI			
Stream discharge rate	USGS, SOP, IMOI			
Water well discharge rate	ASTM, USGS, SOP			
Stream stage	USGS, SOP, IMOI			
HF operating parameters, quantities, Groundwater level, Stream discharge rate, Water well discharge rate, Stream stage	Methods notably different from government approved or universally accepted methods, if any exist	Method requirements	Meets method requirements and EPA approval	Difference(s) defined
	Method or DQO unavailable or undefined	Method or DQO unavailable or undefined	EPA approval	Missing information identified
	Method unavailable or undefined	DQO unavailable or undefined and no indication that data did not meet method requirements	EPA Approval	Quality Unknown

2.9.5 Other Acceptance Requirements for Data

Non-Direct measurement data requirements are described below:

- Methods used to discover and collect non-direct measurement data must be documented and generally accepted from both a technical and QA standpoint. The quality objective is to assure that discovery, identification and selection efforts are thorough, comprehensive and unbiased. Current methods are to investigate known sources that have been recommended by EPA and documented in this QAPP as well as secondary sources derived from these sources and evaluate available data for acceptance as outlined in Section, 2.9. Other methods may be searching for information in specific professional society journals, which shall be documented as a note identifying the journal, how it was selected and the time frame covered by the journal dates searched. This documentation shall be retained in the OATS project file.
- Data shall meet the spatial and temporal requirements as described in Section 2.9.2.
- In general, data must be from a known or recognized and accepted source, such as; a state or federal government agency (e.g., USGS, COGCC, PADNR); a peer reviewed publication; from a source with a documented Quality Management System (QMS) including a QMP; or from a source with documented evidence that generally accepted methods were used in generation of data. See requirements for specific data types in Sections 2.9.2, 2.9.3 and 2.9.4.
- The source of non-direct measurement data used in the study must be identifiable and recorded along with the data as a reference. Copies of reference source data documentation must be retained in the project file and be readily retrievable.
- The reference source of data will be an element of the data record and will be included with the compiled data and database, at a minimum.
- Data qualifiers from either the source or from the Data Collector as to its quality or applicability will be an element of the data record and will be included with the compiled data and database, at a minimum.
- Data collected must be independently reviewed and approved by the Data Reviewer. This review must be documented.

2.10 DATA MANAGEMENT

Information from data collection activities will be collected and saved on computers from network websites and other sources using standard commercially available internet, internet search and data handling/manipulation software, such as Word, Excel, or other EPA approved software.

Data collection computer system requirements are specified below.

- The computers must be Shaw corporate computers with appropriate and active security system software installed including firewall and web protection with regular scans for viruses and spyware.
- The computers shall not have any software installed that is not approved by Shaw Information Technology (IT) Services.

- The data collected will be secure, either by encryption of computer hard-drives when used off-site or on a secure computer network such as the ShawNet intranet system, which has additional server based security and scheduled backup systems.
- The computers will also be rebooted daily to ensure system updates are installed in a timely manner.
- Computer connections to the internet will be made using Shaw computers equipped with active and Shaw-approved anti-virus and anti-phishing software and that are compliant with Shaw corporate policy regarding computer safety.

The data collected will subsequently be used to prepare GIS maps and assemble a database of information for the study. Ideally data can be compiled directly from collected source data files into selected database software, but it may be necessary to manipulate source data and prepare compiled data files in specified format for uploading into or use by selected GIS and database software. Compiled data files are distinguished from the information database that will be developed for end user use.

Data management requirements are the following:

- Project files will be segregated from non-project files on an electronic portal site or storage location by a specified storage drive or drive subdirectory location.
- Source data files will be stored on an electronic portal site or in a designated location on a computer and organized by data type, source or source type and by format. File manipulations performed to consolidate data, files or file types to simplify file organization and data handling will be reviewed by the Data Reviewer to assure the integrity of the data file or data set is intact. This review must be documented. Approved data files will be designated as such in the file name and will be saved to a segregated location. The approved designation will be consistently applied to source data files and will be readily identifiable. Data manipulation such as that for conversion of geospatial reference coordinates or measurement data units for consistency must be approved by EPA, documented, and the process must be reviewed and checked for errors by the Data Reviewer.
- The EPA will be informed by the Data Collector of the format and method for preparing compiled data files. This applies to the process of preparing compiled data files from data set files (source data files) for use by or uploadable into GIS or database software. Data manipulation steps as well as prepared (final) compiled data files must be reviewed and approved by the Data Reviewer. This review must be documented. Approved compiled data files will be designated as such in the file name and will be saved to a segregated location. A file naming convention that is consistent and renders file types and QA approval status readily identifiable will be used to identify compiled data files.
- Compiled or constructed data files must have 100% of manual data entries checked independently against source documentation for correct transcription. This check must be documented. If errors are discovered, they will be presented to the Data Collector for

