

A. Project Management
A1. Title Page

Upper White River Basin Geospatial Database Development Quality Assurance Project Plan

**The Ozarks Environmental and Water Resources Institute
(OEWRI)**

Missouri State University

May 5, 2006

A1a. Approval Sheet

Jacquelyn Ferguson, EPA Technical Project Manager/Officer

Signature: _____ Date: _____

Dianne Harris, EPA QA Manager/Officer

Signature: _____ Date: _____

Steve Stewart, Director, Upper White River Basin Foundation

Signature: _____ Date: _____

Robert T. Pavlowsky, Ph.D., Director OEWRI (PI)

Signature: _____ Date: _____

Rex G. Cammack, Ph.D., Cartography/GIS (co-PI)

Signature: _____ Date: _____

Marc R. Owen, Project Supervisor OEWRI

Signature: _____ Date: _____

Gopala Borchelt, Project Assistant

Signature: _____ Date: _____

A2. Table of Contents

Section

A. Project Management	1
A1. Title Page	1
A1a. Approval Sheet	2
A2. Table of Contents	3
A2a. List of Figures	4
A2b. Glossary of Terms	4
A3. Distribution List	5
A4. Project/Task Organization	5
A5. Problem Definition/Background	7
A6. Project/Task Description	7
A7. Quality Objectives and Criteria	8
A8. Special Training/Certification	8
A9. Documents and Records	8
B. Data Generation and Acquisition	9
B1. Sampling Process and Design	9
B2. Sampling and Image Acquisition Methods	9
B3. Sample Handling and Custody	9
B4. Analytical Methods	9
B5. Quality Control	10
B6. Instrument/Equipment Testing, Inspection, and Maintenance	10
B7. Instrument/Equipment Calibration and Frequency	10
B8. Inspection/Acceptance for Supplies and Consumables	10
B9. Data Acquisition (Non-direct Measurements)	10
B10. Data Management	11
C. Assessment and Oversight	13
C1. Assessments and Response Actions	13
C2. Reports to Management	14
D. Data Validation and Usability	14
D1. Data Review, Verification and Validation	14
D2. Verification and Validation Methods	14
D3 Reconciliation with User Requirements and Data Objectives	14

A2a. List of Figures

Figure 1. Organizational Chart 6
Figure 2. Data Processing Flow Chart 12

A2b. Glossary of Terms

- Co-PI – Co-Principal Investigator
- ESRI – Environmental Systems Research Institute
- GGP – Department of Geography, Geology, and Planning
- MDNR – Missouri Department of Natural Resources
- MSU – Missouri State University
- NAD – North American Datum
- NPDES – National Pollutant Discharge Elimination System
- OEWRI – Ozarks Environmental and Water Resources Institute
- PI – Principal Investigator
- QA – Quality Assurance
- QC – Quality Control
- SDE – Spatial Data Engine
- UTM – Universal Transverse Mercator
- UWRB – Upper White River Basin
- UWRBF – Upper White River Basin Foundation

A3. Distribution List

Jacquelyn Ferguson, EPA Technical Project Manager
Dianne Harris, EPA QA Manager
Steve Stewart, Director, Upper White River Basin Foundation
Robert T. Pavlowsky, Ph.D., Director OEWRI (PI)
Rex G. Cammack, Ph.D., Cartography/GIS (co-PI)
Marc R. Owen, M.S., Project Supervisor, OEWRI
Jim Richburg, M.S., Project Supervisor, OEWRI
Gopala Borchelt, Project Assistant, Missouri State University
Ralph Davis, University of Arkansas

A4. Project/Task Organization

Principal Investigator – Provides project oversight, budgetary control, and approves all documentation and project plans. This position is the primary contact for the project.

Co-Principal Investigator – This position is tasked with providing geospatial database management expertise and advice during creation and QA/QC of the database and project plan. This includes hardware/software support, database structure/delivery implementation, and oversight of QA/QC procedures.

QA Manager – The QA manager will be in charge of performing QA procedures for verification and validation of datasets to be placed in the geospatial database.

Project Supervisor – This position oversees the work of the Graduate Assistant and manages the project. This person is in charge of writing the QAPP and associated SOPs and preparing all project related documentation.

Project Assistant – The primary project worker tasked with executing the project plan.

Principal Data Users – The principal users of the data are watershed groups and managers engaged in the water quality management planning process. These data will be available via the internet for download.

