

13. Data Management

Quality data and effective data management are critical to well informed decision-making, long-range planning, and cooperative efforts. An effective data management approach supports informed decision making by providing efficient access to critical information and, through cost-effective cooperative data collection, provides access to more and higher quality data than participants might obtain individually. An effective approach also supports long-range planning and adaptive management by documenting results of watershed projects for effective adjustment of program objectives. It also fosters clearer understanding and agreements between cooperating entities and ensures integration of future data collected.

Data management is included in the Integrated Regional Water Management (IRWM) Plan Guidelines as one of the Plan Standards and requires that the IRWM Plan describe:

“...the process of data collection, storage, and dissemination to IRWM participants, stakeholders, the public, and the State.” (Propositions 84 & 1E IRWM Guidelines page 22)

This section describes how data are collected, validated, and shared in the region. It also includes recommendations that will result in effective sharing of existing data and protocols for collection of future data that will maintain/improve data integration. It is organized to discuss regional data needs, the proposed means and protocols for management of existing data, integration of future data, and plans for regional data sharing.

It is important to clarify that for this section the term “data” will be used in a general sense to describe all information collected to support decisions, carry out plans, and document results. These data may describe such things as conditions of infrastructure or natural systems, operational issues, project or program effectiveness. The general term refers to three types of information; (1) field/lab data which represents quantitative, scientific observations and measurements; (2) reports which may include studies, references, and evaluations; and (3) plans which may include designs, drawings, maps, or other spatial information.

13.1 Overview of Data Needs in the Region

The proposed data management system will provide a means of managing available and future data. The recommended management system should consider the following data needs, as described below.

13.1.1 Specific Data Needs Identified by the Regional Watershed Action Group

In development and review of issues and interests, the members of the Regional Watershed Action Group (RWAG) identified the following data needs for the region:

- Description of the hydrologic cycle of the region. Specifically, RWAG members noted limited information describing the hydrology and hydrogeology around and under Mt. Shasta and in the Medicine Lake Highlands area
- Historical water resources and watershed conditions
- Description of ground and surface water interactions
- Climatologic field data and patterns
- Description of potable water supply (spring and groundwater sources) timing, discharge, age
- Documentation of water quality conditions throughout the region and specifically within the Upper Sacramento River

13.1.2 Other Potential Data Needs

In addition to these specific needs, the following data may be needed to support project development and implementation:

- Municipal water use, reuse, and discharge
- Water rights information
- Industrial water use, reuse, and discharge
- Population information and trends
- Regional economic information
- Biological and physical habitat condition indicators
- Watershed condition indicators
- Information regarding available education and outreach programs and facilities
- Existing infrastructure condition and future needs
- Fuels management programs/protocols/studies
- Fire protection requirements
- Silvicultural practices
- Project effectiveness relative to basin objectives (future projects for monitoring the IRWM Plan effectiveness)

13.1.3 Data Management Needs

A data management approach will be designed and implemented in order for the aforementioned data to be effectively used by the members of the RWAG. Some data listed above already exist but need to be curated (reviewed and managed to ensure a certain level of data set quality and reliability) and made discoverable (be in a location and format that can be searched efficiently and used effectively). Some data have yet to be gathered. In general, information should be collected and stored to maximize discoverability. Field/lab data should be collected in a manner that is consistent with current collection standards so as to make it reliable and compatible with existing databases.

13.2 Data Management Approach

As efforts are made to find existing or collect new data, management of that data is imperative. Inaccessible data (dark data) are no better than data that don't exist, and proper qualification of data is important to their proper use. This section describes some principles and elements to be incorporated into the overall data management approach to ensure the most reliable and accessible data. To clarify, the data management approach refers to a holistic method of planning for, collecting, and managing data. As a tool within the approach, the RWAG will utilize a Data Management System (DMS) that will include some combination of Internet links, web portal(s), data library, and/or GIS interface. The data management approach includes the following components:

- Use of the Environmental Protection Agency (EPA) Data Quality Objectives (DQO) process for collecting new or assembling existing data and for evaluating the utility and reliability of that data
- Standardized techniques for collecting new reports, plans, and field/lab data
- Protocols for stakeholder data contribution
- Selection and implementation of a DMS

- Data validation and Quality Assurance/Quality Control (QA/QC) measures
- Protocols for sharing data collected for project implementation, for data distribution and maintaining compatibility with state databases.