concurrence and resolution or correction. Corrections will be reviewed by the Data Reviewer for approval. If concurrence between the Data Collector and Data Reviewer cannot be achieved on potential errors, the Shaw QAO will be consulted for resolution. Data files that have been checked and approved for manual entries will be designated as such in the file name using a consistent and readily identifiable naming convention.

- Compiled data files must have ten percent (10%) of electronic data transfers checked against source data for correct transcription. This check must be documented. If errors are discovered, they will be presented to the Data Collector for concurrence and resolution or correction. Corrections will be reviewed by the Data Reviewer for approval. If concurrence between the Data Collector and Data Reviewer cannot be achieved on potential errors, the Shaw QAO will be consulted for resolution. The discovery of a transcription error will trigger an increase in the percent data transfer checks to twenty-five percent (25%) for the file. Data files that have been checked and approved for electronic data transfers will be designated as such in the file name using a consistent and readily identifiable naming convention.
- Source documentation used for manual entries must be maintained in a project file and managed as prescribed in Section 1.6, Documents and Records.
- Collected and compiled data files must be backed up on an independent computer, network server or independent data storage device daily if changes, additions or modifications have been made.
- Products from data collection, such as the source data inventory and compiled data files will be forwarded to the EPA or saved to an electronic portal site available to the EPA, or at a minimum, forwarded to the OATS File Administrator for backup within the EPA computer network system on a weekly basis if changes, additions or modifications have been made.
- Data collected off-site of the OATS contract intended for input into an onsite OATS/EPA database will be transferred electronically to the Database Manager by email or saved to an electronic portal site or common computer drive available to the developer.

The process for preparing the raw data for upload to the database is given in OATS SOP-03 (Shaw, 2011). The general process is briefly outlined as follows.

- Raw Data is provided to the Database Manager and stored on a local computer file folder, folders are named by month and date data are received;
- A sample format file or other translation matrix, which indicates how the required fields (USEPA, 2003, 2008a) are to be populated, if the data in the source files are absent or inconsistent with the EPA Valid Values (EPA, 2008b) is developed by the Technical Lead working with the Database Manager.
- If necessary, a script is developed to extract or transform the data to a form compatible with the required EDD format. The script developed will be available for review;
- Output is loaded into an EDD for upload into the database. ;
- The EDD is reviewed to assure data handling process(es) performed as intended.

- An electronic data processor (EDP) tool is used to check the data against the EDD format requirements and load the data into the database;
- The final database entries are spot-checked for correctness.
- Data files reside on local network drive, the network drives are periodically backed up on tape drives.

3.0 ASSESSMENT AND OVERSIGHT

3.1 ASSESSMENTS AND RESPONSE ACTIONS

Assessors do not have stop work authority, however, they can advise the EPA Work Assignment Manager (WAM) if a stop work order is needed in situations where data quality may be significantly impacted. The EPA WAM makes the final determination as to whether or not to issue a stop work order. The EPA WAM may consult with the EPA technical lead to assist with this determination. . This does not preclude the WAM from working through the proper contract channels to accomplish this activity.

3.1.1 Assessments by Shaw

Technical Systems Assessments (TSA) will be used to monitor project activities for implementation and conformance to the requirements of this QAPP and related Quality documents. The TSA will include assessment of data collection activities, documentation, quality checks, record management and reporting.

A TSA will be performed quarterly by the Shaw QAO. Results of the TSA will be reported to the OATS Program Manager (PM), EPA Quality Assurance Manager (QAM), and EPA WAM. Nonconformances will be identified as findings, recommendations or observations. Corrective actions for findings will be developed, documented and proposed by the OATS PM within 15 days of TSA report issuance to the Shaw QAO, EPA QAM, and EPA WAM for concurrence and approval. The Shaw QAO is responsible for ensuring resolution of findings. The Shaw QAO will monitor implementation and completion of corrective actions. After all corrective actions have been implemented and confirmed to be completed, the Shaw QAO shall send documentation to the OATS PM, EPA QA Manager and EPA WAM that the TSA is closed. TSA reports and responses shall be maintained by the OATS File Administrator in the project file.