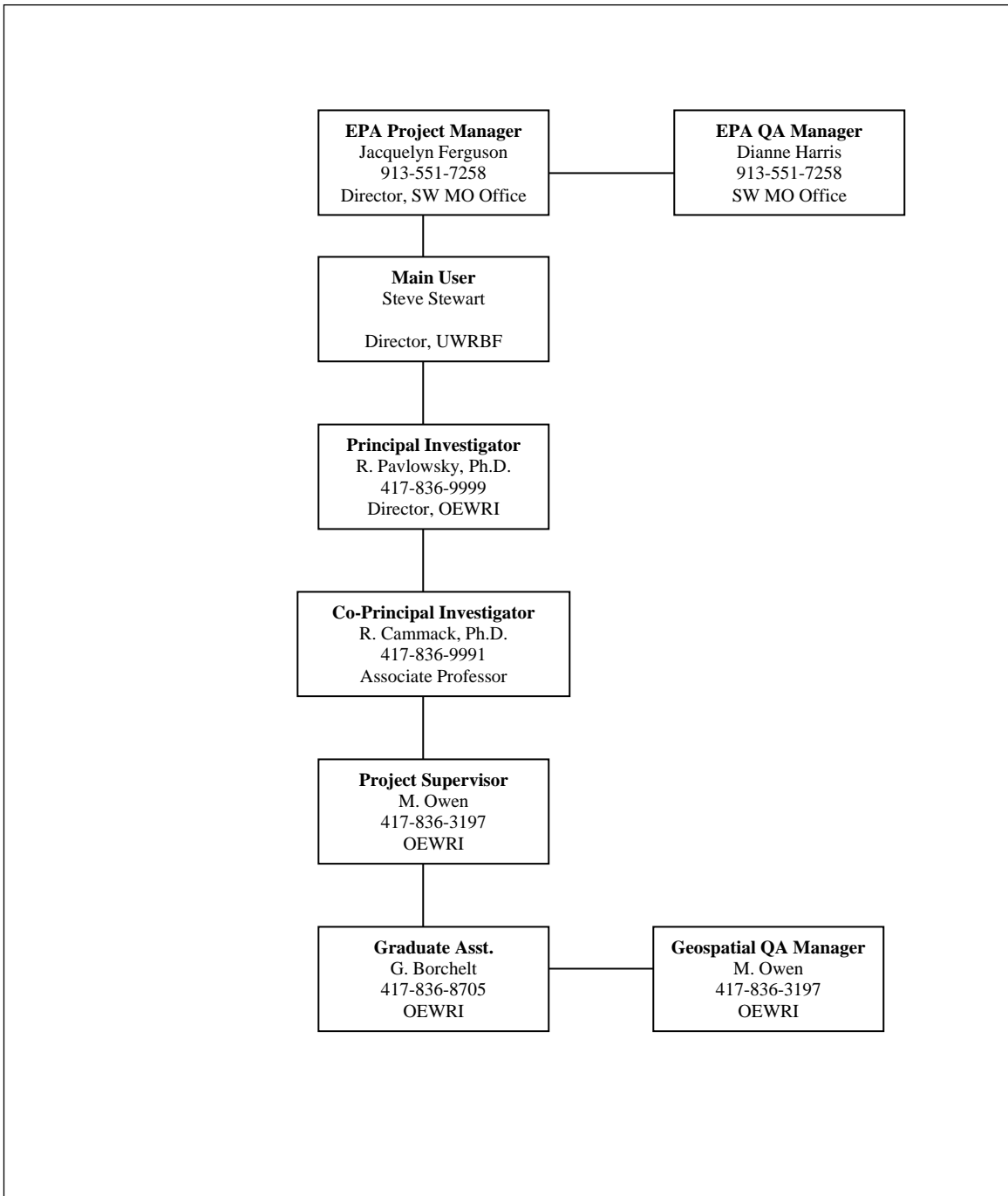


Figure 1. Organizational Chart

A5. Problem Definition/Background

The goal of this project is to develop a seamless geospatial database between Missouri and Arkansas for the Upper White River Basin. These data will be used by watershed management groups/stakeholders to develop watershed management plans. Spatial datasets will be focused on watershed scale features important to water quality. Watershed management groups/stakeholders involved in creating and implementing the water quality management plan will use these data to help identify and model sources of pollution in their specific watersheds.

