

# Demand Dataset Procedure

Payman Alemi and Aakash Prashar

Last Updated: 5/7/2025

## Table of Contents

1. Purpose .....	2
2. Getting Started .....	3
3. Flat File Download .....	4
3.1 Description .....	4
3.2 Procedure .....	4
3.3 Output .....	4
4. GIS Pre-Processing .....	5
4.1 Description .....	5
4.2 Procedure .....	5
4.3 Output .....	10
5. GIS Semi-Automated Manual Review .....	10
5.1 Description .....	10
5.2 Procedure .....	11
6. GIS Full Manual Review .....	17
7. Preliminary Demand Dataset .....	18
7.1 Description .....	18
7.2 Procedure .....	18
7.3 Output .....	18
8. Demand Dataset QA/QC Flags .....	19
8.1 Description .....	19
8.2 Procedure .....	19
8.3 Output .....	20
9. Manual Review: Duplicate Reporting Errors .....	20
9.1 Description .....	20
9.2 Review Example: A024524 and A024526 (2020 Direct) .....	24
9.3 Review Example: D030704 (2017 Direct and Storage) .....	24

9.4 Review Example: A024943, A025060, S008688, S019745, S019747, S019754, and S019775 (2017–2023 Storage) .....	24
9.5 After Completion of the Manual Review .....	25
10. Manual Review: Unit Conversion Errors .....	26
10.1 Description .....	26
10.2 After Completion of the Manual Review .....	26
11. Manual Review: Empty Reports.....	27
11.1 Description .....	27
11.2 After Completion of the Manual Review .....	28
12. Updated Demand Dataset.....	28
13. Sub-Basin Delineation .....	28
13.1 Description .....	28
13.2 Procedure .....	29
13.3 Output.....	30
14. Manual Review: Sub-Basin Delineation .....	30
14.1 Description .....	30
15. Master Demand Table .....	31
15.1 Description .....	31
15.2 Procedure .....	31
15.3 Output.....	31
Appendix A: Compiled List of Requirements.....	32
Appendix B: Troubleshooting Errors While Running Scripts.....	32
Appendix C: Water Years and Calendar Years .....	33

## 1. Purpose

This procedure involves downloading and cleaning data from the Electronic Water Rights Information Management System (eWRIMS) for use in modeling a watershed with DWRAT (Drought Water Right Allocation Tool). Given water supply and demand data, DWRAT optimizes allocations, handling first riparian water right claims and then distributing the remaining supply among appropriative water rights (in order of priority).

For the demand input data, DWRAT requires files that list all appropriative and riparian water rights located within the watershed. These files should include information on the

rights' priority dates, average monthly water usage (in acre-feet), and sub-basin location (only one sub-basin per water right). DWRAT also requires a file that expresses the connectivity between different sub-basins within the watershed.

## 2. Getting Started

Later steps will have additional requirements, but these are the minimum to get started. Make sure you have the following on your computer:

- R and R Studio
- Git for Windows
- R Packages (preferably the latest versions)
  - tidyverse
  - sf
  - openxlsx
  - mapview
  - lwgeom
  - httr
  - data.table
  - odbc
  - DBI
  - readxl
  - janitor
  - writexl
- **DWRAT\_DataScraping GitHub Repository** (with the latest version of the scripts on your branch)—this means performing a pull request on the remote repository from the main branch into your branch and then performing a local git pull.
- **OneDrive** and synced **SDA SharePoint** repository
  - The SDA SharePoint repository can be locally accessed via OneDrive after clicking on the “Sync” button located [on the webpage](#)
  - OneDrive must be running to ensure that the files are up-to-date
- CalEPA VPN Connection
  - Required only for downloading the flat files
- At least 30 GB of free space on your hard drive (or SSD)

After that, navigate to the “DWRAT\_DataScraping” folder on your computer, open the “Demand” sub-folder, and open both “Demand.RProj” and the script “Demand\_Master\_Script.R”.

## 3. Flat File Download

### 3.1 Description

This script downloads several eWRIMS flat files via SQL queries. It also prepares a file for the GIS steps. As of 9/6/2024, the flat files are hosted on the ReportDB database under the ReportManager, 1542 server.

### 3.2 Procedure

1. While connected to the GlobalProtect VPN, run the script “GIS\_POD\_Flat\_File\_Prep.R” if you have access to the ReportDB database.
  - a. This can be accomplished by running Line 32 of “Demand\_Master\_Script.R”.
2. If you lack read permissions to the ReportDB database, then you will have to run “GIS\_POD\_Flat\_File\_Prep\_Old\_Method.R” instead.
  - a. Running “GIS\_POD\_Flat\_File\_Prep.R” without the necessary permissions will result in the error message shown in Figure 1.

```
! ODBC failed with error 42000 from [Microsoft][ODBC SQL Server
Driver][SQL Server].
✗ The SELECT permission was denied on the object
'ewrims_flat_file_pod', database
'ReportDB', schema 'FLAT_FILE'.
• <SQL> 'Select * from ReportDB.FLAT_FILE.ewrims_flat_file_pod'
i From nanodbc/nanodbc.cpp:1722.
Run `rlang::last_trace()` to see where the error occurred.
```

Figure 1. Error message in R for insufficient access to the ReportManager SQL server.

- b. “GIS\_POD\_Flat\_File\_Prep\_Old\_Method.R” instead downloads the flat files from the [Division of Water Rights internal flat file webpage](#), which should be accessible to all Division of Water Rights staff. However, since this script downloads the entire files (without any filters) from the Internet instead of a SQL server, it runs far more slowly. Consequently, it may take more than an hour to download all necessary flat files—make sure you have at *least 30 GB of storage available on your computer as well*.

### 3.3 Output

Both the *GIS\_POD\_Flat\_File\_Prep* and *GIS\_POD\_Flat\_File\_Prep\_Old\_Method* R scripts will generate several CSVs and save them in the “RawData” folder:

- “ewrims\_flat\_file\_pod.csv”

- “ewrims\_flat\_file.csv”
- “water\_use\_report.csv”
- “water\_use\_report\_extended.csv”
- “ewrims\_flat\_file\_use\_season.csv”
- “ewrims\_flat\_file\_party.csv”

Also, in the “IntermediateData” folder, the scripts will add “Flat\_File\_eWRIMS\_[DATE].csv” (where “[DATE]” is yesterday’s date in YYYY-MM-DD format). This will be the input file for the GIS steps.

## 4. GIS Pre-Processing

### 4.1 Description

The next step is to identify which water rights operate within the watershed. Rights with at least one point of diversion (POD) in or around the watershed boundaries will be selected. However, PODs may be misplotted. These steps involve researching flagged PODs to check for errors in the dataset. The result will be a finalized list of PODs that divert water **within** the watershed.

### 4.2 Procedure

1. Procure a GIS layer that contains the boundaries of the watershed of interest as a single polygon.
2. Save this layer on the Supply and Demand Assessment SharePoint (SharePoint) and update “Watershed\_Demand\_Dataset\_Paths.xlsx” with information about the watershed, including a path to that GIS file.
  - a. “Watershed\_Demand\_Dataset\_Paths.xlsx” is located in the “4. Demand Data Tracking” folder on SharePoint.
    - i. Values need to be specified in the “Main\_Sheet” worksheet. The “Dictionary” worksheet contains definitions and explanations for every column in “Main\_Sheet”.
  - b. The columns to update with watershed information are:
    - i. “INDEX”
      1. This number will be important in later steps, so be sure to remember it.

- ii. “NAME”
  - iii. “ID”
- c. The three columns that should be updated with information about the GIS layer are:
  - i. “WATERSHED\_BOUNDARY\_DATABASE\_PATH”
    - 1. This should typically be the path to the GIS container (such as a geodatabase or geopackage file). If, instead, a layer file is used, specify that name here and leave the “WATERSHED\_BOUNDARY\_LAYER\_NAME” column blank.
    - 2. If a local file is used, specify the path to the file, either as an **absolute path** (e.g., “C:\Users\aprashar\Documents\ArcGIS\Example.gdb”), or a **relative path** from the “Demand” GitHub sub-repository folder.
    - 3. If a SharePoint file is used, specify the relative path to the file from the SharePoint home directory (e.g., “Watershed Folders\Navarro River\Data\GIS Datasets\Example.gpkg”).
    - 4. Both backslashes and forward slashes are acceptable in these file paths (they can even be mixed together).
  - ii. “WATERSHED\_BOUNDARY\_LAYER\_NAME”
    - 1. If the watershed boundary layer is contained within a GIS container file, the layer name should be specified in this column.
  - iii. “IS\_SHAREPOINT\_PATH\_WATERSHED\_BOUNDARY”
    - 1. If the GIS file is located on SharePoint, this should be “TRUE”. Otherwise, it should be “FALSE”.
    - 2. See the “Dictionary” worksheet in “Watershed\_Demand\_Dataset\_Paths.xlsx” for additional descriptions of these three columns.
- 3. For this watershed, ensure that a name is specified in each of the three search string columns in “Watershed\_Demand\_Dataset\_Paths.xlsx”
- 4.

