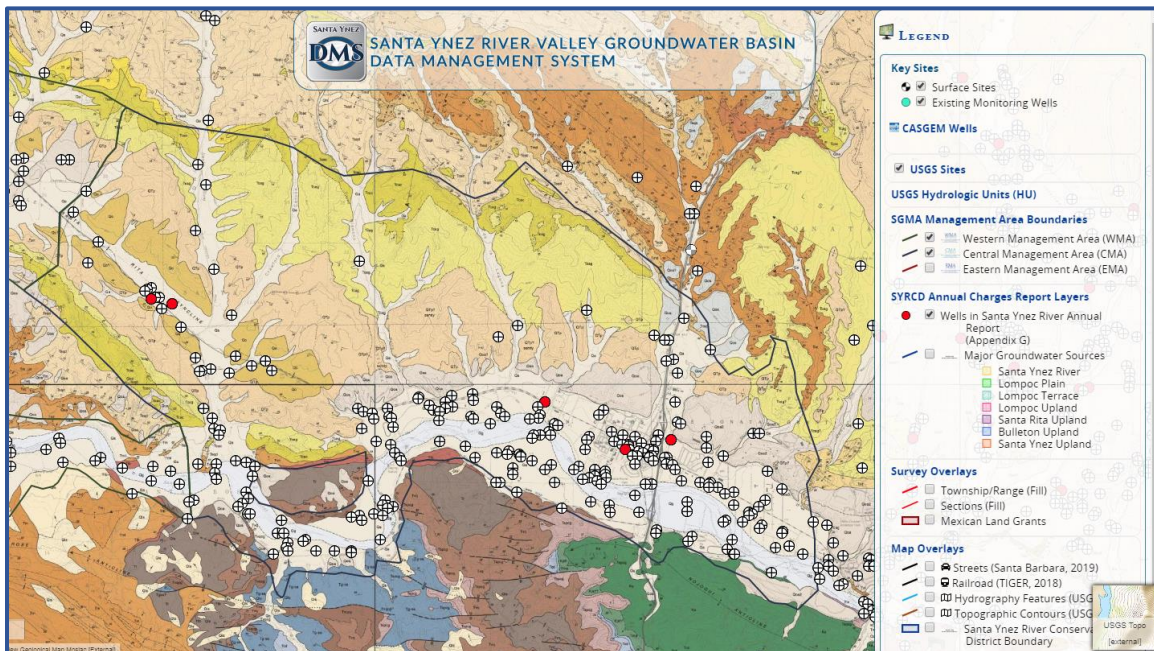


# CMA

Santa Ynez River Valley Groundwater Basin  
Central Management Area  
Groundwater Sustainability Agency

DRAFT

## DATA MANAGEMENT PLAN



Screenshot including streaming geologic maps from the USGS.

SEPTEMBER 2019

INTENTIONALLY LEFT BLANK



## TABLE OF CONTENTS

---

1	Introduction .....	1
2	Goals of DMS.....	5
2.1	Support of Hydrogeologic Conceptual Model Development.....	5
2.2	Monitoring Network.....	5
3	System Architecture.....	7
3.1	Platforms .....	7
3.2	Scripts .....	7
3.3	Data Location.....	7
3.4	Data Types .....	8
3.4.1	Database.....	8
3.4.2	Technical Report Format .....	9
3.4.3	Well and Surface Site Data .....	9
3.4.4	Geospatial Data .....	9
3.5	Backup .....	10
3.6	Intra-basin Consistency / Formats .....	10
4	Data Sources and Quality.....	13
4.1	Data Sources.....	13
4.1.1	Federal Data Sources.....	13
4.1.2	State Data Sources .....	13
4.1.3	County Data Sources .....	14
4.1.4	Municipal, Water District, and Other Data Sources.....	14
4.2	Data Quality and Quality Control Plan .....	14
5	User Access .....	17
5.1	Login Procedures.....	17
5.1.1	Account Recovery.....	17
5.2	Queries .....	17
5.2.1	Site Specific Query.....	17
5.2.2	Data Source Approach.....	18
5.2.3	Mixed Graph Approach .....	18

5.2.4	Library Search .....	18
5.3	Data Export.....	18
5.3.1	Water Data Export.....	18
5.3.2	Library Metadata Export .....	18
5.3.3	Library Document Download .....	19
5.4	Data Import .....	19
5.4.1	Library Data Import .....	19
5.4.2	Water Data Import .....	19
5.5	Annual Reports and Monitoring Network.....	19
6	Administration.....	21
6.1	Security.....	21
6.1.1	Linux User Access Limitations .....	21
6.1.2	Database Access Limitations .....	21
6.1.3	Database Access User Levels.....	21
6.1.4	Software Database Protection .....	22
6.1.5	Mapserver Access Limitations.....	22
6.1.6	Web User Password Protection.....	22
6.2	Administration.....	22
6.2.1	Web user Access and Roles.....	22
6.2.2	Database Administration.....	22
6.2.3	Other Data Administration.....	23
6.2.4	Server Administration .....	23
7	Summary .....	25
8	References.....	27

## LIST OF FIGURES

---

Figure 1-1	Western Management Area Boundary.....	3
------------	---------------------------------------	---