13.2.1 EPA Data Quality Objectives Process

Data management is more than just a library list of documents or quality control of field samples. For an overall data management approach this plan will follow the EPA DQO methodology (see Guidance on Systematic Planning Using the Data Quality Objectives Process, EPA/240/B-06/001, Feb 2006). It is a systematic means of collecting new data or assembling existing data to support a decision (e.g. project to fund) or a finding (e.g. quantity of water available). As participants plan for and implement data collection and management efforts, the principals of Performance Criteria and Ranking/Prioritizing Criteria will be utilized to enhance efficiency.

Performance Criteria: Outlining performance criterion prior to data collection and management provides a framework to guide these efforts. For collection of new data, it provides the basis for setting protocol and limits on data collection. Even for existing information, it provides guidelines for deciding what to use and how to use it. Adherence to these principles provides defensible products and decisions utilizing available resources through:

- Logical project development: project needs and objectives are clearly defined and data management efforts are tailored to these objectives;
- Selection of compatible data collection protocol: review existing database(s) to which the data could be submitted and plan data collection/management accordingly;
- Target quality data: based on project needs and database protocol, appropriate data quality is targeted (to avoid over or under defining the item of interest) for the task and available resources (for existing data, an appropriate ranking of data quality, applicability, and compatibility is assigned);
- Maintain transparency of intent and direction: by providing clear and accessible records of this process, confidence in the integrity of data is maintained;
- Support sound conclusion: data are utilized to support sound conclusions, which may include peer or regulatory review of data use; and
- Maintain proper documentation: efforts to plan for data management, including decisions, assumptions, and recommendations will be appropriately documented to support confidence in the outcomes and information for future modification.

Rating / Prioritizing Criteria: As noted under Item 3 of the Performance Criteria, there is a need for a system of ranking and/or prioritizing data collected for this process. Data will be collected from a variety of sources, will come in varying degrees of compatibility, and may have a wide range of data quality. The Ranking/Prioritizing Criteria will be utilized to properly annotate the data so that users have a clear sense of their appropriate use. Existing and new data will be evaluated according to the following:

- Applicability and Utility for a proposed project. This criterion refers to means of categorizing data to identify to what extent they may be useful to support findings regarding a specific project or process. This applicability may be functional or geographical.
- Clarity and Completeness of data presentation. Ultimately, the objective of data collection and management is to provide a basis for decision making and planning. There needs to be some indication of the clarity and completeness of field/lab data, a report, or a plan set when it is being relied on to make decisions. The intent of this criterion is not to exclude data, but to provide an understanding of how representative the data set is and where there might be data gaps.

- Uncertainty and Variability of submitted data. An evaluation of the uncertainty and variability associated with submitted data, although similar in nature to clarity and completeness, provides another means of understanding how to use data. A data set may extensively and clearly characterize a particular condition, but due to the nature of the information desired, there may still be significant uncertainty. For example, a stream gage field data set might be extensive (100 years of information) and clear (based on United States Geological Survey (USGS) protocol with clearly defined accuracy of individual measurements) but may still have a fairly high level of uncertainty regarding a specific field data need (e.g. What will be the instantaneous peak flow on March 2, 2015?). Coupling these two criteria provides a more robust understanding of how the data should be used.

13.2.2 Description of Typical Data Collection Techniques

As noted in the description of process framework, part of the data management approach will include evaluating existing database protocols to assure compatibility. Stakeholders will evaluate state programs for applicability on specific items (e.g. Surface Water Ambient Monitoring Program (SWAMP) field data collection and submittal protocols for surface water field data). Specifics of data collection techniques will depend on the project and the type of data being collected but the techniques will follow those outlined in the data collection plan of each applicable state database. If a project seeks to accomplish objectives that do not include the collection of data that would fit into a particular state database, the best principles approach will be used, along with discussions with the project technical advisory committee, to ensure that effective, efficient, and defensive methods are identified and employed.

A number of different databases are described below, categorized by data type. This list is not exhaustive, but includes all databases described in the Department of Water Resources' (DWR) November 2012 IRWM Guidelines. The last category includes searchable databases that don't accept direct data entry. Data available through these sources must be entered through an alternate pathway. However, they represent significant data sources that can be useful when designing a project or assessment.