- a. These columns are:
    - i. "WATERSHED\_COLUMN\_SEARCH\_STRING"
    - ii. "SOURCE\_NAME\_COLUMN\_SEARCH\_STRING"
    - iii. "TRIB\_DESC\_COLUMN\_SEARCH\_STRING"
  - b. The "WATERSHED", "SOURCE\_NAME", and "TRIB\_DESC" columns in the eWRIMS flat file will be searched for matches using these name strings. Therefore, the specified values should be unique to the watershed (but not too specific).
    - i. Technically, the R programming language treats these search strings as regular expressions, so be wary of special characters like hyphens or question marks.
  - c. If you do **not** want to use this search method for flagging PODs, leave these three fields blank.
5. Make sure your local SharePoint files are synced with their remote versions.
- a. You must sync the parent folder ("Documents") in order for the *makeSharePointPath()* function to work!
  - b. It helps to open the file locally on your computer to force OneDrive to sync the changes.
  - c. These recent edits to "Watershed\_Demand\_Dataset\_Paths.xlsx" should appear on your local version of the file.
  - d. Ensure that the containing folder is synced locally by right clicking and selecting "Always keep on this device".

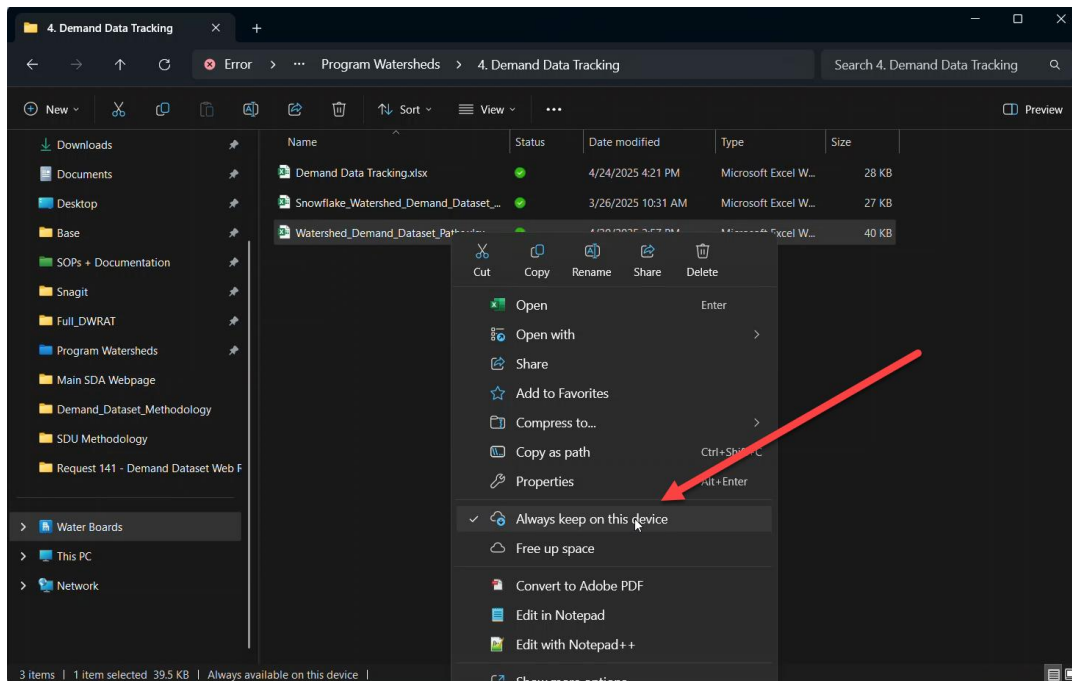


Figure 2. Select "Always Keep On This Device."

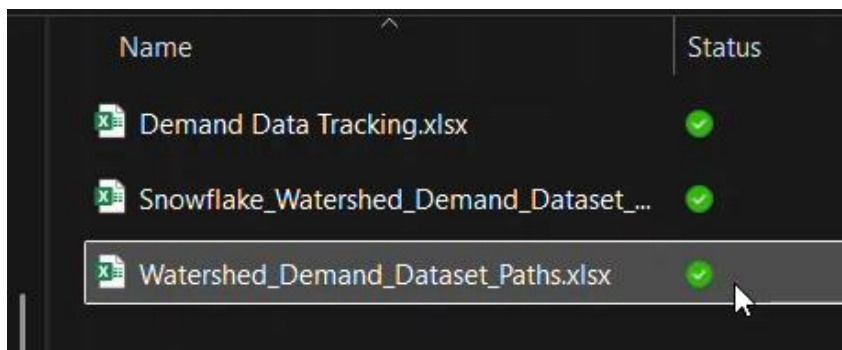


Figure 3. The green checkmark indicates that the file is simultaneously saved to the cloud and fully synced locally on the device.

6. Open "Watershed\_Selection.R" by navigating to the "Scripts" folder from within *Demand.Rproj*. On Line 23 of "Watershed\_Selection.R", where the ws dataframe is defined, change the number to the row index value that corresponds to this watershed.
  - a. The value should be the same as the number in the "INDEX" column of "Watershed\_Demand\_Dataset\_Paths.xlsx" (see Figure 4). **Aside from this number, please do not change anything else in the script.**

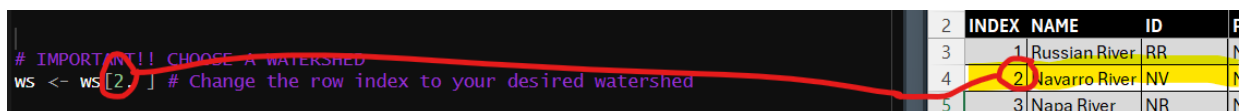


Figure 4. The ws dataframe in "Watershed\_Selection.R".



- b. Be sure to close “Watershed\_Demand\_Dataset\_Paths.xlsx” when you are done with this step. If the spreadsheet is still open when later scripts are executed, you will receive the error shown in Figure 5.

```
Error in utils::unzip(zip_path, list = TRUE) :  
  zip file 'C:\Users\aprashar\Water Boards\Supply and Demand Assessment -  
Documents\Watershed Folders\Watershed_Demand_Dataset_Paths.xlsx' cannot be  
opened
```

Figure 5. Error message when “Watershed\_Demand\_Datasets\_Paths.xlsx” is open.

7. Open “Dataset\_Year\_Range.R”, and on Lines 8 and 9, ensure that the years listed in the vector are appropriate (see Figure 6). **Aside from these two numbers, please do not change anything else in the script.**
  - a. The two numbers should be the starting year and ending year for the demand dataset.
  - b. The dataset bounds affect the number of rights included in the dataset as well as the number of reports used for each right when computing their average monthly diversions.
  - c. These years correspond to the reporting years for RMS submissions available on eWRIMS. Some of these reports cover calendar years, while others represent water years. See **Appendix C: Water Years and Calendar Years** on page 33 for more details.

```
7 # Change the numbers here!!  
8 yearRange <- c(Start = 2017, # The start year  
9               End = 2023) # The end year  
10
```

Figure 6. In “Dataset\_Year\_Range.R”, the numbers written in the “yearRange” vector on Lines 8-9 define the boundaries of the demand dataset.

8. Close the “Watershed\_Demand\_Dataset\_Paths” spreadsheet if it is still open and return to “Demand\_Master\_Script.R”.
9. Run the first part of “Demand\_Master\_Script.R” (Lines 2 to 28).
  - a. Make sure that the output statements in the console about the watershed and dataset bounds are as expected (see Figure 7 for an example).

```

>
> # IMPORTANT!!
> # Update "watershed_Selection.R" to select a watershed
> source("Scripts/watershed_Selection.R") # DATA ACQUISITION SCRIPT
Running script for Test Navarro
>
> # IMPORTANT!! x2
> # Specify the years to be included in the demand dataset
> source("Scripts/Dataset_Year_Range.R") #DATA ACQUISITION SCRIPT
The demand dataset will use reports submitted for 2017 through 2023
>

```

Figure 7. The R console output indicates that the demand scripts will be run for the Test Navarro watershed using the 2017-2023 reporting timeframe.

10. If you intend to perform the **semi-automated** GIS review procedure, run the script “GIS\_Preprocessing.R” by executing Line 37 of “Demand\_Master\_Script.R”.

- a. If you are conducting the GIS review process **entirely manually**, do NOT run this script and skip to [Section 6](#).