## GLOSSARY OF TERMS/ABBREVIATIONS

Acronym/Abbreviation	Definition
<b>API</b>	Application Programming Interface
<b>APN</b>	Assessor Parcel Number
<b>ASCII</b>	American Standard Code for Information Interchange
<b>BibTeX</b>	<i>Software Name (Bibliography for TeX)</i>
<b>CA</b>	California
<b>CASGEM</b>	California Statewide Groundwater Elevation Monitoring
<b>CEDEN</b>	California Environmental Data Exchange Network
<b>CIMIS</b>	California Irrigation Management Information System
<b>CMA</b>	Central Management Area
<b>CSS</b>	Cascading Style Sheets
<b>CSV</b>	Comma-separated values
<b>DEHS</b>	Department of Environmental Health Services
<b>DMS</b>	Data Management System
<b>DOGGR</b>	Division of Oil, Gas, and Geothermal Resources
<b>DOI</b>	Digital Object Identifier
<b>DWR</b>	California Department of Water Resources
<b>elog</b>	electrical log
<b>EMA</b>	Eastern Management Area
<b>ESRI</b>	<i>Company name (formerly Environmental Systems Research Institute)</i>
<b>GIS</b>	Geographic Information System
<b>GSA</b>	Groundwater Sustainability Agency
<b>GSP</b>	Groundwater Sustainability Plan
<b>HTML 5</b>	Hypertext Markup Language, version 5
<b>HTTP</b>	Hypertext Transfer Protocol
<b>JPEG</b>	Joint Photographic Experts Group
<b>JSON</b>	JavaScript Object Notation

<b>LAMP</b>	Linux, Apache, MySQL, PHP (or Perl, or Python)
<b>LAS</b>	Log ASCII
<b>MariaDB</b>	<i>Software Name</i>
<b>MLA</b>	Modern Language Association; in the context of a citation style
<b>MySQL</b>	<i>Software Name</i>
<b>NWIS</b>	National Water Information System
<b>PDF</b>	Portable Document Format
<b>PHP</b>	PHP: Hypertext Preprocessor
<b>QGIS</b>	<i>Software Name (formerly Quantum GIS)</i>
<b>REST</b>	Representational state transfer
<b>RIS</b>	Research Information Systems
<b>SFTP</b>	SSH File Transfer Protocol
<b>SGMA</b>	Sustainable Groundwater Management Act
<b>SQL</b>	Structured Query Language
<b>SSH</b>	Secure Shell
<b>SWRCB</b>	California State Water Resourced Control Board
<b>SYRVGB</b>	Santa Ynez River Valley Groundwater Basin
<b>URL</b>	Uniform Resource Locator
<b>USBR</b>	United States Bureau of Reclamation
<b>USGS</b>	United States Geological Survey
<b>VPS</b>	Virtual Private Server
<b>WAT</b>	California Water Code
<b>WMA</b>	Western Management Area
<b>WMS</b>	Web Map Service
<b>XML</b>	Extensible Markup Language
<b>XMP</b>	Extensible Metadata Platform

---

## 1 INTRODUCTION

The State of California law requires that all groundwater basins that the State has identified as medium and high priority basins achieve sustainability by January 2042. To meet this goal, State law requires the creation and implementation of a Groundwater Sustainability Plan (GSP) or multiple plans submitted under a coordination agreement. Santa Ynez River Valley Groundwater Basin (SYRVGB) as defined by the State of California Department of Water Resources (DWR) Bulletin 118 is medium priority groundwater basin in California's Central Coast. This document is part of the development of the GSP for the Central Management Area (CMA) portion of the SYRVGB (Figure 1-1).

The Sustainable Groundwater Management Act (SGMA), signed into law on September 16, 2014, includes a structure and schedule to achieve sustainable groundwater management within 20 years. The SYRVGB was identified as a medium priority basin and must submit a SGMA compliant GSP, or GSPs under a coordination agreement, by January 31, 2022. The CMA GSA is responsible for preparing the GSP for its portion of the SYRVGB. Two other GSPs are being prepared for other portions of the SYRVGB including the Western Management Area (WMA) and Eastern Management Area (EMA), respectively. Stetson Engineers was hired by the Santa Ynez River Water Conservation District (District) on behalf of the WMA and CMA to develop GSPs for those areas.