A6. Project/Task Description

- The Upper White River Basin is located in SW Missouri and NW Arkansas. Existing geospatial data is housed at various state and federal sponsored geodatabase websites and is available for download. Unfortunately, datasets for any one particular feature do not always match across state boundaries. This project will:
 1. Identify datasets needed for watershed managers to create water quality management plans,
 2. Locate these data on the internet,
 3. Consolidate these into one database for the basin with matching features and consistent attribute information across state lines, and
 4. Make these data accessible over the internet.
- Datasets needed for the water quality management plan will be identified by the following groups at a database planning meeting:
 - Upper White River Basin Foundation
 - Missouri State University (OEWRI)
 - University of Arkansas

Data acquisition, processing, merging, and database management will be completed by the Project Assistant under the guidance of the Co-Principal Investigator and the Project Supervisor. Quality assurance procedures will be performed by the Geospatial QA Manager.

- Milestones
 1. Geospatial database needs identification by group, completion year 0.5
 2. Dataset acquisition, processing, and QA/QC, completion year 1
 3. Database creation and structuring in ArcSDE, completion year 1.5
 4. Website creation and FTP capability, completion year 2
- Deliverables
 1. Digital GIS Database
 - A complete watershed geospatial database for the Upper White River basin will be delivered to the UWRBF.
 2. Five Thematic Maps
 - Multiple copies of 5 different thematic maps with a water quality theme will be designed with 3 copies of each to be printed and delivered to the UWRBF.

3. Internet Accessible Geospatial Database
 - Development of a seamless database harmonized across state boundaries in conjunction with the University of Arkansas will be stored on a server housed at Missouri State University in the Department of Geography, Geology, and Planning with on-line access to geospatial and water quality data.
- Tasks
 1. Geospatial database needs will be identified by the group. This will be accomplished during a database planning meeting between the Missouri and Arkansas groups. This task will complete milestone #1 and is the first step in accomplishing deliverable #1.
 2. Project Assistant will acquire data and combine to create a digital GIS database for the completion of milestone #2 and complete deliverable #1.
 3. Project Assistant will prepare and print 5 thematic maps to complete deliverable #2.
 4. Project Assistant will combine Arkansas and Missouri data into one database and match features and attributes across state lines. These data will be managed in ArcSDE on the Missouri State Server. This completes milestone #3, and a portion of deliverable #3.
 5. QA Manger will use approved procedures to QA/QC database for posting on the internet to complete portion of deliverable #3.
 6. Project Supervisor and Project Assistant will gather and organize data for posting on project website to complete deliverable #3.

A7. Quality Objectives and Criteria

Six quality objectives have been identified for this project.

1. Datasets will be gathered from federal or state sponsored internet sites
2. Data must be accompanied by metadata that is up to date.
3. All data will be projected into the following coordinate system and datum:
UTM Zone 15N NAD 83
4. All line work and raster images will match visually across state lines
5. All attribute data will be standardized between the states
6. Data created for this project will adhere to the same standards as the acquired data.

A8. Special Training/Certification

No special training or certification is necessary for this project. Project Assistant is proficient at ESRI ArcGIS software and is able to perform all tasks. Further assistance for database management will be provided by Co-Principal Investigator.

A9. Documents and Records

- All documents and records will be distributed to project staff via email. Final copies and backup copies will be managed by the project supervisor.
- The following documentation will be included with each final product
 1. Digital GIS database
 - A meeting summary of the database planning meeting.

- Metadata detailing all procedures for each dataset this includes metadata from secondary data sources. All metadata will adhere to the Federal Geographic Data Committee standard.
 - A Geospatial Data Quality Control Checklist for each dataset.
 - A Quality Assurance Checklist for each dataset.
 - 2. Five thematic maps
 - A map creation summary for each map.
 - 3. Internet accessible database
 - A schematic of database structure with summary.
- All data will be housed on the OEWR server located in the GGP Department at Missouri State University. Data will be updated as required. After five years a data review will determine what will happen to the data in the future.