## 4.3 Output

The script’s output will be a spreadsheet titled “[ID]\_GIS\_Preprocessing.xlsx” (where “ID” is a unique identifier for the watershed, matching the value in the “ID” column of “Watershed\_Demand\_Dataset\_Paths.xlsx”). It will be exported to the “OutputData” folder of the DWRAT\_DataScraping/Demand sub-repository. This spreadsheet will be necessary for the semi-automated review process described in the next section.

The script also produces a Geopackage file in the “OutputData” folder called “[ID]\_PODs\_of\_Interest.gpkg”. This file contains spatial data for all of the PODs that were flagged by “GIS\_Preprocessing.R”. It can be used to plot the PODs in GIS applications like ArcGIS Pro.

# 5. GIS Semi-Automated Manual Review

## 5.1 Description

A manual review of water rights’ documentation is required at this stage. There is a script that can assist with this process, but some manual work is unavoidable. The end goal is a list of water rights that divert water in the watershed.

“GIS\_Preprocessing.R” performed several different subroutines to identify PODs that may divert from the watershed. In “[ID]\_GIS\_Preprocessing.xlsx”, the worksheet titled “POD\_Selection\_Info” summarizes which selection method captured each water right.

With this information, you can choose which PODs to review (if a full review of every POD is infeasible).

The four subroutines are:

- **MATCHING\_MTRS\_OR\_FFMTRS**
  - Flags PODs whose Public Land Survey System (PLSS) location information on eWRIMS matches a PLSS section that overlaps with the watershed polygon
- **LESS\_THAN\_ONE\_MILE\_WITHIN\_WATERSHED\_BOUNDARY**
  - Flags PODs whose geographic coordinates fall within the watershed boundary, but they are less than one mile from the boundary line
- **ONE\_MILE\_OR\_MORE\_WITHIN\_WATERSHED\_BOUNDARY**
  - Flags PODs whose geographic coordinates fall within the watershed boundary, but they are at least one mile away from the boundary line (i.e., the PODs are plotted at least one mile inside of the watershed)
  - These PODs are less likely to be misplotted in the wrong watershed than the ones flagged by “LESS\_THAN\_ONE\_MILE\_WITHIN\_WATERSHED\_BOUNDARY”. If the total number of PODs to review is excessive, then all of these PODs could be assumed to be correctly plotted.
- **MENTIONS\_WATERSHED\_IN\_SOURCE\_INFORMATION**
  - Flags PODs whose “WATERSHED”, “SOURCE\_NAME”, and/or “TRIB\_DESC” fields on eWRIMS match with at least one of the search strings specified in Step 3 of [Section 4](#).

The “R\_Review” worksheet in “[ID]\_GIS\_Preprocessing.xlsx” provides a table for performing the manual review.

It is intended to be compatible with the (optional) script “POD\_StreamStats\_Analysis.R”.

## 5.2 Procedure

1. For each POD that will be reviewed, open the water right’s documentation on eWRIMS.
  - a. The “URL” column in the “R\_Review” sheet contains links to the rights’ primary documentation.
2. In the “R\_Review” sheet, check the text describing the POD location in these three fields:

- a. "WATERSHED"
- b. "SOURCE\_NAME"
- c. "TRIB\_DESC"

Ideally, these fields should mention the watershed of interest. If the text lists a different watershed—the POD should be excluded from the final list, regardless of its coordinates. For example, the Butte Creek watershed contains PODs within its boundaries that actually divert from the Sacramento River instead, so these PODs were excluded from the final demand dataset.

- d. In the "R\_Review" worksheet, if "POD\_StreamStats\_Analysis.R" will be used, to exclude a POD from the dataset without checking its coordinates on USGS StreamStats, type "TRUE" into the "MANUAL\_OVERRIDE: REMOVE POD" column.

Blanks in these three fields might be okay because sometimes eWRIMS simply lacks this information. But you will still need to confirm the watershed by reviewing the original documentation.

3. After that, investigate the documentation for descriptions of the POD locations.
  - a. The locations may be described using one of several different methods:
    - i. Geographic coordinates (latitude and longitude values, with the datum often being "WGS84" or "NAD83")
    - ii. Northing and Easting coordinates (a projected coordinate system that uses either the NAD27 or the NAD83 datum)
    - iii. Deviations from a PLSS section corner
  - a. If "POD\_StreamStats\_Analysis.R" will **not** be used, convert these descriptions into geographic coordinates (with the "WGS84" datum) and use the "Flow (Raindrop) Path" tool on [USGS StreamStats](#) to confirm that the flow path exits through the watershed outlet.
4. If "POD\_StreamStats\_Analysis.R" will be used, fill out the corresponding "REPORT" fields in the "R\_Review" sheet with information from the documentation's POD descriptions.
  - a. For geographic coordinates, use these fields:
    - i. "REPORT\_LATITUDE"
    - ii. "REPORT\_LONGITUDE"

- iii. "LAT\_LON\_CRS"
- b. For northing and easting coordinates, use these fields:
  - i. "REPORT\_NORTHING"
  - ii. "REPORT\_EASTING"
  - iii. "NOR\_EAS\_CRS"
- c. For deviations from a PLSS section corner, use these fields:
  - i. "REPORT\_SECTION\_CORNER"
  - ii. "REPORT\_NS\_MOVE\_FT"
  - iii. "REPORT\_NS\_DIRECTION"
  - iv. "REPORT\_EW\_MOVE\_FT"
  - v. "REPORT\_EW\_DIRECTION"
  - vi. "REPORT\_SECTION"
  - vii. "REPORT\_TOWNSHIP"
  - viii. "REPORT\_RANGE"
  - ix. "REPORT\_DATUM"
- d. Please consult "[R\\_Review Sheet Data Dictionary.xlsx](#)" for detailed instructions on how to enter values into the fields listed above so that they are compatible with the "POD\_StreamStats\_Analysis.R" script. Spreadsheets from prior manual reviews can also be used as reference material.
- e. Most rights' documentation will only use one of the three POD description methods. Therefore, some of the columns will likely remain blank during the manual review process.
- f. For "LAT\_LON\_CRS" and "NOR\_EAS\_CRS", sometimes, the datum will not be specified explicitly in the documentation.
  - i. For geographic coordinates, "WGS84" is generally the best guess in most cases due to its widespread usage. (If information from USGS was used in specifying the POD location, "WGS84" is very likely to be the coordinate reference system that was used.)
  - ii. For northing and easting coordinates, the values should be tested with both datums (for example, by converting them to a different system like "WGS84" or "NAD83"). Because of the differences between the two datums, one coordinate pair will be more sensible than the other one.
- g. Even if you opt not to use "POD\_StreamStats\_Analysis.R", filling out the "REPORT" fields accurately may assist with tracking POD information in your manual review.

5. If “POD\_StreamStats\_Analysis.R” will **not** be used, skip to Step 18.
6. Before running “POD\_StreamStats\_Analysis.R”, open “Watershed\_Demand\_Dataset\_Paths.xlsx”. Specify the location of the “[ID]\_GIS\_Preprocessing.xlsx” file whose “R\_Review” worksheet you completed.
  - a. The columns to update are:
    - i. GIS\_PREPROCESSING\_SPREADSHEET\_PATH
      1. The general file path specifications mentioned in Step 2 of Section 4.2 apply to this column and other similarly named columns.
    - ii. GIS\_PREPROCESSING\_WORKSHEET\_NAME
      1. This is the worksheet name of the relevant table in the spreadsheet file. Other columns with a similar name as this one have a similar purpose.
    - iii. IS\_SHAREPOINT\_PATH\_GIS\_PREPROCESSING\_SPREADSHEET
      1. The general file path specifications mentioned in Step 2 of Section 4.2 apply to this column and other similarly named columns.
      2. The “Dictionary” worksheet in “Watershed\_Demand\_Dataset\_Paths.xlsx” has explanations for each of these columns.
7. Run “POD\_StreamStats\_Analysis.R”.
  - a. This can be accomplished by running Line 42 of “Demand\_Master\_Script.R”.
8. If this is your first time running the script, you should receive an error message (shown below in Figure 8).

```
Error: No exit point was chosen for watershed Gualala River.  
Please use the code 'mapview(wsPoints)' to view the points in 'wsPoints'. Then,  
choose the point (using its ID/row number) that best represents the watershed's  
exit area.
```

Figure 8. An error message that appears if the watershed exit point has not yet been specified.

9. Open “POD\_StreamStats\_Analysis.R” and run Lines 27 through 47.
10. In the console, run this code: **mapview(wsPoints)**. See Figure 9 for an example.
  - a. An alternative option is to remove the comment hashtag (“#”) from Line 66 and execute that line.