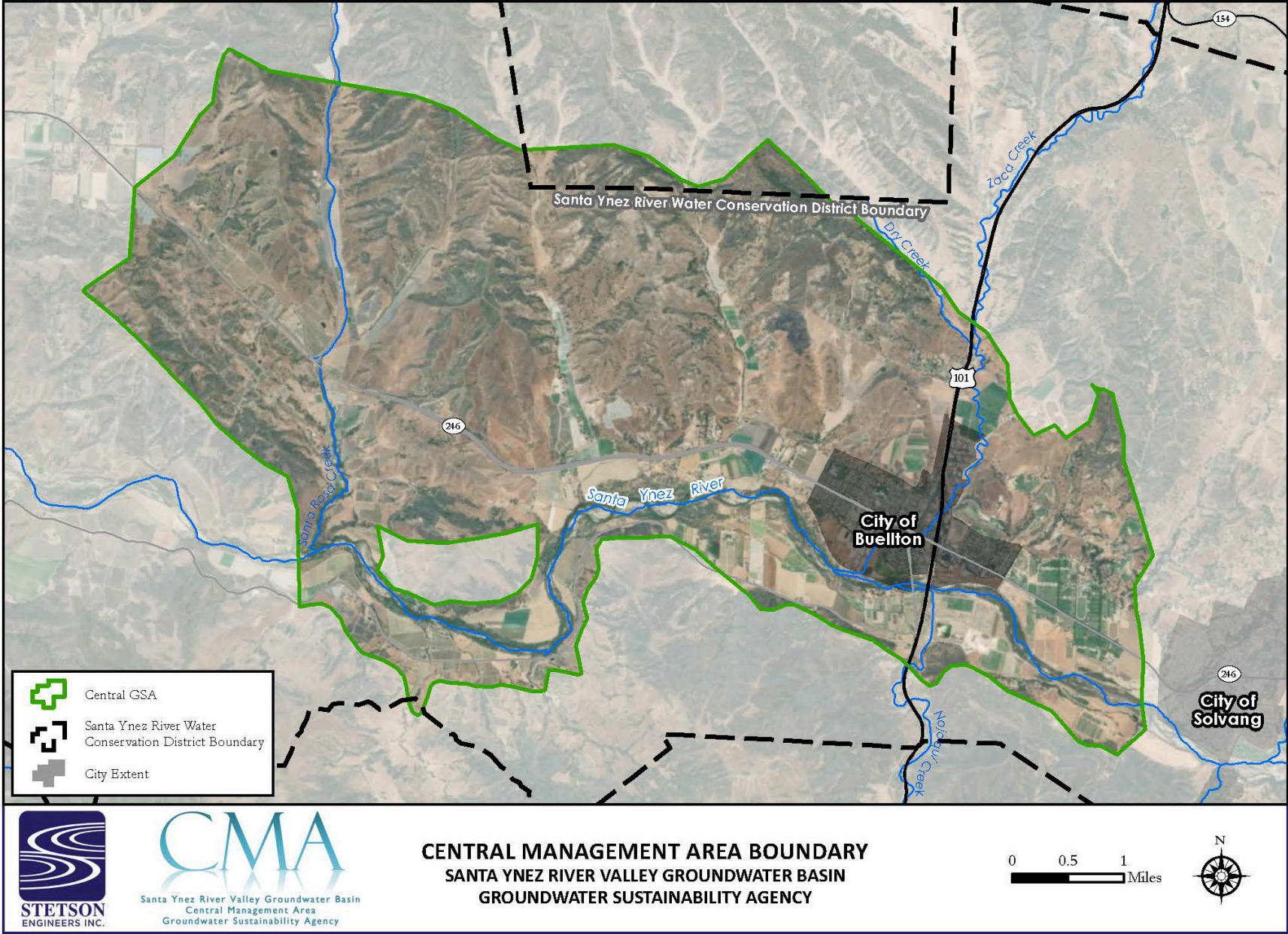
The CMA is in the Central portion of the Santa Ynez River Valley Groundwater Basin as shown in Figure 1-1. The CMA is comprised primarily of the Santa Ynez River alluvium and Buellton Uplands.

The SGMA requires **Data Management System (DMS)** preparation and maintenance as part of GSP preparation. This report describes the structure and content of the DMS being prepared for the CMA. Section 2 reviews the goals of the DMS, which include meeting the statutory requirements under SGMA as well as aiding in the development of the GSP. Section 3 describes the architecture of the DMS, including the technical computer software, hardware, and data storage components. Section 4 describes the data sources, including Federal, State, and local resources, that will be housed in the DMS. Section 5 describes user access features, including the procedures to login, query, and import and export data from and to the DMS. Section 6 identifies the security considerations in the DMS as well as the various administrative duties and roles in developing and maintaining the DMS.

INTENTIONALLY LEFT BLANK



Document Path: J:\m2710\CMA\_logo.mxd



INTENTIONALLY LEFT BLANK

---

## 2 GOALS OF DMS

The goals of the DMS are to improve data collection and storage and assist in the understanding and future reporting about groundwater conditions in the CMA. The SGMA GSP Regulations § 352.6. on Data Management Systems states:

Each Agency shall develop and maintain a data management system that is capable of storing and reporting information relevant to the development or implementation of the Plan and monitoring of the basin.

Source: CA WAT § 10733.2. Reference: CA WAT § 10727.2, 10728, 10728.2, and 10733.2.

Per these regulations, there are two main goals of the DMS including: 1) to support the development of the GSP, and 2) to provide a data framework for the continued monitoring of the CMA. The DMS will serve as the central repository of information for these efforts.

### 2.1 Support of Hydrogeologic Conceptual Model Development

One of the first uses of the DMS will be to support the development of the hydrogeologic conceptual model. The hydrogeologic conceptual model will describe the regional geologic structural setting and current conditions of the CMA groundwater basin as well as the components of water exchange throughout the hydrogeologic system.

The DMS will house information about the existing wells in the basin. For each of these wells, existing data will be populated within the tables of the DMS, including groundwater level data, well construction information, well logs, geophysical data, pumping test data, water quality data, and pumping data. In addition, the DMS will house data related to land subsidence, surface water flows, and total water use in the CMA.

Use of the DMS will allow for rapid determination of which parameters currently have data gaps and/or uncertainty, in order to prepare the Data Gaps Analysis and the plan to acquire any additional data that is needed to support sustainable groundwater management.