3.1.2 Assessments by EPA

Technical Systems Audits (TSAs) and Audits of Data Quality (ADQs) will be conducted early in the project by the EPA to allow for identification and correction of any issues that may affect data quality. Detailed checklists, based on the procedures and requirements specified in this QAPP, will be prepared and used during these TSAs. ADQs will be done on a representative sample of the data. These audits will be conducted with contract support from Neptune and Co.

Audit reports will be prepared by the QA support contractor, which will be reviewed and

approved by the EPA QA Manager prior to release. Audit reports shall be sent and copied to the EPA PO and supervisor, EPA Technical Lead, EPA WAM, GWERD Division Director, and OATS PM. Specific actions will be identified in the reports. For assessments that identify deficiencies requiring corrective action, the OATS PM must provide a written response, within 10 working days, to each finding and observation to the EPA QA Manager and WAM, which shall include a plan for corrective action and schedule. The EPA QA Manager will review the written response to determine its appropriateness and provide feedback to the OATS PM, if necessary. The OATS PM is responsible for ensuring resolution of audit findings. The EPA QA Manager will monitor implementation and completion of corrective actions. After all corrective actions have been implemented and confirmed to be completed, the EPA QA Manager shall send documentation to the same parties that received the audit report that the audit is closed. Audit reports and responses shall be maintained by the OATS File Administrator in the project file and the EPA QA Manager in the QA files, including the QLOG database.

3.2 REPORTS TO MANAGEMENT

Meetings with the EPA will be held as scheduled or directed by the EPA in the Technical Directive (TD). Meeting invitees will include the EPA Project Officer, EPA WAM, EPA Technical Lead, EPA QA Manager, Shaw PM and Shaw Database Manager with the Shaw Data Collector, Shaw Data Reviewer and Shaw QAO participating by teleconference. If the Shaw PM does not participate in these meetings he shall be supplied the information provided in these meetings in an update document, notes or verbal briefing by the Shaw QAO.

In addition Monthly Status Reports will be made by the Shaw PM to the EPA Project Officer, EPA WAM and EPA QA Manager.

A Quality report to management (Shaw PM) will be made quarterly in conjunction with the TSA report. The Quality/TSA report will discuss the results of the TSA, status of the quality management program and other related issues.

4.0 DATA VALIDATION AND USABILITY

4.1 DATA REVIEW, VERIFICATION AND VALIDATION

As stated in Sections 2.9 and 2.10, products of project data collection activities will be reviewed and approved from a technical and QA standpoint by the Data Reviewer using criteria in 2.9.2, 2.9.3, 2.9.4 and 2.10. This is a documented verification process in which the existing data mined from sources will be reviewed for the following:

- Applicability of the data for the project objectives will be reviewed by examining the source documentation for concurrence with the collector's interpretation and use.
- Usability of the data will be assessed based on the documented or perceived quality of the source, metadata, available data quality indicators, measurement performance or

methods of generation.

- Completeness in collection of data and quality related information available from the source
- Accuracy in interpreting and assembling the information in a usable format
- Format in which the data is organized and compiled with respect to EPA direction, user understanding, ease of use, and achieving project goals.

In addition, Shaw will use procedures described in 2.10 to verify that data transcribed from existing data sources have been transcribed accurately. Compiled or constructed data files must have manual data entries checked independently against source documentation for correct transcription. Compiled or constructed data files must have ten percent (10%) of electronic data transfers checked against source data for correct transcription. These checks must be documented.

4.2 VERIFICATION AND VALIDATION METHODS

The Shaw Data Collector will request a review be performed of compiled data files that have undergone his review and determination as to their acceptability. The review request directed to the Data Reviewer will be documented and will include a description of the data files to be reviewed. The Data Reviewer will have access to the project data files in a restricted electronic portal location and will review them against the criteria in 2.9.2; 2.9.3, 2.9.4 and 2.10 to determine their acceptability, and document the review. The results will be reported to the Shaw Data Collector who shall then identify this data as validated as described in Section 2.10, if it was found acceptable. Only verified and validated data will be included in data compiled for database entry.