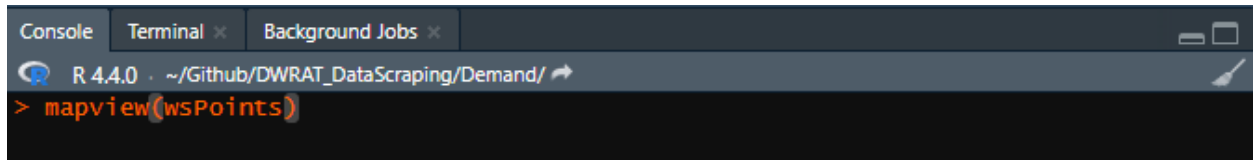


Figure 9. A line of code that can generates a map of the variable “wsPoints”.

11. Examine the map that appears in the “Viewer” pane. You should see the boundary of your watershed defined with many points (like in Figure 10).

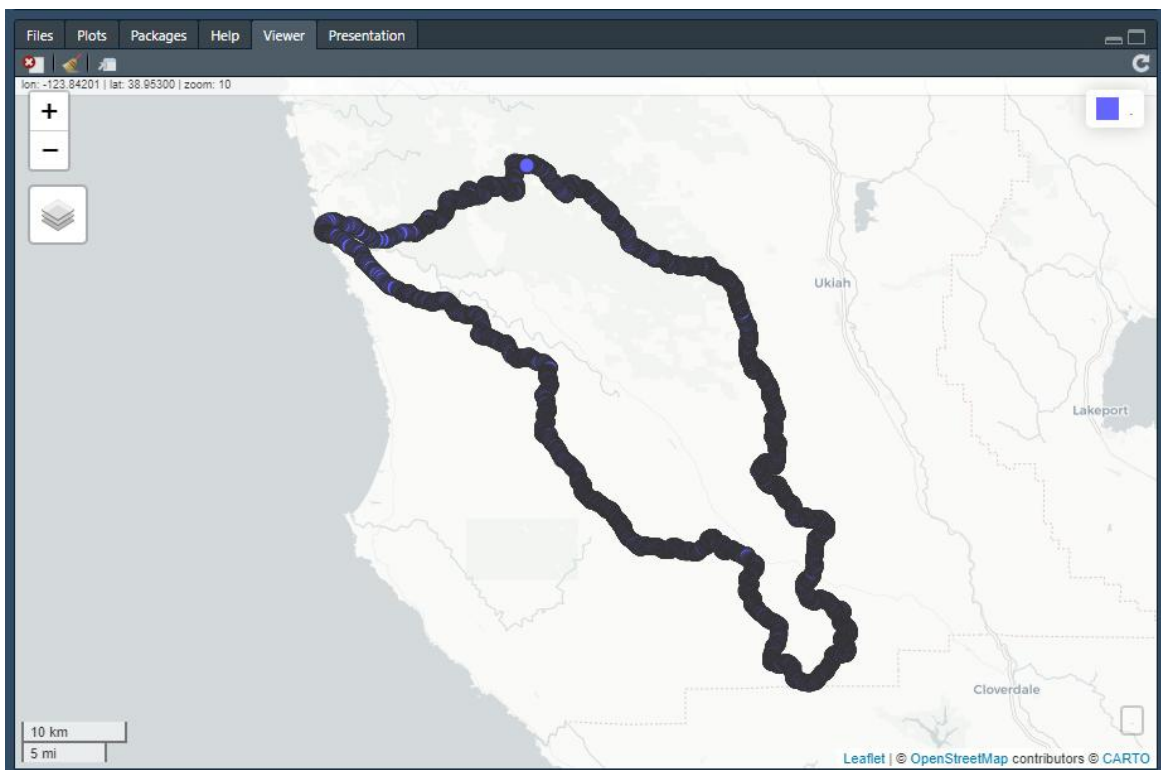


Figure 10. The boundary line of the Navarro River watershed as a series of points.

12. Zoom into the exit of the watershed. Identify the point that is closest to the middle of the exit.
13. Hover over the chosen point. Note the integer value that appears in the pop-up box (as shown in Figure 11).

14. Input this number into “WATERSHED\_EXIT\_POINT\_INDEX” in “Watershed\_Demand\_Dataset\_Paths.xlsx”.

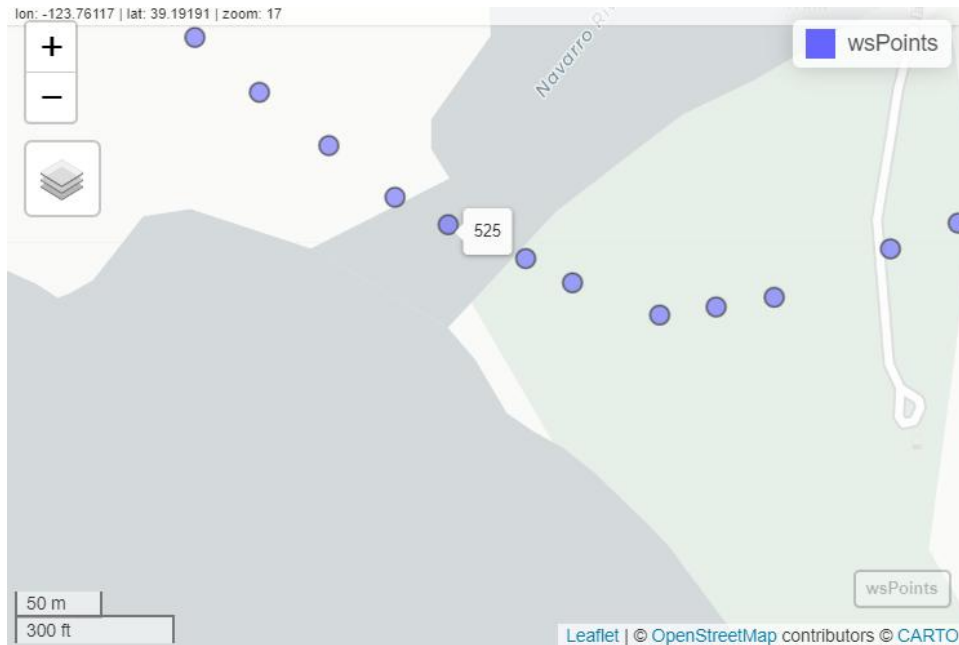


Figure 11. A map of “wsPoints” for the Navarro River watershed that is zoomed in on the watershed exit. The point closest to the middle of the exit has an ID of 525. This number will be entered into “Watershed\_Demand\_Dataset\_Paths.xlsx” as the exit point of the watershed. When POD coordinates are fed into USGS StreamStats, the resultant flowlines will be compared to this point to verify that the flow path from the POD exits through this location (thereby confirming that the POD is within the watershed boundaries).

15. After updating and closing “Watershed\_Demand\_Dataset\_Paths.xlsx”, clear the R environment.

16. Rerun “POD\_StreamStats\_Analysis.R”.

- a. This can be accomplished by running Line 39 of “Demand\_Master\_Script.R”.
- b. This script may take a while to run because it is checking each POD’s coordinates with USGS StreamStats.

17. Once the script has completed its procedure, a spreadsheet called “[ID]\_POD\_StreamStats\_Review.xlsx” will appear in the “OutputData” folder.

- a. The “Final\_List” worksheet lists the PODs that divert water from the watershed.



- i. This table contains:
  - 1. PODs whose raindrop flow path was found to drain through the watershed exit (according to USGS StreamStats).
  - 2. PODs that have a non-empty value for “MANUAL\_OVERRIDE: KEEP POD”.
- ii. PODs with a non-empty value for “MANUAL\_OVERRIDE: REMOVE POD” will **not** be included in the final list, regardless of whether the aforementioned conditions apply.

18. If you did **not** use “POD\_StreamStats\_Analysis.R”, create a spreadsheet that contains all PODs that you reviewed and confirmed to divert from the watershed.

- a. Also include the PODs that you did not review but assumed to be present within the watershed boundaries (for example, if you did not review the PODs that are one mile or more within the watershed, include them in this spreadsheet).
- b. This spreadsheet should contain at least two fields:
  - “APPLICATION\_NUMBER” and “POD\_ID”.
  - i. To assist with the sub-basin assignment process later in the procedure, also include the eWRIMS coordinates as “LATITUDE” and “LONGITUDE” fields in this spreadsheet.

19. Regardless of whether “POD\_StreamStats\_Analysis.R” was used, consider uploading the final spreadsheet to SharePoint. Then, update “Watershed\_Demand\_Dataset\_Paths.xlsx”.

- a. The columns to fill in are:
  - i. POD\_APPLICATION\_NUMBER\_SPREADSHEET\_PATH
  - ii. POD\_APPLICATION\_NUMBER\_WORKSHEET\_NAME
  - iii. IS\_SHAREPOINT\_PATH\_POD\_APPLICATION\_NUMBER\_SPREADSHEET

## 6. GIS Full Manual Review

A manual procedure for flagging and reviewing PODs is discussed in “[SDA GIS Full Manual Review Methodology.docx](#)”.