### 2.2 Monitoring Network

The DMS will be used to store and access the CMA data, including the CMA Monitoring Network data. Existing data will be reviewed for compliance with the criteria as outlined in the SGMA regulations for the CMA Monitoring Network and supplemented as needed. Data collected from the CMA Monitoring Network will follow the GSP Regulations Best Management Practices (California Department of Water Resources, DWR, 2016), specifically *Monitoring Protocols, Standards, and Sites*, and *Monitoring Networks and Identification of Data Gaps*. The output from the DMS will be constructed to easily input to the DWR GSP submittal tool, which includes the process for the annual monitoring report submittal.

INTENTIONALLY LEFT BLANK

---

## 3 SYSTEM ARCHITECTURE

The DMS system architecture is based upon the understood needs and requirements of the CMA. If during the development of the SGMA project additional or different needs are identified, aspects of the architecture may be adjusted to satisfy these needs.

The plan for the DMS is that a user's primary mode of interaction will be to open and interact with a web application through a modern web browser. Several user levels and roles have been established with different access privileges, and some roles have limited administrative capacity.

### 3.1 Platforms

The DMS platform for the CMA will be a web application built on the Linux Apache MySQL PHP (LAMP) web stack, which is a mature open source platform, scalable to the needs of the CMA. The LAMP web stack consists of the following set of software:

- L Linux operating system, currently the DMS is on a Fedora Linux distribution
- A Apache webserver
- M MySQL compatible database server, currently the DMS is on a MariaDB installation
- P PHP scripting

In addition to the database server, a mapserver is also being run on the system to provide access to certain kinds of complex geospatial data. A mapserver is an intermediary program that takes the source GIS data and provides it on demand in a format that client interface programs can use. Currently this mapserver is the QGIS server program. Additional user notification is provided through an email service, currently through the Postfix program.

End user interaction with the DMS is through a web application, which interfaces with the LAMP stack with a standard web front end, using JavaScript, CSS, and HTML 5 which requires the user to have a modern web browser.

### 3.2 Scripts

In addition to the components of the DMS which react to the user input or push telemetry, the DMS as a system includes programming that is run on a scheduled basis. Most of these scripts are written in Perl or Python. Scripting is for various automated items which include automated pull requests to telemetry, automated quality control, automated user notification, and general automated application maintenance.

### 3.3 Data Location

The DMS is planned to be contained and stored within a single server, which will include the data and scripting as described in the data types section, below. The DMS web application is designed to

function without any external dependencies, however some external third-party map data may be provided as links rather than mirrored within the system. Externally linked map data is not controlled by the DMS and may become unavailable, and requires the user have an internet connection to access.

The current location of the DMS is on a virtual private server (VPS) rented from a datacenter. VPS hosting is used as it is a fixed server with dedicated resources for a set price, unlike cloud hosting where resources are not fixed, and price is related to metered usage of resources. The VPS was selected due to the more predictable pricing. The current VPS provider for the CMA DMS is Host Winds.

Because the DMS data is contained within a single server, the DMS can easily be transferred to many other server configurations, maintaining flexibility for future requirements.

### 3.4 Data Types

There are a range of data types that are included as part of the DMS. To the extent possible data will be inserted in the MySQL-compatible database (database), however, there is additional information which is not easily included in the database.

#### 3.4.1 Database

The primary use of the database will be to host primarily indexed data which can contain the following types of data:

- Time-*invariant* location data for indexing and describing locations such as wells and surface sites such as stream gages.
- Time-variant data such as groundwater water levels, pumping data, or streamflow. Generally, this data consists of a location index, a measurement time, a measurement type identifier, a value, and a value qualifier.
- General information used in the interpretation of the above data types. This includes tables such as the USGS parameter code list, various set regulatory tables, etc. Each well will have corresponding database fields containing the well ID data, site information, construction details, and well screen information.
- Metadata about key documents regarding conditions within the basin. Metadata fields include publication data, author, alternative DOI or URL location, and geographic extents; not all documents will have all metadata fields.
- Data related to the access of the web application including web users and web user roles. This would include items such as the web user contact information, specific access-granted roles, and encrypted copies of web user passwords. Other data included here would be access logs which track usage of the web application, including information such as web user, IP addresses, login times, and browser details.

The DMS database plan is to exclude stakeholder information used for communication and outreach.

### 3.4.2 Technical Report Format

A second set of data, which are indexed by the database but not contained within, are digital copies of published and unpublished documents regarding conditions within the basin. These are saved in the standard portable document format (PDF). These will be provided and saved using unique identifiers, and the metadata will be stored in the database.

### 3.4.3 Well and Surface Site Data

Additional data types indexed by, but not contained, in the database include the following:

- Photographs of the wells and surface sites are expected to be stored outside of the database in JPEG format. Panoramic images, if they are included, are expected to have the included metadata (XMP format) set properly.
- Well completion reports, expected to be stored as portable document format (PDF). Most of the data interpreted from the well completion reports will be entered into the database tables. These may include as part of the report package pump test summaries, geophysical data including electrical logs (elogs).
- Interpreted well logs, expected to be stored as comma separated value (CSV) files.