General Databases:

Sacramento River Watershed Information Module

The Sacramento River Watershed Information Module (SWIM) is a data management tool managed by the Sacramento River Watershed Program (SRWP) as a network for coordinating and utilizing Sacramento River Watershed information. This site provides a clearinghouse and is not necessarily intended to set protocol for data collection. The RWAG may consider utilizing the SWIM as a data management system for the region or at least linking data managed through the IRWM Plan to the SWIM system. The Upper Pit IRWM Region utilizes the SWIM as its data management system and there is a baseline of data in this tool for the Upper Pit IRWM Region. Instructions on use of this website are available at: www.sacriver.org.

California Environmental Data Exchange Network

California Environmental Data Exchange Network (CEDEN) is a system designed to facilitate integration and sharing of data collected by many different participants. The system allows for both finding and submitting data. The CEDEN data templates are available on the CEDEN website: <http://www.ceden.org>.

California Environmental Information Catalogue

The California Environmental Information Catalogue (CEIC) is a statewide metadata clearinghouse for geospatial data. Entering field/lab data and information into SWIM (above) automatically enters it into the CEIC. This database can be accessed at <http://ceic.resources.ca.gov/>.

Water Quality Databases:

Surface Water Ambient Monitoring Program

The State Water Resources Control Board created the Surface Water Ambient Monitoring Program (SWAMP). This database offers protocols for quality assurance and offers guidelines for standard operating procedures. Any group receiving state funds for surface water quality work must ensure collection of their field data is done in coordination with SWAMP standards. The website for SWAMP is: www.swrcb.ca.gov/water_issues/programs/swamp/about.shtml.

Groundwater Databases:

Groundwater Ambient Monitoring and Assessment program

Groundwater Ambient Monitoring and Assessment program (GAMA) provides a comprehensive assessment of water quality in water wells throughout the state. If you are going to complete a project that will have a groundwater quality component, it's very important that you contact the GAMA program manager before you design the field/lab data output format to enable easier transfer of field/lab data and information into the GAMA database. The preferred format is GeoTracker ESI (www.waterboards.ca.gov/ust/electronic_submittal/); simple Excel files can be problematic. Additional information on the GAMA program, as well as staff contact information, is available at www.swrcb.ca.gov/gama.

California Statewide Groundwater Elevation Monitoring

California Statewide Groundwater Elevation Monitoring (CASGEM) builds on many long-term monitoring programs and anticipates the entry of new data from recently initiated monitoring programs. Monitoring is done by local jurisdictions; there is an extensive user guide available, as well as program staff contact information. Data entry requires a monitoring entity to create a login for entering field/lab data and information into the database. CASGEM is available at www.water.ca.gov/groundwater/casgem/.

Reference-only Databases and/or Data Management Systems:

Water Data Library

This Water Data Library (WDL) collects field/lab data and information from other sources to house it all in one place. It is a good place to go for information, but primary data must be added in another place (SWIM, GAMA, SWAMP, or CASGEM). The information can be accessed at www.water.ca.gov/waterdatalibrary/.

Integrated Water Resources Information System (IWRIS)

DWR maintains this data management tool for water resources data. It is a web-based application that allows access to data in GIS-thematic layers and includes entries from the WDL. Map data can be accessed by users at www.water.ca.gov/iwris.

California Environmental Resources Evaluation System

California Environmental Resources Evaluation System (CERES) was developed to facilitate access to a variety of electronic data describing California's rich and diverse environments through a focus on three related components: technology, data, and community. It incorporates several other state databases and, as such, has found data standards to be an important consideration; it has several examples of data standards and data transfer/exchange formats. The database is available at <http://ceres.ca.gov>.

California Data Exchange Center

California Data Exchange Center (CDEC) installs, maintains, and operates an extensive hydrologic data

collection network including automatic snow reporting gages for the Cooperative Snow Surveys Program and precipitation and river stage sensors for flood forecasting. The database provides a centralized location to store and process real-time hydrologic field data and information gathered by various cooperators throughout the state. CDEC also operates a data exchange program with various federal, state, and other public agencies. Information on the tool is available at <http://cdec.water.ca.gov/intro.html>.

California Irrigation Management Information System

The California Irrigation Management Information System (CIMIS) is a program in the Office of Water Use Efficiency in DWR that manages a network of over 120 automated weather stations in the State of California. The primary purpose of CIMIS has always been to make real-time weather field data and information — useful in estimating crop water use for irrigation scheduling— publicly available. CIMIS information is available at www.cimis.water.ca.gov/cimis/welcome.jsp.