Once that procedure is complete, perform Step 19 of Section 5.2, using the file specifications outlined in Step 18.

## 7. Preliminary Demand Dataset

### 7.1 Description

The GIS Pre-processing steps finalize the list of water rights in the watershed. The next step is to extract the diversion data from this water right list. The selections made in “Dataset\_Year\_Range.R” will define the boundaries of this extraction.

This process will also assign priority dates to each water right. DWRAT requires this information to prioritize allocations. No manual efforts are required for this component of the procedure, but the assigned priority dates will be viewable in a spreadsheet. That output spreadsheet will also be automatically referenced by later scripts.

### 7.2 Procedure

1. Run several R scripts.
  - a. The names of these scripts are:
    - i. “Priority\_Date\_Preprocessing.R”
    - ii. “Priority\_Date.R”
    - iii. “Priority\_Date\_Postprocessing.R”
  - b. This can be accomplished by running Lines 50 through 58 of “Demand\_Master\_Script.R”.

### 7.3 Output

After all three scripts have run successfully, in the “IntermediateData” folder, there will be a file titled “[ID]\_[YEAR1]\_[YEAR2]\_Statistics\_FINAL.csv”, where “ID” is the two-letter watershed ID and “YEAR1” and “YEAR2” are the starting and ending bounds of the demand dataset (e.g., “NV\_2017\_2023\_Statistics\_FINAL.csv”). In this initial run—before any corrections have been implemented—the CSV file contains the raw eWRIMS diversion data for the water rights identified during the GIS manual review.

In addition, the “OutputData” folder will contain “[ID]\_Priority\_Date\_Scripted.xlsx”. No manual changes are required. Later scripts will reference this spreadsheet automatically, though moving the file to another location would cause an error at that step.

## 8. Demand Dataset QA/QC Flags

### 8.1 Description

The next scripts will flag records in the dataset for potential errors. A manual review will be required. The two main types of errors that these scripts will attempt to identify are **duplicate reporting errors and unit-conversion errors**. There is also a third potential issue related to **empty Report Management System (RMS) reports**, though it is quite rare.

**Duplicate Reporting Errors (Chapter 9):** Another issue can arise when a party owns multiple water rights. They may sometimes report the exact same values for each right. This is not an issue if they are simply splitting their total usage across the rights. However, if they are instead repeating their total usage in each report, this would be a case of duplicate reporting. “Multiple\_Owner\_Analysis.R” flags rights with the same owner and total volume in the same reporting year, and a manual review is needed to determine the reason for this duplication.

**Unit-Conversion Errors (Chapter 10):** The annual RMS reports should have monthly values reported in units of acre-feet. However, reporters may sometimes submit volumes in different units (usually gallons). The “Expected\_Demand.R” and “Expected\_Demand\_Units\_Issue\_Flagger.R” scripts attempt to identify these cases by flagging reports where the total volume is substantially different from some comparison value (e.g., a water right’s initial diversion amount).

**Empty RMS Reports (Chapter 11):** The final issue addressed in this procedure is related to empty RMS reports, where the diverter did not provide monthly volumes, but a total volume is specified. In the eWRIMS database, these reports’ volumes are presented as “NA” (“missing”). However, since a report exists, the values should be considered in the analysis, so a manual review may be needed to convert these “NA” values into proper numbers.

### 8.2 Procedure

1. Run several R scripts.
  - a. Their names are:
    - i. “Multiple\_Owner\_Analysis.R”
    - ii. “Expected\_Demand.R”
    - iii. “Expected\_Demand\_Units\_Issue\_Flagger.R”
    - iv. “Check\_Empty\_Reports.R”
  - b. This can be accomplished by running Lines 66 through 78 of “Demand\_Master\_Script.R”.

2. Confirm that the following spreadsheets are present in the “OutputData” folder:

- a. “[ID]\_Duplicate\_Reports\_Manual\_Review.xlsx”
- b. “[ID]\_Expected\_Demand\_Units\_QAQC.xlsx”
- c. “[ID]\_Expected\_Demand\_Units\_QAQC\_Median\_Based.xlsx”
- d. “[ID]\_Empty\_Reports\_Manual\_Review.xlsx”
  - i. This spreadsheet will not appear in the directory if a manual review is not required; when executing “Check\_Empty\_Reports.R”, the script will notify you whether a manual review is necessary.

## 8.3 Output


Each script will generate a manual review spreadsheet containing flagged records. The next step will be to review the flagged records and decide which ones need to be remediated.

# 9. Manual Review: Duplicate Reporting Errors

## 9.1 Description

The records in the duplicate reporting manual review spreadsheet (“[ID]\_Duplicate\_Reports\_Manual\_Review.xlsx”) should be sorted by owner (“PARTY\_ID”). There may be multiple cases of duplicate reporting across several years, and owners with more than two water rights may make this mistake across several different rights. In addition, a few flagged records may be for the same water right and year—but with different diversion types.

To investigate these records, check the rights’ RMS reports, including those from before and after the flagged years. Estimate typical usage values for each water right and compare them to the values in each flagged report. If the values are not unusual for each right, then this may not be a case of duplicate reporting.



[Menu](#) | [Help](#)

Navigate to:

You are logged-in as: palemi . If this account does not belong to you, please


### e-WRIMS Advanced Water Right Search

Enter query criteria below and click "Search".

Search
Reset Search

Water Right Type: <small>(Ctrl-click to select multiple values)</small>	Adjudicated Appropriative Appropriative (State Filing) Cert of Right - Power
Sub Type: <small>(Check all that applies)</small> <b>(FOR APPROPRIATIVE/TEMP PERMIT ONLY)</b>	<input type="checkbox"/> Standard Groundwater Streamlined <input type="checkbox"/> Temporary 180-day <input type="checkbox"/> Temporary 5-year
Sub Type: <small>(Check all that applies)</small> <b>(FOR STATEMENTS ONLY)</b>	<input type="checkbox"/> Riparian <input type="checkbox"/> Pre 1914 <input type="checkbox"/> Court Decree/ Adjudication ==> #: <input type="text"/> <input type="checkbox"/> Other ==> Specify: <input type="text"/> <input type="checkbox"/> Un-Authorized Diversion <input type="checkbox"/> Incomplete Statement
Water Right Status: <small>(Ctrl-click to select multiple values)</small>	Active Adjudicated Cancelled
Application ID:	<input style="background-color: #f0f0f0;" type="text" value="A024524"/>
Permit ID:	<input type="text"/>
License ID:	<input type="text"/>
Special Use Area:	<input type="text"/>
PWSS ID:	<input type="text"/>
County: <small>(Ctrl-click to select multiple values)</small>	Alameda Alpine Amador Butte
Primary Owner:	<input type="text"/>
Entity Type: <small>(Ctrl-click to select multiple values)</small>	Corporation Estate Federal Government Government (State/Municipal)
CDTFE Account Number:	<input type="text"/>
Assignment Type: <small>(Ctrl-click to select multiple values)</small>	<b>Data Management</b> Ownership Change Address/Agent Update Phone Log
Assignment Status: <small>(Ctrl-click to select multiple values)</small>	Pending On Hold Cancelled Completed
Assigned To:	<input type="text"/>

Figure 12. eWRIMS water rights record search page.

 **CIWQS**

Navigate to:  [Menu](#) | [Help](#) | [Log out](#)

You are logged-in as: palemi . If this account does not belong to you, please log out.

### eWRIMS Water Right Search Results

[\[Return to Water Right Search\]](#)

Criteria: Displaying Water Rights where Application Number like "A024524".

Search Results: [previous](#) 1-1 of 1 [next](#)


Application ID	Permit ID	License ID	Water Right Type	Status	Primary Owner	Status Date	Face Amt	County	HUC 8 Watershed	Source Stream	View Reports	View Document	Open in GIS	At a Glance	Export to Excel
A024524	017726	011813	Appropriative	Licensed	JAMES C GOWAN	01/14/1974	8.4 acre-ft/yr	Mendocino	Big-Navarro-Garcia	UNST	<a href="#">View Reports</a>	<a href="#">View Document</a>	<a href="#">Open in GIS</a>	<a href="#">At a Glance</a>	<a href="#">Download to Excel</a>

[Generate Up/Down stream Mailing Address](#)
[Generate Mailing Address Export](#)
[Generate Mailing Address Export With IDs](#)
[Download to Excel](#)
[Open In GIS](#)

\* Primary Water Holder Name data changes may experience up to a 24 hour delay.