### 3.4.4 Geospatial Data

In addition to the geospatial data that is being included in the database, there are other geospatial data sets that are included as part of the DMS. These include both vector and raster datasets and a summary of these geospatial data types are as follows:

- Geographic vector data sets that are relatively simple in terms of styling and small in terms of file size are generally saved in as GeoJSON format. This format is a structured version of the JSON (JavaScript Object Notation), a JavaScript data-interchange format, specifically for geospatial data. Additionally, the DMS may have programming (JavaScript) that adds interactivity based on the fields contained in the file.
- For large or complex vector data sets or raster data, the data is stored in the original format such as “ESRI” shapefile and made accessible through the mapserver following the Web Map Service (WMS) protocol. When data is requested by the user the mapserver renders the GIS format data into image tiles which are then sent to the user.
- For some large or complex data sets data may be pre-rendered and stored as a series of image tiles.

The selection of the method of storing and transmitting a geospatial data set depends on the details of the data set and needed output, as well as on constraints, such as available computing resources.

In addition to the key geospatial data which are hosted on the DMS server, the DMS may link to external geospatial data hosted by third parties. Currently this linked external third-party geospatial data is primarily from Federal and State of California servers, and includes various aerial imagery, supplemental topographic data, and geological maps with copyright restrictions. Third party data by nature is not controlled or managed by the DMS, so availability may be subject to change.

During coordination efforts with the EMA team, the EMA is not expected to provide a linked geospatial service. Additionally, to protect confidentiality of data, access to the mapserver and other data requires an active login to the DMS website and would not be available to third parties.

### **3.5 Backup**

Whole DMS backups consist of a transfer of a copy of the entire DMS to a dedicated physical hard drive, which is generally unplugged and located at a different physical location from DMS sever location, currently on a shelf in the consultants' office. These backups are conducted on a periodic basis, currently once a quarter. The process currently has several manual steps in downloading and transferring copies of the files. Use of an automated cloud storage backup strategy may also be adopted, which would have some additional storage hosting costs, but would be run on more frequent basis of monthly or weekly.

In addition to the whole DMS backup, portions of the code base common to other DMS projects are entered in to one of several distributed version control repositories in order to track changes and quickly roll out patches and improvements. The centralized location of these files (the repositories) are currently on GitHub, a subsidiary of Microsoft. These repositories are utilized whenever changes are made to the common code base.

### **3.6 Intra-basin Consistency / Formats**

The Central Management Area (CMA) is one of three management areas of the Santa Ynez River Valley Groundwater Basin; the others are the Western Management Area (WMA) and Eastern Management Area (EMA). The WMA will be using the same DMS architecture as the CMA and will be entirely compatible with the CMA. The EMA is being organized through a different consultant utilizing a separate and different system. Currently the plan is later to develop a common protocol to share data.

The primary method of consistency between the three management areas of the Santa Ynez River Valley Groundwater Basin will be to use a common data set generated by third parties. The CMA and WMA team has provided the EMA team with source datasets from the United State Geological Service (USGS), County of Santa Barbara, and United States Bureau of Reclamation (USBR).

Where there are unique data sets generated in one or more of the management areas, the EMA consultant has agreed to work together to ensure that any unique data can be shared across the basin. If there is data to be shared on a regular basis, the EMA consultant and the CMA and WMA



consultant team have agreed to work together to develop a common protocol for sharing data (such as an XML, JSON, or structured Excel file) through which all three management areas can communicate through.

INTENTIONALLY LEFT BLANK

---

## 4 DATA SOURCES AND QUALITY

The existing historical and current water resources monitoring and management programs within the CMA will be utilized and incorporated into the DMS as described below, including Federal, State, and Local programs.

### 4.1 Data Sources

#### 4.1.1 Federal Data Sources

A key Federal source of data will be the United States Geological Survey (USGS) which includes historical groundwater elevations and surface water flows. Data are stored electronically in the National Water Information System (NWIS) files and are retrievable from the USGS Water Resources Internet site. This data is reviewed by the USGS and available through well formatted interfaces, called REST APIs, which provide data in a structured XML format upon request. Included is location information, necessary measurement information in addition to the measurement result, a description of the measurement being conducted, and the units of measurement. In addition, the CMA has numerous USGS hydrogeological studies, whose data will be incorporated in the CMA DMS.

Another Federal source of data is the United States Bureau of Reclamation (USBR). The USBR holds the water rights permits for the Cachuma Reservoir, located about 35 miles upstream of Lompoc. As part of the conditions of this permit, USBR collects monthly groundwater level data along the Santa Ynez River alluvium and within the Lompoc Plain.

#### 4.1.2 State Data Sources

State of California (State) sources of data include the DWR's California Statewide Groundwater Elevation Monitoring (CASGEM) Program. DWR works cooperatively with local agencies (County of Santa Barbara), referred to as CASGEM "Monitoring Entities," to collect and maintain groundwater elevation data in a manner that is readily and widely available to the public through the CASGEM online reporting system.

The State will be a source for well drilling information. DWR has compiled well completion reports for successful and unsuccessful groundwater wells and has made these available online with redacted personal information (per CA WAT § 13752). These well completion reports describe aspects of the installed well, and generally include driller well logs that describe the nature of the formations encountered while drilling. Because there are thousands of wells in the SYRVGB, data from wells determined to be "key wells" will be included in the DMS. These "key wells" are the wells which are most useful for assessing the basin. A "key well" has a known and accurate well location (geographically and vertically), depth of the well, availability and completeness of the lithological log, availability of geophysical logs, and proximity to other wells or key features.

Additionally, information about petroleum and gas wells will be retrieved from the California Division of Oil, Gas, and Geothermal Resources (DOGGR) database. This data set includes geophysical well

logs, generally in an analog PDF format rather than the digital Log ASCII (LAS) format. Key wells in this dataset will be identified and information such as geological horizons and other pertinent geologic data will be entered in the DMS.