13.2.3 Protocol for Stakeholder Data Contribution

Part of the integration of resources in the region will be through shared use of data. As part of the IRWM process a website has been developed as a means of sharing data and to serve as a central location for participants to seek data and keep track of the IRWM process. This website will serve as an access point for project data. The website would function as a web portal for shared access to existing state and local databases. If a local entity does not have an electronic database, information would be submitted to the entity managing the website [currently River Exchange (REX)] for curation and addition to the database. This website includes a searchable library of available data. The website will maintain a password protected section for data that have limitations on distribution. In the future, the website could also link data to the SWIM web portal or IWRIS data management tool. A proposed future enhancement to the website would be the integration of a GIS-based mapping component, e.g. ESRI GeoPortal (information on this tool may be found at <http://www.esri.com/software/arcgis/geoportal>), for spatial reference to available data.

When stakeholders submit data or links to data for inclusion in regional or state database, the RWAG utilizes the following protocol.

1. Data submitted for inclusion in the IRWM database must be accompanied by metadata for tracking the submitted information. Required information includes:
 - Location
 - Contributing entity
 - Contact information
 - Type of data (field/lab data, report, plan)
 - Title of submittal
 - Author/publisher of information
 - Limitations on distribution
 - QA/QC information [Peer reviewed journal, data collection protocol followed, rating /prioritizing criteria (applicability/utility, clarity & completeness, uncertainty & variability)]
 - Keywords/document reference: for improved searchability of the database.
2. The information is reviewed to ensure that complete metadata is provided for acceptance in accordance with DQO principles discussed in Section 13.2, above.
3. Web portal access is as follows:
 - Web address: <http://www.uppersacirwm.org/>

- Password requirement: there is no password required for plan viewing or basic data access. The administrator may grant a username and password for data uploading or access to password-protected areas.
 - Upload instructions:
 - Obtain login and password from program administrator;
 - Select “login” button on website;
 - Enter login and password; and
 - Follow website instructions for the desired upload.
4. Special provisions for data sharing: Some data may include limitations on its use. Because the database is intended for data sharing, it will be the responsibility of the contributor to ensure that data published for use by the RWAG is appropriate for that purpose.

Note that this protocol is for data sharing within the region. Participants are encouraged to also publish data to state and other databases in accordance with their established protocols (as described in Section 3.2, above).

13.2.4 Proposed Data Management System

As noted, there are additional opportunities and options for data management that should be explored as part of this planning process. This section provides guidelines for development of a long term, robust data management tool. Based on plan development and proposed projects, the following phased approach to data management is recommended.

Phase 1: DMS Maintenance

At a minimum, the RWAG should implement plans to maintain the IRWM website as a data repository and sharing tool. Since the website has been developed through the IRWM planning project, this is the most cost- and time-effective means to provide a basic repository and accessible system for collection and sharing of information. It provides a central location for reports from projects so that RWAG members have a central location to get information to evaluate project success and inform the next round of prioritization. The group knows the location as well as the protocol for data and document upload. In addition, the website has built in levels of function (general public info as well as password protected) that allows for varying levels of access.

RWAG responsibility would be to commit to maintaining the website and data library. This could be accomplished through volunteer staff time or minimal investment of entity funds for staffing. Standardizing data collection (e.g. to conform to state database requirements) at this stage will be the responsibility of project sponsors. They will provide data to the library and separately provide data to state databases. Efforts to enforce data collection protocols would be limited.

This approach initially is best because it is simple. RWAG members are resource-limited, have project details to develop, and are likely to support the continuation of the IRWM process through in-kind efforts. For the entities that make up this RWAG, staff time is a precious commodity typically not available to manage an extensive and complex DMS.

Phase 2: Basic DMS

As the RWAG develops and projects progress, the group may look to adding dedicated staff. As mentioned, funding may be sought for a plan administrator or director to help take on this role. A plan administrator would take over the duties of maintaining the website and library. This management approach, however,

would likely include an objective to develop a basic DMS as opposed to simply maintaining a repository. The administrator would take over the responsibility of some data curating (reviewing, organizing, possibly soliciting data). This would include a more active application of the data collection and organizing protocols outlined in this chapter. The administrator would also take a more active role in working with project sponsors to standardize data collection, although it will still be the responsibility of project sponsors to assure compliance and submit to state databases.