© 2025 State of California. [Conditions of Use](#) [Privacy Policy](#)

Figure 13. Click *View Reports* to pull up a list of all the annual water usage reports.

 **CIWQS**

Navigate to:  [Menu](#) | [Help](#) | [Log out](#)

You are logged-in as: palemi . If this account does not belong to you, please log out.

### eWRIMS Reporting

[\[Return to Water Right Search\]](#) [\[Return to Water Right Search Results\]](#)

Application ID: A024524 [View Document](#)      Water Right Type: Appropriative      [Open in GIS](#) [Paperless Documents](#)

Permit ID: 017726      Water Rights Status: Licensed (01/14/1974)      [View Map](#) [Report Data Errors](#)

License ID: 011813      Primary Owner: JAMES C GOWAN

[Application](#) [Parties](#) [Status](#) [Permit/License](#) [Uses and Seasons](#) [Special Uses](#) [PODs and Sources](#) [Measurements](#) [Notices](#) [Protests](#)  
[Petitions](#) [Complaints](#) [Inv/Viol/Enf](#) [Related Records](#) [Documents](#) [E-mail](#) [Returned Docs](#) [Forms and Surveys](#) [Reporting](#) [Annual Fees](#)  
[Payments](#) [Assignments](#) [Summary](#)

#### Reports

Select	Year	Revision	Report Type	Date Received	Method	Fee Req'd	Fee Paid	Validated	Needs Further Review	Submitter Phone & Email	Action
<input type="checkbox"/>	2024	1	Report of Licensee	01/31/2025	Electronic	Yes	No			707 205 1545 don@gowanscider.com	<a href="#">View</a>
<input type="checkbox"/>	2023	1	Report of	01/07/2024	Electronic	Yes	No			7072051545	<a href="#">View</a>

Figure 14. Click *View* to pick the water usage report for reporting year 2024.

[SUMMARY OF FINAL SUBMITTED VERSION]

REPORT OF LICENSEE FOR REPORTING PERIOD

October 1, 2023 to September 30, 2024

Primary Owner: JAMES C GOWAN  
Primary Contact:

Date Submitted: 01/31/2025

Application Number: A024524  
License Number: 011813

Source(s) of Water      POD Parcel Number      County  
UNST      260-360-07      Mendocino

MAX Direct Diversion Rate: 0.41 CFS  
MAX Collection to Storage: 0.9 AC-FT  
Face Value: 8.4 AC-FT

Permitted Use(s)	Acres	Direct Diversion Season	Storage Season
Frost Protection	3.50	03/15 to 05/15	02/01 to 07/01
Irrigation	3.50	04/01 to 11/01	02/01 to 07/01

Purpose of Use
Data Not Available

Special Use Categories	
Are you using any water diverted under this right for the cultivation of cannabis?	No

Amount of Water Diverted	
Month	Amount diverted or collected to storage
2023	
October	0 Acre-Feet 0 Gallons
November	0 Acre-Feet 0 Gallons
December	0 Acre-Feet

Figure 15. The top portion of the 2024 water usage report for A024524.

If included, these report elements may also provide useful context:

- Comments and remarks
- Flow measurement spreadsheets
- Purpose of Use

When inputting corrective actions (or no action) in the “QAQC\_Action\_Taken” column, please be sure to use the exact same language described in “QAQC Action Dictionary.xlsx” (spelling, capitalization, and punctuation matter). Otherwise, subsequent R scripts will produce errors.

## 9.2 Review Example: A024524 and A024526 (2020 Direct)

In 2020, these two rights had the exact same “Direct Diversion” total (2 AF). However, the monthly values were different, and the rights typically have different totals in most reporting years. Therefore, this was likely a coincidence, so “None” may be the most appropriate action for this case (see Figure 16).

APPLICATION_NUMBER	YEAR	DIVERSION_TYPE	PARTY_ID	AnnualTotal	PK_WR	QAQC_Action_Taken
A024524	2020	DIRECT	403983	2	2020_A024524_DIRECT_2	None
A024526	2020	DIRECT	403983	2	2020_A024526_DIRECT_2	None

Figure 16. The entries for A024524 and A024526 in the Navarro River watershed’s review spreadsheet for duplication errors. This occurrence may have been a coincidence.

## 9.3 Review Example: D030704 (2017 Direct and Storage)

In 2017, this right had the same total “Direct Diversion” and “Diversion to Storage” volumes (4 AF). Other RMS reports for D030704 typically had only one column filled in—in addition, the right is only authorized to perform “Diversion to Storage”. Therefore, this was likely a mistake, and “Keep Storage” may be the most appropriate action (see Figure 17).

APPLICATION_NUMBER	YEAR	DIVERSION_TYPE	PARTY_ID	AnnualTotal	PK_WR	QAQC_Action_Taken
D030704	2017	STORAGE	425883	4	2017_D030704_STORAGE_4	Keep Storage
D030704	2017	DIRECT	425883	4	2017_D030704_DIRECT_4	None

Figure 17. The entries for D030704 in the Navarro River watershed’s review spreadsheet for duplication errors. Please note that the corrective action “Keep Storage” was only mentioned once, and the second entry simply had “None” listed.

## 9.4 Review Example: A024943, A025060, S008688, S019745, S019747, S019754, and S019775 (2017–2023 Storage)

These seven water rights all had the same owner between 2017 and 2023. At various points in this period, some of the rights had the same total reported volume. Since the duplicated pairs are not consistently the same rights, these may be coincidences. Still, further investigation is warranted.

Most reports lack a measurement spreadsheet and do not contain information about the measurement methodology. In some prior submissions, the reporter mentioned relying on visual estimates, and RMS reports for A024943 contain details about a staff gauge (with



spreadsheets containing the recorded levels). However, this information does not clarify whether the user’s total water usage is split between the rights.

Another source of insight is the “Reported Use” sections in the RMS reports. The stated uses differ somewhat between the rights; all have some combination of “Irrigation”, “Frost Protection”, “Heat Protection”, and “Recreational” uses listed. This suggests that the reporter submitting this information may have been aware that each RMS report and use volume corresponds to a different water right.

Overall, there are many ambiguities in this data, so it may be best to leave the volumes unchanged (see Figure 18).

APPLICATION_NUMBER	YEAR	DIVERSION_TYPE	PARTY_ID	AnnualTotal	PK_WR	QAQC_Action_Taken
S019747	2017	STORAGE	641840	5	2017_S019747_STORAGE_5	None
S019754	2017	STORAGE	641840	5	2017_S019754_STORAGE_5	None
A025060	2018	STORAGE	641840	4	2018_A025060_STORAGE_4	None
S019747	2018	STORAGE	641840	2.5	2018_S019747_STORAGE_2.5	None
S019754	2018	STORAGE	641840	4	2018_S019754_STORAGE_4	None
S019775	2018	STORAGE	641840	2.5	2018_S019775_STORAGE_2.5	None
A024943	2019	STORAGE	641840	11	2019_A024943_STORAGE_11	None
A025060	2019	STORAGE	641840	3.5	2019_A025060_STORAGE_3.5	None
S008688	2019	STORAGE	641840	11	2019_S008688_STORAGE_11	None
S019775	2019	STORAGE	641840	3.5	2019_S019775_STORAGE_3.5	None
A025060	2020	STORAGE	641840	2.5	2020_A025060_STORAGE_2.5	None
S019747	2020	STORAGE	641840	2.5	2020_S019747_STORAGE_2.5	None
S008688	2021	STORAGE	641840	1.75	2021_S008688_STORAGE_1.75	None
S019754	2021	STORAGE	641840	1.75	2021_S019754_STORAGE_1.75	None
A025060	2022	STORAGE	641840	2	2022_A025060_STORAGE_2	None
S019775	2022	STORAGE	641840	2	2022_S019775_STORAGE_2	None
S019745	2023	STORAGE	641840	1	2023_S019745_STORAGE_1	None
S019775	2023	STORAGE	641840	1	2023_S019775_STORAGE_1	None

Figure 18. The flagged entries of A024943, A025060, S008688, S019745, S019747, S019754, and S019775 that appear in the Navarro River watershed’s review spreadsheet for duplication errors. The duplicated volumes are not consistently for the same water rights, and stated uses generally differ between the rights, so the most prudent decision may be to leave the data unchanged.

## 9.5 After Completion of the Manual Review

Although the manual review spreadsheet is initially accessed from the “OutputData” folder in the GitHub repository, it is recommended that you save this spreadsheet to the appropriate subfolder within the SharePoint folder that corresponds to your watershed.