These databases will be reviewed and well sites with useful information will be incorporated into the CMA DMS. The SWRCB water rights database will also be queried for information to import into the CMA DMS such as location information.

For water quality, two additional State databases will be utilized for the CMA DMS including the State Water Resources Control Board's (SWRCB) Groundwater Ambient Monitoring and Assessment Program (GAMA) database and California Environmental Data Exchange Network (CEDEN).

For climate data, the California Irrigation Management Information System (CIMIS) stations in the Santa Ynez River watershed will also be utilized in the CMA DMS. This data may also be used for the determination of water use in the basin.

#### 4.1.3 County Data Sources

The Santa Barbara County Water Agency currently conducts precipitation monitoring and, as of spring 2019, conducts annual groundwater level monitoring that was previously conducted by the USGS. Precipitation and groundwater data from the County will be included in the CMA DMS.

The Water Agency provided copies of their staff "field notebook" which documents the water level collection activities. The field documentation was originally developed by the USGS, and includes various digital images, some of which are photos of wells, scans of water level documents such as owner contact information, site sketches, and other notes. As appropriate, this data will be incorporated into the DMS.

The Santa Barbara County Department of Environmental Health Services (DEHS) has well records of wells that were drilled within the CMA. The data is organized by the Assessor's Parcel Number (APN). These records are in hard copy form and located at the DEHS Santa Maria office. Many of the records were digitized as part of the data collection effort and are under review for possible inclusion in the DMS. Confidential or personal information will be redacted.

#### 4.1.4 Municipal, Water District, and Other Data Sources

Data obtained from the CMA member agencies will be imported into the CMA DMS. This includes hydrogeologic data from the City of Buellton. In addition, available groundwater data from the Santa Ynez River Water Conservation District will be obtained and imported into the CMA DMS. Data that is confidential will not be included in the CMA DMS.

## 4.2 Data Quality and Quality Control Plan

The SGMA GSP Regulations § 354.44 (c) states that "Projects and management actions shall be supported by best available information and best available science." The above sources constitute

the “best available information” for the CMA that is consistent with scientific and engineering professional standards of practice.

Data will be evaluated for validity and acceptable use for the GSP preparation. Data compilation and review will identify potential data gaps or unacceptable levels of uncertainty, which may facilitate focused discussions with the CMA GSA. When different sources of data have different values for the same parameter (i.e. well location or land surface elevation), a source and comments data field (column) will be associated with the current value.

Initially, all data will be collected and imported into the CMA DMS, but then sites will be reviewed and screened in a three-tiered process for the purposes of evaluating the data for potential inclusion in the CMA DMS and sharing with the other two Management Areas:

Tier 1: Data Meets All Criteria for Inclusion in the GSP

Tier 1 Data will be used in the future monitoring program for the CMA GSP. These data meet all the compliance criteria outlined in the SGMA regulations for inclusion in a SGMA Monitoring Network (i.e. SGMA GSP Regulations § 352.4). Measurable objectives and minimum thresholds will be established as part of the evaluation of the SGMA sustainability indicators. Data will be field validated for inclusion in the CMA Monitoring Network.

Data evaluated as part of the Tier 1 review will be included in the DMS. This tier of data will be shared with the other Management Areas as part of an intra-management area coordination agreement.

Tier 2: Meets Partial Criteria - May or may not be included in the GSP

Data that do not meet all the criteria for inclusion in the CMA GSP monitoring network may still be useful in developing the hydrogeologic conceptual model. For example, if a well has a significant amount of historical water level data but lacks well casing or total depth information, or conversely, if a well has a lithologic well log available but no historical water level data, these wells can still be used to develop the hydrogeologic conceptual model. Because SGMA GSP Regulations § 352.4 (c) (3) states, “Well information used to develop the basin setting shall be maintained in the Agency’s data management system”, these additional wells are an important part of the CMA DMS.

Wells and surface sites that are identified in this tier of the process will be included in the DMS, but professional judgment will be used as to the relevance and usefulness of these data for the GSP. This tier of data may or may not be included as part of intra-management area coordination with the WMA and EMA.

Tier 3: Minimum Criteria - Not for Inclusion in the GSP

Data that do not meet the criteria for the CMA GSP (Tier 1) or have no useful information (Tier 2) will be included in Tier 3. As a default, this tier of data will be “turned off” or not visible in the DMS but

will be held in the DMS in case additional information is obtained in the future that would change the tier classification of the data.

Wells and surface sites that are identified as not containing useful data for monitoring or modeling may also be included in the DMS to track and exclude from other layers. A low amount of effort will be employed on these sites, and wells as part of this layer will generally be excluded from intra-management area coordination with the WMA and EMA.

---

## 5 USER ACCESS

DMS user access will be primarily through a web application; users will be assigned specific roles and given specific permissions to access the DMS. The web interface will require the user to access the DMS through a modern web browser, older browsers may work with less functionality or not at all.

### 5.1 Login Procedures

Access to the DMS will be controlled through a username and password login system with a username having a specific defined role on the website; each role has specific defined privileges to access data or conduct limited administrative actions.

In most cases the user's registered email can be used in lieu of the username. User information will be set to automatically populate the username and login information by default.