B. Data Generation and Acquisition

B1. Sampling Process and Design

The purpose of the geospatial database creation portion of this project is to integrate existing data for the Upper White River Basin (UWRB). For this project, the UWRB is defined as the watershed drainage area (20,000km²) upstream of the dam at Bull Shoals Reservoir covering Southwest Missouri and Northwest Arkansas. “Existing data” will be compiled for this project by acquiring data online from federal and state databases and by sharing data with project partners at the University of Arkansas. While the majority of the data used for this project already exists, two datasets will be created as part of the water quality monitoring portion of the overall project. These datasets are the sample location and delineation of the upstream drainage areas at those sample locations. Sampling sites will be located at 19 USGS gauging stations throughout the UWRB. Each gauging station has a latitude and longitude which will be used as the X and Y for each sampling point in the dataset. This is a routine process for point generation in GIS software. The points generated will then be used as the “pour point” in the watershed delineation process for each sample site. The watershed delineation process will create a polygon data layer of the drainage area upstream of each sampling location.

B2. Sampling and Image Acquisition Methods

The latitude and longitude information for each USGS gauging station used as a sampling location is published on the USGS real-time website for that station. The website for the USGS Real-Time Gauging Station information is < <http://waterdata.usgs.gov/nwis/rt>>. This information is sufficient to locate a “pour point” for the watershed delineation process. Digital Elevation Models to perform the watershed delineation process will be acquired from the Missouri Spatial Data Information Service web site (<http://msdisweb.missouri.edu/>).

B3. Sample Handling and Custody

No hardcopy imagery or data sources will be used for this project.

B4. Analytical Methods

No analytical methods were used for this project.

B5. Quality Control

The quality control methods used by various entities in creating the spatial data layers acquired for this project are expected to be acceptable due to the large area and the scope of this project. These data layers are expected to present an accurate and sufficiently detailed picture of the watershed features to be able to extract general watershed and water quality trends and relationships.

Each feature and raster image will be checked using the Geospatial Dataset Quality Control Checklist. One checklist will be generated for each dataset. A copy of the check list can be found in the Data Management (B10) element.

The quality control procedures for the acquired datasets are described here:

Acquired Data :

- ✓ Data are acquired from a federal or state sponsored website.
- ✓ Data are accompanied by metadata that is up to date.

B6. Instrument/Equipment Testing, Inspection, and Maintenance

No data collection instrumentation will be used for this project.

B7. Instrument/Equipment Calibration and Frequency

No data collection instrumentation will be used for this project.

B8. Inspection/Acceptance for Supplies and Consumables

Not applicable to this project.

B9. Data Acquisition (Non-direct Measurements)

Data gathered for this project will be identified by the overall group. These data will be used for watershed management planning by groups working in the UWRB. Since the objective of this project is to harmonize data across state lines, it is already known that much of the data are available from various federal and state websites.

The acquisition of spatial datasets will be performed by the Project Assistant using a MSU workstation computer to download geospatial data layers from a variety of federal and state websites. Below is a list of websites that will be used, but the project is not limited to these sites.

- Missouri Spatial Data Information Center (MSDIS)
(<http://msdisweb.missouri.edu/>)
- Center for Advanced Spatial Technologies (CAST)
(<http://www.cast.uark.edu/cast/research/lulc/index.html>)
- Geostore (<http://www.geostor.arkansas.gov/Portal/index.jsp>)
- USDANRCS National Spatial Data Gateway (<http://datagateway.nrcs.usda.gov/>)

Acceptance criteria for data from various websites are outlined in Quality Control (B5) element above and detailed in the Geospatial Data Quality Control Checklist. An example of this checklist is included in the Data Management (B10) element.

B10. Data Management

This element describes the following database management procedures:

- Data backup, storage, and security
- Database construction and updating
- Data processing flow charts
- Quality assessment and correction
- Naming conventions

Data backup, storage, and security

Geospatial data will be stored in an ArcSDE (SQL) database working from a RAIDED server in the GGP Department at Missouri State University. ArcSDE allows for a more efficient storage and control of the data. During database construction only the Project Assistant, Project Supervisor, and Co-PI will have access to the data. Once the database is constructed and has been accepted by the QA officer, the database will be made available through a website where a user can request to download data from the server.

Database construction and updating

This section describes the database construction process for the acquired vector and raster datasets:

Vector: Datasets downloaded for Missouri and Arkansas will be projected into UTM Zone 15N NAD 83. These data will be combined using the “merge” command in ArcInfo toolbox. These data will be clipped to the UWRB boundary using the “clip” command in ArcInfo toolbox. These layers will be managed in ArcCatalog with final storage in an ArcSDE database.