Once your manual review is complete, open “Watershed\_Demand\_Dataset\_Paths.xlsx” and fill in the following columns:

- “QAQC\_DUPLICATE\_REPORTING\_SPREADSHEET\_PATH”
- “QAQC\_DUPLICATE\_REPORTING\_WORKSHEET\_NAME”
- “IS\_SHAREPOINT\_PATH\_QAQC\_DUPLICATE\_REPORTING\_SPREADSHEET”

Descriptions of these columns are provided in the “Dictionary” worksheet of “Watershed\_Demand\_Dataset\_Paths.xlsx”.

## 10. Manual Review: Unit Conversion Errors

### 10.1 Description

There are two spreadsheets in the *OutputData* folder to review for unit conversion errors: “[ID]\_Expected\_Demand\_Units\_QAQC.xlsx” and “[ID]\_Expected\_Demand\_Units\_QAQC\_Median\_Based.xlsx”. The two spreadsheets’ source scripts have different methodologies for identifying potentially erroneous reporting (the former compares annual totals to the right’s face value or initial diversion amount while the latter compares annual totals to the right’s median and average annual totals). But in both cases, the review procedure is the same.

When investigating these reports, assess how atypical the flagged values are. Unit conversion errors that result in very large volumes are the easiest mistakes to identify because these values are usually much larger than the right’s face value (the maximum allowable annual diversion amount for an appropriative right) (or other comparison value). In other cases, the report may contain an explanation for the unusual volumes. The water right’s other RMS submissions can also be used in this assessment; the volumes stated in those reports can help clarify what the user’s typical usage volumes are.

In more complicated cases (such as when the user consistently applies a smaller unit conversion error across several years), this manual review may require estimating appropriate volumes for the right’s stated uses and comparing those estimates to the reported values.

Finally, as with the Duplicate Reporting manual review, corrective actions should use the exact same language described in “[QAQC\\_Action\\_Dictionary.xlsx](#)” (spelling, capitalization, and punctuation matter).

### 10.2 After Completion of the Manual Review

It is recommended that you save “[ID]\_Expected\_Demand\_Units\_QAQC.xlsx” and “[ID]\_Expected\_Demand\_Units\_QAQC\_Median\_Based.xlsx” to the appropriate subfolder in your SharePoint watershed folder. Once the manual review is complete, update these columns in “Watershed\_Demand\_Dataset\_Paths.xlsx”:

- “QAQC\_UNIT\_CONVERSION\_ERRORS\_SPREADSHEET\_PATH”
- “QAQC\_UNIT\_CONVERSION\_ERRORS\_WORKSHEET\_NAME”
- “IS\_SHAREPOINT\_PATH\_QAQC\_UNIT\_CONVERSION\_ERRORS\_SPREADSHEET”

- “QAQC\_MEDIAN\_BASED\_UNIT\_CONVERSION\_ERRORS\_SPREADSHEET\_PATH”
- “QAQC\_MEDIAN\_BASED\_UNIT\_CONVERSION\_ERRORS\_WORKSHEET\_NAME”
- “IS\_SHAREPOINT\_PATH\_QAQC\_MEDIAN\_BASED\_UNIT\_CONVERSION\_ERRORS\_SPREADSHEET”

Please consult the “Dictionary” worksheet in “Watershed\_Demand\_Dataset\_Paths.xlsx” for any confusion regarding these columns.

## 11. Manual Review: Empty Reports

### 11.1 Description

This manual review covers two different issues:

- Some RMS reports were submitted with no monthly volumes. The user may have included a total volume for that year, but in the dataset, all monthly volumes are “NA” (“missing”).
- Water rights that began diverting after 2021 will appear in the dataset as having only “NA” values for the year prior to their first actual report. This is due to an operation performed on the eWRIMS dataset in “GIS\_POD\_Flat\_File\_Prep.R”. Starting in 2022, the reported monthly volumes are stored by water year, while prior years’ entries are stored by calendar year. A procedure in the code converts the water years into calendar years to match the pre-2022 submissions. However, if a water right began reporting in 2022 or later, this procedure ends up creating “NA” entries for the prior calendar year.

This review is important due to the significance of “NA” in the dataset. **Only RMS reports with at least one non-missing volume are considered when calculating the average monthly volumes for a water right.** In the case of the first issue, the user submitted a report about their water usage, and if they did include a total volume, that report’s data should be included in the averages used by DWRAT. And “NA” reports introduced by the second issue *should be removed from the dataset* since they add unnecessary confusion.

This manual review spreadsheet is already prepopulated with determinations for the second issue case. Therefore, the primary focus of the manual review is checking the reports of the first issue case. If the reports do contain a total volume, the “NA” entries for that year should be replaced with numeric values. Right now, the scripts are designed to only handle cases where the total volume is 0 AF; updates would be needed to handle non-zero totals. In the spreadsheet, simply entering “TRUE” in the column titled “REPLACE\_NA\_VALUES\_WITH\_ZEROS” will enable this replacement procedure.

## 11.2 After Completion of the Manual Review

Be sure to save “[ID]\_Empty\_Reports\_Manual\_Review.xlsx” to the appropriate subfolder in your SharePoint watershed folder. In “Watershed\_Demand\_Dataset\_Paths.xlsx”, fill out the following columns with information about the file:

- “NA\_REPORTS\_SPREADSHEET\_PATH”
- “NA\_REPORTS\_WORKSHEET\_NAME”
- “IS\_SHAREPOINT\_PATH\_NA\_REPORTS\_SPREADSHEET”

## 12. Updated Demand Dataset

1. In “Watershed\_Demand\_Dataset\_Paths.xlsx”, input paths to the completed review spreadsheets if you have not already done so.
2. Re-run “GIS\_POD\_Flat File\_Prep.R” (to get the latest flat files).
3. Re-run from “Priority\_Date\_Preprocessing.R” to “Check\_Empty\_Reports.R”.
  - a. Issues with the Unit Conversion Errors and Duplicate Reporting manual review spreadsheets will show up when running “Priority\_Date\_Postprocessing.R”.
4. After successfully running these scripts, the output spreadsheets of the four manual review scripts should be empty. If not, append these new rows to your existing manual review spreadsheets (the ones listed in “Watershed\_Demand\_Dataset\_Paths.xlsx”). Then, perform a manual review on these entries and re-run the scripts again from “Priority\_Date\_Preprocessing.R” to “Check\_Empty\_Reports.R”.
5. Once the four manual review scripts output spreadsheets with no flagged records, run “Beneficial\_Use\_Return\_Flow.R”.
  - a. This can be accomplished by running Line 81 of “Demand\_Master\_Script.R”.

## 13. Sub-Basin Delineation

### 13.1 Description

DWRAT expects the watershed of interest to be divided into separate sub-basins. Any kind of subdivision is acceptable for the model (as long as a connectivity matrix properly details

the relationship between the sub-basins). In addition, each water right should be assigned to one sub-basin within the watershed. This step of the procedure involves determining the proper sub-basin for each water right based on the locations of their PODs.

There are two different scripts that can be used for sub-basin assignment.

“Assign\_Subbasin\_to\_POD.R” is a script that will likely produce a manual review spreadsheet. “Assign\_Subbasin\_via\_Connectivity\_Matrix.R” is an automated counterpart to “Assign\_Subbasin\_to\_POD.R”, but it requires the sub-basin connectivity matrix. Both scripts at minimum require a GIS layer that delineates the watershed sub-basins.

If a sub-basin connectivity matrix is available, then you only need to run “Assign\_Subbasin\_via\_Connectivity\_Matrix.R”. This script contains an additional feature: right splitting. If a water right has PODs in sub-basins that are completely disconnected from one another (i.e., there are no flow connections between them), that right is divided into separate water rights (with flows proportioned based on the relative drainage areas of each sub-basin and their upstream sub-basins).

## 13.2 Procedure

1. In “Watershed\_Demand\_Dataset\_Paths.xlsx”, fill in the following columns with the specified files:
  - a. A spreadsheet containing all PODs in the watershed (with “APPLICATION\_NUMBER”, “LATITUDE”, and “LONGITUDE” fields included) as well as the coordinate reference system used by the dataset. This can be the same spreadsheet that you generated during the GIS manual review (if you included all of the aforementioned columns).
    - i. “POD\_COORDINATES\_SPREADSHEET\_PATH”
    - ii. “POD\_COORDINATES\_WORKSHEET\_NAME”
    - iii. “IS\_SHAREPOINT\_PATH\_POD\_COORDINATES\_SPREADSHEET”
    - iv. “POD\_COORDINATES\_REFERENCE\_SYSTEM”
  - b. A GIS layer containing the watershed sub-basins as well as field(s) within the layer that uniquely identify each sub-basin.
    - i. “SUBBASIN\_POLYGONS\_DATABASE\_PATH”
    - ii. “SUBBASIN\_POLYGONS\_LAYER\_NAME”
    - iii. “IS\_SHAREPOINT\_PATH\_SUBBASIN\_POLYGONS”
    - iv. “SUBBASIN\_FIELD\_ID\_NAMES”
      1. If more than one field is specified in this column (separated by semicolons), then the sub-basin scripts will retain all of these fields in their output spreadsheets. However, the first field in this list is assumed to be the sole column necessary for distinguishing between sub-basins (i.e., the sub-basin

connectivity matrix should use the same values as this field for distinguishing between different sub-basins).