In order to protect web user passwords if the DMS is ever compromised, web user passwords will be stored as encrypted hashes.

#### 5.1.1 Account Recovery

The DMS includes automated retrieval of account access if username and/or passwords are forgotten. The application will email the web user to their provided email address on file, sending a recovery link which will allow the user to reset their password and regain access the DMS. This feature is only available if the DMS is on a server that is connected to the internet and requires the web user to maintain control of their email account.

### 5.2 Queries

As described in the System Architecture section, access through to the underlying MySQL-compatible database is mediated through the PHP programming.

DMS data in the database is generally accessed through two approaches: a well/site-specific approach, or a data source approach.

#### 5.2.1 Site Specific Query

The site-specific approach has the user identify which data, well or surface site is of interest. The location of interest is selected by the user either through a map interface, or through pages with a search and list features. Data is then provided about that data, well or surface site.

Well or site information may include well properties, image of the well or well log, geophysical logs, and time-series data such as production, water level elevation, or depth to water, pulled from various databases.

This site-specific approach allows for additional insights to be provided to the user, such as the land surface at the site, well perforations, and relationship between water level depth and water level elevation at that well.

### 5.2.2 Data Source Approach

This approach has the user navigate to a page for each specific source of data. This includes groundwater level data and water quality, geophysical, well construction, surface water and other data. The page consists of map showing the location the data was collected, as well as a listing of sites and the available data at each site. The user can select data either through the map, or through the list, and easily compare several sites for the given source. Using the interface, the user can compare one or more data sets to established thresholds, limits or other criteria established by the GSA, State, or Federal agency.

The data source approaches will be developed as various data sets are incorporated into the DMS.

### 5.2.3 Mixed Graph Approach

This graphing feature allows pulling together two or more data sets that are not necessarily related either by location or source of data. An example of this would be stream gage and depth to groundwater data; surface water and groundwater data; and water quality data. The DMS will include a search feature for identifying what data sets are available.

### 5.2.4 Library Search

The library currently provides several ways to search the metadata including by title, year, and keywords.

## 5.3 Data Export

### 5.3.1 Water Data Export

All available graphs currently have a data export feature that exports the data queried in the graph to a Microsoft Excel file, in addition to providing download options into various image formats.

### 5.3.2 Library Metadata Export

The library functions include export features to a set of selected citation manager formats, including RIS, Microsoft Word XML, and BibTeX.

Citation management software are used in track works cited or used in the document and formatting to match specific bibliography and citation styles. Using citation management software is a best practice when writing for a publication, as various publications generally specify a bibliography and citation style such as the Chicago or MLA.



### 5.3.3 Library Document Download

Individual PDFs can be downloaded directly through the web interface. The DMS supports full pause and resume download functionality. The pause and resume feature on the server require the client software supports the HTTP range request, a feature that is available in all modern web browsers.<sup>1</sup>

## 5.4 Data Import

### 5.4.1 Library Data Import

Certain user roles allow the user to edit the library metadata, as well as to upload, and delete PDFs through the web user interface.

### 5.4.2 Water Data Import

Currently, the ability to import water data is limited to the DMS administrators. Some data sets which originate from third parties with published APIs, such as the USGS data through NWIS, may be automatically fetched and updated on a scheduled basis.

Additionally, if telemetry is deployed, the DMS may be configured to accept specific push requests, and DMS scripting can be configured for automated pull requests.

## 5.5 Annual Reports and Monitoring Network

SGMA § 10728 on Annual Reporting by the GSA to DWR states:

“On the April 1 following the adoption of a groundwater sustainability plan and annually thereafter, a groundwater sustainability agency shall submit a report to the department containing the following information about the basin managed in the groundwater sustainability plan:

- (a) Groundwater elevation data.
- (b) Annual aggregated data identifying groundwater extraction for the preceding water year.
- (c) Surface water supply used for or available for use for groundwater recharge or in-lieu use.
- (d) Total water use.
- (e) Change in groundwater storage.”

---

<sup>1</sup> Browser support for the HTTP Range request: [https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Range#Browser\\_compatibility](https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Range#Browser_compatibility), accessed 2019-08-15.

The CMA DMS will be the primary tool used to compile this data each year for the annual report. The CMA DMS will contain the sites used in the CMA Monitoring Network. The DMS is planned to automate the generation of the tables and figures for the annual report. The output from the DMS will be constructed to easily input to the DWR GSP submittal tool, which includes the process for the annual monitoring report submittal.

---

## 6 ADMINISTRATION

The DMS platform is built on top of a LAMP web stack, and depending on the administrative task, administration requires interaction with different portions of the stack, and different access levels. As mentioned, some web user roles have limited ability to access or administer parts of the DMS platform.

### 6.1 Security

Security covers several topics and concerns related to malicious actors such as: damaging the system, intercepting information, denying access to the system, falsely spoofing the system, or using the system to cause damage to others. Since the DMS is expected to be connected to the internet, there are expected to be constant low-level attacks. A general defense-in-depth strategy has been employed.

#### 6.1.1 Linux User Access Limitations

Currently the DMS firewall allows only specific whitelisted IP addresses to connect to the SFTP/SSH port (22). This denies potential access to the thousands of daily attempted unauthorized logins.

General SFTP (port 22) user access can be granted in order to provide a secure way to share data. SFTP access spaces are in a chroot jail<sup>2</sup> and are denied shell access in order to limit the amount of potential damage from inappropriately disclosed usernames and passwords.