This approach would be implemented either through additional financial support of RWAG members, or in concert with funding for a specific project. For example, projects proposed within this plan relating to staff positions could include an element of data management. For other types of projects, there will be a need to manage some data to meet funding agency reporting requirements. There would likely be enough overlap so the position could be at least partially funded through a project. This phase would build on and improve the existing website library. If done systematically, improvements in the database could be made incrementally as funding is available.

Phase 3: Expanded DMS

Continued collaboration or expansion of data collection could lead to a need for an expanded DMS. This may be a region-specific software package or may utilize other web-based data management tools such as the SWIM web portal or IWRIS data management tool in lieu of the IRWM website. Also, as tracking of geospatial data becomes more important, this tool may also include a GIS mapping component — either through SWIM or through a region-specific development. In this phase, a program manager would seek funding to support an active DMS. Project sponsors could utilize the program manager and staff for data collection in compliance with state database protocols. The program manager would also actively evaluate existing data and support the RWAG in collecting project monitoring and report information for evaluation of project success. Such a program would also support a broader evaluation of regional success in implementing plan objectives. The program would actively pursue data collection and fully utilize the protocols outlined in this chapter.

As an example, this type of DMS may be developed in conjunction with the region-wide groundwater-monitoring project that is being developed. A DMS would be integral to providing transparency and data sharing for such a project, and could be developed to benefit the rest of the region in conjunction with project specific tasks. The following elements should be part of the expansion/transition to maintain the database developed for the project and for connection with state databases:

5. Acquire and launch new or coordinate with existing GIS-based web interface for data acquisition
 - a. Update existing website access
 - b. Develop data sharing agreement
 - c. Develop connections to watershed scientific data
6. Target expansion of internet access for DACs to facilitate data availability
7. Expand data portal to include project monitoring data
8. Expand data portal interface to connect with state databases (listed in Section 3.2)
9. Continue to encourage data submission to databases listed in Section 3.2

13.2.5 Entity Responsible for Maintaining Data in the DMS

At the time of the 2018 IRWM Plan update, the region is in the process of transitioning from a rotating volunteer-based administration to establishing administrative staff to oversee the administrative duties. The River Exchange (REX) was responsible for the acquisition, management, and evaluation of data associated with the 2018 IRWM Plan update. The 2013 IRWM Plan, an interactive map of the region, project

information, and project videos are all maintained on the USR RWAG website. The new administrator will be responsible for managing data on behalf of the region, including publishing the 2018 IRWM Plan update to the website, and will be expected to review new data relevant to the region and to make the information available to the region's stakeholders.

13.2.6 Data Validation and Quality Assessment and Control Measures

Data to be utilized in the IRWM process will undergo a QA/QC process including:

- Verification that required metadata (as listed in Paragraph 3.3) is provided with submitted information to document original data collection/preparation conditions
- Submittal of QA/QC documentation for the given data generation effort will be requested. This documentation may include reference to QA/QC for an existing database or project specific quality assessment protocol (QAP). The data will be rated relative to its verifiable level of QA/QC as part of the evaluation process described in Section 3.1.
- The entity submitting the data may provide additional documentation of data validation or QA/QC measures employed.
- The QA/QC rating of data will be regularly available for comment

13.2.7 Sharing Collected Data

Data sharing will be through the IRWM website or by regular email distribution to the RWAG group. Information currently anticipated for sharing includes:

- RWAG meeting information (e.g. participant information, meeting summaries, outreach efforts)
- IRWM Plan section drafts, completed sections, and project proposals
- Reference materials supporting and/or cited in the IRWM Plan sections
- Photos, videos, and maps
- Contact information for participating groups and their representatives
- GIS / geospatial data
- Plans, reports, studies completed by participating entities

The RWAG is committed to data sharing and the DMS will be an important part of that in multiple ways. First it provides a means by which RWAG can access data that might not otherwise be available to them. Second it provides a protocol for maintaining and documenting the quality and applicability of data and therefore improving the value of data shared.

As noted previously, efforts will be made to assure proper access to data. Different levels of website access and security will be assigned to individuals. For example, most entities will be allowed to upload data to the portal. A select few will be given access to remove or edit data placed there. In every case, data utilized on the project will be backed-up locally by REX during the plan preparation phase. A database administrator may be needed after the project to maintain the data backup. The plan will direct participants to make data compatible with applicable state agencies. The data management approach will maintain data compatibility by systematically documenting the compatibility of existing data and requiring compatibility for new data that becomes part of the system.