4.3 RECONCILIATION WITH USER REQUIREMENTS

The development of this information database will facilitate use of the information to achieve overall study goals, however, this is an early stage of the study process. The information database is intended to provide a basis for additional activities to continue the study. To that end, the product of this work should provide a thorough and open framework to support further work. Limitations on the use of the collected information will be based on the final outcomes of the information development process.
Project deliverables will be reviewed internally by Shaw and externally by the EPA Technical Lead and WAM to ensure they meet EPA's requirements. Shaw shall describe data quality and data limitations in its project deliverables so that later data users may determine if the data are of sufficient quality for their use.

5.0 REFERENCES

ANSI/ASQC E4-1994, *Specifications and Guidelines for Environmental Data Collection and Environmental Technology Programs,* ASQ Quality Press, Milwaukee, WI, www.asq.org.

FGDC, 1998, *Content Standard for Digital Geospatial Metadata,* FGDC-STD-001-1998, Federal Geographic Data Committee, June 1998.

Shaw, 2009, *On-Site Analytical and Technical Support Services Quality Management Plan for the NRMRL/GWERD, QA-001, March 01,2009, Edited October 10, 2010.*

Shaw, 2011, *Hydraulic Fracturing Data Handling and Database Management, OATS SOP-03, December 13, 2011.*

USEPA, 2002a, *Overview of the EPA Quality System for Environmental Data and Technology,* U.S. Environmental Protection Agency, EPA/240/R-02/003, November 2002.

USEPA, 2002b, *EPA's Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency (Information Quality Guidelines),* U.S. Environmental Protection Agency, EPA/260R-02-008, October 2002.

USEPA, 2002c, *Guidance for Quality Assurance Project Plans,* U.S. Environmental Protection Agency, EPA QA/G-5, EPA/240/R-02/009, December 2002.

USEPA, 2003, Electronic Data Deliverable (EDD) Specification Manual, Version 1.1, USEPA, Region 5; June

USEPA ORD, 2006, *Paper Laboratory Records,* Office of Research and Development, Policies and Procedures Manual, Chapter 13.2, December 2006.

USEPA, 2005, *National Geospatial Data Policy,* U.S. Environmental Protection Agency, CIO Policy Transmittal 05-002, August 2005.

USEPA, 2006, *Latitude/Longitude Data Standard,* Standard No.: EX000017.2, EPA/EDSC Data Standard, January 6, 2006.

USEPA, 2008a, Comprehensive Electronic Data Deliverable (EDD) Specification Manual, Version 2.0, USEPA, Region 5; August, 2008

USEPA, 2008b, Comprehensive Electronic Data Deliverable (EDD) Valid Value Appendix, Specification Manual, Version 2.0, USEPA, Region 5; August, 2008

USEPA, 2011a, *Plan to Study the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources*, Office of Research and Development, U.S. Environmental Protection Agency, EPA/600/11/001, November, 2011. Available online at: http://water.epa.gov/type/groundwater/uic/class2/hydraulicfracturing/index.cfm

USEPA, 2011b, *Hydraulic Fracturing Support*, Work Assignment No. 2-10 (HF), EPA Contract No. EP-C-08-034, February 7, 2011.

ATTACHMENT

STANDARD OPERATING PROCEDURE
HYDRAULIC FRACTURING DATA HANDLING AND DATABASE MANAGEMENT
OATS SOP-03

1.0 SCOPE/OVERVIEW

This SOP details the required elements for receiving, manipulating, loading and storing of hydraulic fracturing data for GWERD Hydraulic Fracturing Research case studies. The QAPP that covers these activities is entitled "Data Collection/Mining for Hydraulic Fracturing Case Studies." The intent of the SOP is to document the approved process including QA/QC requirements for preparing data that has been received from various sources, and importing it into an EQuIS-built database.

2.0 RESPONSIBILITIES

2.1 **Database Manager** – Receive raw data intended for use in Hydraulic Fracturing Research case studies and process the data for loading into a project database. Notify the Shaw Technical Lead of problems associated with the raw data package. Move the raw data through the process and document the steps involved from data receipt all the way to importing the data to the EQuIS database. Respond to reviewer's comments, resolve issues and perform corrective actions in a timely manner to expedite the data receiving / loading process. Assure that processed data progress through the review process in a timely manner. Assist with the review process by performing raw data QC and transcription checks as needed or requested.

2.2 **Shaw Technical Lead / Data Collector -** Communicate instructions to the Database Manager regarding handling of source data files and processing of source data to be compliant with EPA Electronic Data Deliverable (EDD) requirements prior to loading into the EQuIS database. Perform final review of uploaded data if needed in a timely manner and document the review and results. Notify the Program Manager, as appropriate, of any problems associated with the data or the data receiving / loading process.