Raster: Datasets downloaded for Missouri and Arkansas will be projected into UTM Zone 15N NAD 83. These data will be combined using the “append” command in ArcInfo toolbox. These data will be clipped to the UWRB boundary using the “gridclip” command in ArcInfo workstation. These layers will be managed in ArcCatalog with final storage in an ArcSDE database.

Quality assessment and correction

Data layers collected for Missouri and Arkansas portions of the UWRB study area will not match exactly across state boundaries due to their being originally created by different companies using different classification schemes. An example is land use datasets for both states which use a different classification scheme. These datasets will be reclassified to a simpler classification that will be able to encompass both of the classifications in the two state data sets. A similar reclassification will be performed for any datasets that are not matching completely across state boundaries. All analytical methods used in creating the datasets, both created and acquired, will be defined and described in the metadata file associated with each spatial data file.

Assessment techniques for both created and acquired data are described here:

Acquired Data -

All spatial data layers will be checked by the Project Assistant for display errors and attribute errors. Specific assessment techniques are described in Assessments and Response Actions (C1) element.

Created Data -

Each watershed boundary will be assessed visually and by area comparison. Specific assessment techniques are described in Assessments and Response Actions (C1) element.

Data processing flow charts

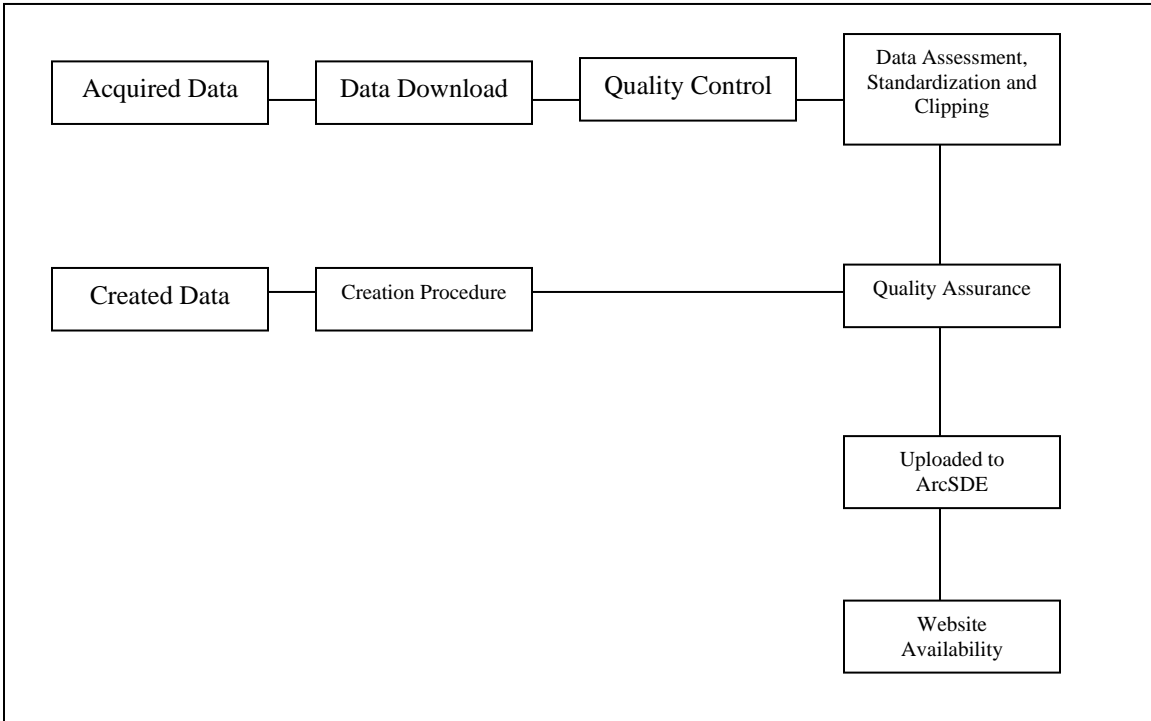


Figure 2. Data Processing Flow Chart

Naming Conventions

The naming conventions for both acquired and created data are described below:

Acquired:

Data from both Missouri and Arkansas are named with the state abbreviation followed by the feature name. When these data are combined “Moark” is used as the prefix to the feature named. When that feature is clipped with the UWRB boundary, the prefix is changed to “UWRB”.