#### 6.1.2 Database Access Limitations

Access to the MySQL-compatible database is not directly available to remote users through the standard MySQL port (3306) for direct login. All access must be from the local server, such as from PHP, or from a Linux user logged-in through an encrypted connection. This limits the potential for unencrypted data to be intercepted.

This is enforced at several levels: at the MySQL database level all users are required to login from the localhost, additionally the server firewall blocks all incoming connections to that port.

#### 6.1.3 Database Access User Levels

Access to the database is limited through a series of database users, each with specified user privileges allowing for certain actions on the database, and unique passwords. The DMS web interface communicates with the database using the least privileged user level in each instance.

---

<sup>2</sup> Term meaning setting an apparent root directory. Users and processes cannot identify files outside the root directory, and so this has the effect of disallowing access to any files outside of the specified directory tree, separately from any file permissions-based restrictions.

#### 6.1.4 Software Database Protection

The DMS web interface has been programmed with an understanding of the potential for SQL injections. Strategies employed to limit this attack vector include input sanitation and parameterized SQL queries, as well as using connections with limited privileges.

#### 6.1.5 Mapserver Access Limitations

The mapserver runs through the Apache web server, but access to the mapserver is intentionally limited so all access is from the local server, such as the PHP program. The primary reason for this limitation is to limit potential denial of service attacks against the DMS server, as the mapserver can be resource intensive.

In addition, the PHP program provides additional checks that the web user has logged in, as well as additional cache support enabling an overall faster experience for the web user.

#### 6.1.6 Web User Password Protection

Strong passwords are encouraged for all web users of the DMS by providing examples of strong passwords, and by providing calculation of the information density of the proposed password. User passwords are partially protected by a several second time out when incorrect passwords are entered, in order to limit the rate at which web passwords can be tested by a potential attacker.

As described in Section 5.1.1, an automated account recovery is provided. This automated recovery emails a recovery URL to the email account on file. This recovery URL is a random, time-sensitive, unique URL. This method of account recovery relies on the user securing and maintaining control of their associated email account.

## 6.2 Administration

### 6.2.1 Web user Access and Roles

Web user roles and access privileges can generally be modified through the web interface, if the web user has been granted 'administrator' role privileges. Otherwise user privileges can be directly modified through modifying the database.

### 6.2.2 Database Administration

Administration of the database is to be primarily conducted by the CMA consultant team.

Currently database administration requires access to Linux user to connect to the server, as well as username and password for the database user with the required privileges for the administration task. A web user role with database administration through the web interface may be developed if needed.

### 6.2.3 Other Data Administration

Administration of the non-database data, such as well and surface site images, or GIS data, will be performed by the CMA consultant team. This requires modifying files in specific locations or modifying configuration files in the case of the mapserver.

These modifications require access to the primary Linux user. A web user role with a file manager administration through the web interface may be developed if needed.

### 6.2.4 Server Administration

Administration of the server (root access) will be performed by the CMA consultant team. Server administration requires the appropriate password and connection from a whitelisted IP address.

INTENTIONALLY LEFT BLANK

## 7 SUMMARY

This Data Management Report describes both the content and structure that is planned for the Data Management System (DMS) that will meet the statutory requirements under SGMA. Data for the CMA will now be collected, reviewed, stored, and made available as described in this document; however, this plan will be amended as needed based on ongoing needs of the CMA in developing the Groundwater Sustainability Plan.

The plan includes a discussion of the general architecture of the DMS, including aspects of the software that is planned to be used, and how various types of data are planned to be incorporated into the system. As described the DMS uses open source software for most of the architecture components. The plan identifies how all data types, including GIS data and reports, will be handled in the DMS.

The plan discusses the expected sources of relevant data, including federal, state, county, local, and municipal, and how they will be collected for inclusion into the DMS. There is an identification of a tiered scheme for data collection and verification efforts, in order to focus efforts on higher impact data.

The plan also includes a general description of the web interface works and provides access to the data stored within the system, as well as outlining a process for exporting and importing various data sets into the system.

Finally, more details are provided with regards to various administration concerns, such a general description of several of the security steps taken to protect the system, as well as various ways in which administration of the system is planned.

The next step in the DMS process will be the continued population of the various data sets as outlined in this plan for the data compilation effort.

As the data compilation effort and population nears a completion, a Technical Memorandum will be produced to describe the data compilation effort as completed, including the data collected and sources. The technical memorandum will also provide updates and significant changes to the functions of the web based DMS.

INTENTIONALLY LEFT BLANK



---

## 8 REFERENCES

California Code of Regulations-Title 23 Waters. 2014. Groundwater Sustainability Plan Regulations (“GSP Emergency Regulations”).

Department of Water Resources (DWR). 2016. Bulletin 118. Interim Update 2016. California’s Groundwater: Working Towards Sustainability. December 2016.

Department of Water Resources (DWR). 2016. Monitoring Networks and Identification of Data Gaps. Best Management Practices for the Sustainable Management of Groundwater. December 2016.

Department of Water Resources (DWR). 2016. Monitoring Protocols, Standards, and Sites. Best Management Practices for the Sustainable Management of Groundwater. December 2016.

Sustainable Groundwater Management Act (SGMA) and related Provisions. 2014.

INTENTIONALLY LEFT BLANK