2.3 **Data Reviewer** – Perform review of processed data and the data handling process steps from the time the data is received all the way through importation into the data warehouse. Document reviews and resolve review issues with the Database Manager.

3.0 PROCEDURE

The data loading process is initiated upon receipt of a data file to be loaded into a database. A data handling process for the data file is developed based on decisions regarding the extraction of data from complex source fields, how to translate codes from the source files to EPA Valid Values (EPA, 2008b) and what values to use where required fields of the electronic data deliverable (EDD) are not present in the source file. The specific transformations/translations are performed on a case by case basis and will be documented as part of the discovery and review process. as described in the QAPP (USEPA, 2011).

The **Database Manager** plays a central role in guiding the data through the data handling process. The work flow from receipt of data to upload to the EQuIS database is illustrated in the flowchart shown in Figure 1 and described below.

1. Raw Data are provided to the **Database Manager** either through a secure FTP Site, or in an email, or as a package from a Shaw data sharing site (Shaw XNet);
2. On receipt, the data are downloaded into a local computer file folder (e.g. HF Meeting 05.11), named by month and date data are received;
3. The data is reviewed and compared to EPA Region 5 (R5) EDD requirements by the **Technical Lead** working with the **Database Manager** to determine if data needs to be extracted or transformed, if data is absent for required fields (EPA, 2003, 2008a), or if data in the source files are inconsistent with the EPA Valid Values (EPA, 2008b)**.**
4. A sample format file or other translation matrix, which indicates how the required fields are to be populated, if data in the source file(s) are absent or inconsistent with the EPA Valid Values is developed by the **Technical Lead** working with the **Database Manager**.
5. If necessary, the **Database Manager** assures that a script is developed to extract or transform the data to a form compatible with the required EPA R5 EDD format (EPA, 2003, 2008). Needed scripts will be developed by the **Database Manager** in consultation with Shaw personnel familiar with writing database scripts;
6. The script developed will be available for review;
7. The EDD script product is reviewed to assure the script process(es) perform as intended. The review is documented.
8. The source file code definitions in the EDD product are compared and translated to EPA Valid Values (EPA, 2008b) according to the sample format file or translation matrix.
9. Constant data for required fields that are absent in the EDD are added to conform to EPA R5 EDD requirements per the sample format file or translation matrix.
10. A QA review of the EDD is performed to assure code definitions are translated and required fields are populated accurately as described in the QAPP. The review is documented.
11. The EDD file is given an appropriate name based on the format file and source file, and saved in a local computer folder;
12. The EQuIS 5 Professional Electronic Data Processor (EDP) tool is used to load the saved text files into the EQuIS database;
13. The final database entries are spot-checked for correctness as determined by decisions made in steps 3 and 4 above. The spot-check process including details about which data was checked and how it was checked is documented.
14. Data files reside on local shared **L Drive** under (L:\Lab\CSMOS\Hydraulic Fracturing 2011\Chem), which will be backed up every day and also periodically backed up on tape drives and stored in three different locations. The EQuIS database resides in Oracle, also on the network drive.
15. All folders mentioned above, which are on the **Database Manager's** local computer, are synchronized with the **L Drive** under (L:\Lab\CSMOS\Hydraulic Fracturing 2011).