- Example: Missouri roads = “Moroads.shp”
- Arkansas roads = “Arkroads.shp”
- Missouri and Arkansas roads combined “Moarkroads.shp”
- Combined roads clipped to basin boundary “UWRBroads.shp”

Created:

Sample site data points are named “site 1-19” which corresponds to the water sampling data from the water quality monitoring portion of this project. Watershed boundary names will correspond to stream name followed by “watershed”.

Example: Site 1 sample point = “site1.shp”

Wilson’s Creek Watershed = “Wilsonwatershed.shp”

C. Assessment and Oversight

C1. Assessments and Response Actions

Upon acquiring the geospatial datasets, the Project Assistant will perform a data quality assessment by looking for display discontinuities and attribute discrepancies between Missouri and Arkansas data. The following describes the assessment process looking for both types of errors with corrective actions:

Display Errors –

Finding display errors between the two states will be accomplished through visual inspection of the line work at the border. Any line work that does not match will be checked to make sure it is in the proper projection. Any projection conflicts will be corrected and the line work will be rechecked for border match. The second check will involve verification that the two datasets represent the same feature. If this action shows the two datasets do not represent continuous data, a new search for matching data will take place. If these inconsistencies cannot be corrected, these datasets will not be included in the final database.

Attribute Errors -

Anomalous data will be identified by its deviation from the expected or normal range of spatial location or value. For example, the point sources of nutrient water pollution, such as the NPDES permit data acquired from the MDNR, will be checked by the Project Assistant to determine if the location information is correct. This will be done by locating addresses listed in the attribute data or in the online permits and verifying that the listed address matches the coordinate data and are within the specified sub-watershed.

Raster datasets, such as the digital elevation models and land use images will be assessed for general accuracy by looking at the values of different types or classes of pixels and their associated spatial patterns. For example, the impervious classification in the Land use layers is expected to fall along major transportation routes and in and around populated areas. The metadata and resolution information associated with the Land use layers will be used to further determine the adequacy of this data for this project. Line data such as streams and roads will be verified by their associated metadata and can also be compared to aerial photography in certain parts of the study area. These datasets however, will not be used in trend analysis or modeling, so are not required to meet accuracy criteria outside of being visually correct for mapping purposes.

Watershed Boundary Error -

Each watershed boundary will be visually assessed for accuracy onscreen by comparing the boundaries of two adjacent watersheds. The final step will be to sum the areas for each of the smaller watershed boundary polygons (n=18) and they should add up to the area of the Upper White River Basin polygon.

If problems are identified in any of the datasets, the Project Assistant will contact the project management team for further input, clarification and assessment of possible solutions. Action will be taken by the Project Assistant to correct any data inaccuracies when there is sufficient information to support these corrective changes. Spatial data that cannot be verified or that appears to have errors that cannot be explained by resolution, acquisition date or other metadata entries may be discarded from this project.

C2. Reports to Management

The Project Assistant will report to the project management team as needed. Information regarding data acquisition, problems with data display or with matching across state boundaries will be included. Possible solutions to any problems will also be addressed.

D. Data Validation and Usability

D1. Data Review, Verification and Validation

Data will be reviewed by the QA officer prior to final acceptance and release of these data. The QA procedures will be documented by a Quality Assurance Checklist for each dataset. Basic verification of the QC and assessment processes will be performed.

D2. Verification and Validation Methods

The QA officer will be responsible for final verification and validation. For each dataset, the QA officer will verify the final products for the following items:

1. Dataset was acquired from a state or federal sponsored website
2. Dataset has updated metadata from source
3. Data is projected into UTM Zone 15N NAD 83
4. Data matches visually across state lines
5. Data attribute information matches across state lines
6. Watershed boundary data the sum of the sub-basins equal the area of the entire basin

These QA procedures will be documented of the Quality Assurance Checklist for each dataset.

D3 Reconciliation with User Requirements and Data Objectives

Data not conforming to the criteria set forth in the QC and assessment procedures will not be used in the final database until these criteria are met. If these criteria cannot be met then these data will not be used. Again, because of the large size of the Upper White River Basin Project study area and the fact that it is a baseline study to determine the general relationship trends between watershed variables and water quality, the accuracy

of the acquired datasets are determined to be adequate if they present a level that allows for sufficient extraction and interpretation of these general trends.