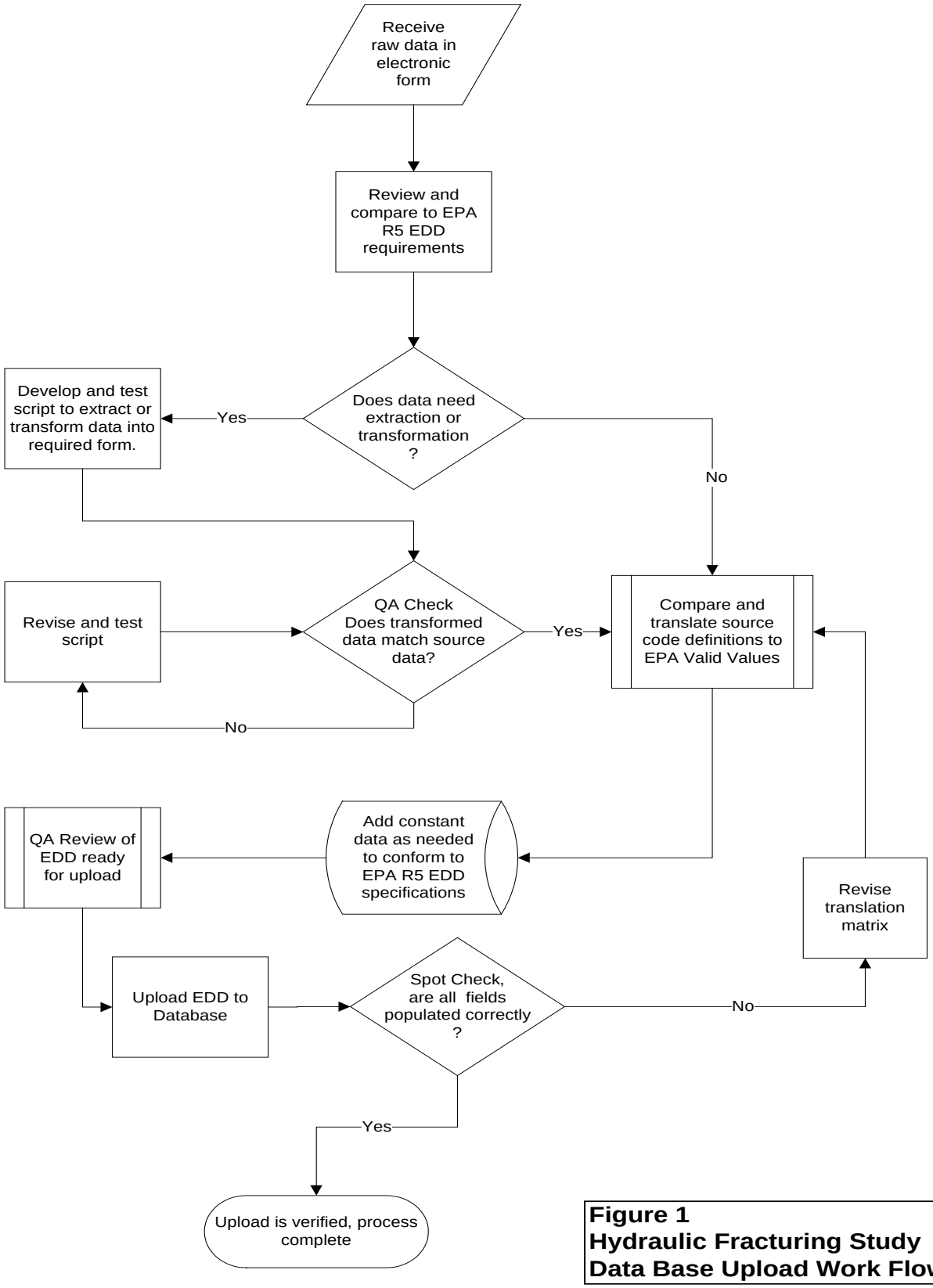

Figure 1
Hydraulic Fracturing Study
Data Base Upload Work Flow

4.0 QUALITY ASSURANCE

Throughout the process, all steps will be carefully documented in dedicated notebooks according to the procedures in the GWERD QAPP for Data Collection/Mining for Hydraulic Fracturing Case Studies, 2011.

For each data file that is loaded into a database the steps in the data handling process (Section 3.0) must be independently verified by the **Data Reviewer** and this verification must be documented. Because of the large number of records that may be involved in this process, the verification will be performed by randomly spot checking end database entries against received data to assure the data handling process has performed as intended. The spot-check verification should include data that represents all data processes performed and span across data fields or types and the entire data set. The review should be such that all data handling process steps across the entire data set are verified on as many different data types as is reasonably achievable.

5.0 REFERENCES

U.S. Environmental Protection Agency , 2011, GWERD Quality Assurance Project Plan for: Data Collection/Mining For Hydraulic Fracturing Case Studies, May 2011.

U.S. Environmental Protection Agency, 2003, Electronic Data Deliverable (EDD) Specification Manual, Version 1.1, USEPA, Region 5; June

U.S. Environmental Protection Agency, 2008a, Comprehensive Electronic Data Deliverable (EDD) Specification Manual, Version 2.0, USEPA, Region 5; August

U.S. Environmental Protection Agency, 2008b, Comprehensive Electronic Data Deliverable (EDD) Valid Value Appendix, Specification Manual, Version 2.0, USEPA, Region 5; August