2. If “Assign\_Subbasin\_via\_Connectivity\_Matrix.R” will be used, then additional fields must be completed:
  - a. A spreadsheet containing the connectivity matrix:
    - i. “CONNECTIVITY\_MATRIX\_SPREADSHEET\_PATH”
    - ii. “CONNECTIVITY\_MATRIX\_WORKSHEET\_NAME”
    - iii. “IS\_SHAREPOINT\_PATH\_CONNECTIVITY\_MATRIX\_SPREADSHEET”
  - b. The specified worksheet should contain **only** the connectivity matrix —no additional tables or values.
3. Run either “Assign\_Subbasin\_to\_POD.R” or “Assign\_Subbasin\_via\_Connectivity\_Matrix.R”
  - a. Either script can be executed using “Demand\_Master\_Script.R”

## 13.3 Output

Both scripts produce the same output file: “[ID]\_POD\_Subbasin\_Assignment.xlsx” in the “OutputData” folder of the DWRAT\_DataScraping Demand sub-repository. While running “Assign\_Subbasin\_to\_POD.R”, the script would have output a message if a manual review is necessary. If that is the case, there will be a worksheet titled “Review” in the output spreadsheet that contains the water rights with multiple potential sub-basins.

# 14. Manual Review: Sub-Basin Delineation

## 14.1 Description

Water rights with multiple PODs may have diversion locations in more than one sub-basin. “Assign\_Subbasin\_to\_POD.R” has no methodology to choose a single sub-basin, so this must be manually decided.

In general, the sub-basin with the greatest flow availability should be selected. For example, the most downstream sub-basin would be preferred, and a mainstem sub-basin should be chosen over a tributary sub-basin. Sub-basins with larger drainage areas or fewer neighboring PODs would also likely have greater flow availability.

“Assign\_Subbasin\_to\_POD.R” does not support right splitting, so one sub-basin must be chosen for each of the flagged water rights.

# 15. Master Demand Table

## 15.1 Description

The final step of the procedure is to generate the Master Demand Table that will be input into DWRAT. Previously generated files will automatically be read, arranged, and merged as needed to produce the table.

## 15.2 Procedure

1. If “Assign\_Subbasin\_to\_POD.R” was used and a manual review was required, update “Watershed\_Demand\_Dataset\_Paths.xlsx” to include a reference to the completed review spreadsheet.
  - a. The following fields should be filled in:
    - i. “SUBBASIN\_MANUAL\_ASSIGNMENT\_SPREADSHEET\_PATH”
    - ii. “SUBBASIN\_MANUAL\_ASSIGNMENT\_WORKSHEET\_NAME”
    - iii. “IS\_SHAREPOINT\_PATH\_SUBBASIN\_MANUAL\_ASSIGNMENT”
2. If “Assign\_Subbasin\_to\_POD.R” was used and a manual review was required, rerun “Assign\_Subbasin\_to\_POD.R” and verify that the script successfully completes its assignment procedure.
  - a. A new version of “[ID]\_POD\_Subbasin\_Assignment.xlsx” should be generated. This spreadsheet will be necessary in the next step.
3. Regardless of which sub-basin assignment script was used, update “Watershed\_Demand\_Dataset\_Paths.xlsx” to reference the file path to “[ID]\_POD\_Subbasin\_Assignment.xlsx”, which contains water rights and their sub-basin assignments.
  - a. The fields to update are:
    - i. “SUBBASIN\_ASSIGNMENT\_SPREADSHEET\_PATH”
    - ii. “SUBBASIN\_ASSIGNMENT\_WORKSHEET\_NAME”
    - iii. “IS\_SHAREPOINT\_PATH\_SUBBASIN\_ASSIGNMENT\_SPREADSHEET”
4. Run “MasterDemandTable.R”
  - a. This script can be executed using “Demand\_Master\_Script.R”

## 15.3 Output

The main outputs of “MasterDemandTable.R” are “[ID]\_[YEAR1]\_[YEAR2]\_DemandDataset\_MonthlyValues.csv” and “[ID]\_[YEAR1]\_[YEAR2]\_MDT\_[DATE].csv”, which are exported to the “OutputData” folder.

The former contains the monthly diversions for each water right that are averaged together for DWRAT. The latter file contains the proper master demand table. Not all fields in the file are used by DWRAT, but all required data is present.

## Appendix A: Compiled List of Requirements

- R and R Studio
- R Packages
  - tidyverse
  - sf
  - openxlsx
  - mapview
  - lwgeom
  - httr
  - data.table
  - odbc
  - DBI
  - readxl
  - janitor
  - writexl
- GIS layer containing the watershed boundaries
  - A single polygon should represent the entire watershed
- GIS layer containing the sub-watershed boundaries
- Watershed connectivity matrix

## Appendix B: Troubleshooting Errors While Running Scripts

1) Error: `path` does not exist:

The image shows a screenshot of a terminal window with a dark background. A red error message is displayed: "Error: `path` does not exist: 'C:/Users/palemi/water Boards/Supply and Demand Assessment - Documents/watershed Folders/Napa/GIS/Test\_POD\_StreamStats\_Review.xlsx'".

Figure 19.

Payman encountered this error when running this code snippet in the **Priority\_Date\_Preprocessing.R** script:

The image shows a screenshot of an R script editor with a dark background. The code is as follows:

```
19 # Read in the results of the GIS Pre-Processing steps
20 # The filename will differ depending on its location (i.e., whether or not it's on SharePoint)
21 Application_Number <- getXLSX(ws = ws,
22                               SHAREPOINT_BOOL = "IS_SHAREPOINT_PATH_POD_APPLICATION_NUMBER_SPREADSHEET",
23                               FILEPATH = "POD_APPLICATION_NUMBER_SPREADSHEET_PATH",
24                               WORKSHEET_NAME = "POD_APPLICATION_NUMBER_WORKSHEET_NAME")
25
```

Figure 20.

To troubleshoot this error, I checked the 4 arguments of the `getXLSX()` function:



- ws is correctly set to the ws variable in the environment, which represents the row in the Watershed\_Demand\_Dataset\_Paths (Watershed spreadsheet) that corresponds to the watershed of interest; in this case, the *Test* watershed.
- I looked at these 3 columns in the Watershed spreadsheet and checked the values for the *Test* watershed row:
  - IS\_SHAREPOINT\_PATH\_POD\_APPLICATION\_NUMBER\_SPREADSHEET was set to *TRUE*;
  - POD\_APPLICATION\_NUMBER\_SPREADSHEET\_PATH was set to
  - POD\_APPLICATION\_NUMBER\_SPREADSHEET\_PATH was set to *Napa/GIS/Test\_POD\_StreamStats.Review.xlsx*, which was incorrect. I corrected it to *Watershed Folders/Navarro/Data/GIS Preprocessing/Test\_POD\_StreamStats\_Review.xlsx*.
  - POD\_APPLICATION\_NUMBER\_WORKSHEET\_NAME was set to *Final\_List*.
- After I corrected the value in the relevant column of the Watershed spreadsheet, the error disappeared.

Many errors that you encounter while running the scripts probably pertain to mis-entries in the Watershed spreadsheet so check that you have typed the correct file paths in the relevant columns.

## Appendix C: Water Years and Calendar Years

Before 2022, RMS report submissions use **calendar years**. Parties would submit their usage volumes for each month of that year. For instance, a submission for Reporting Year 2020 contains monthly volumes from January 2020 to December 2020 (see Figure C-1 for an example).

Starting in 2022, RMS reports use **water years**. A water year starts from October of the previous calendar year and ends in September of the subsequent calendar year. Therefore, a submission for Reporting Year 2022 contains monthly volumes from October 2021 to September 2022 (see Figure C-3 for an example).

Because a water year has data from the previous calendar year, Reporting Year 2021 is a **special case**. It only covers January 2021 to September 2021. The last three months are part of Water Year 2022, so the 2021 submissions only contain data for nine months (see Figure C-2 for an example).

The current versions of the R scripts properly account for these differences, but any future fixes and updates to the procedure must always consider this discrepancy. In addition, the lack of data for October through December in Reporting Year 2021 can cause missing entries (“NA”) to appear in the script output files (see Figure C-4 for an example). This also means that water rights with RMS submissions prior to 2022 will have comparatively less

data available for October through December when calculating the average monthly volumes (because Reporting Year 2021 lacks these entries).

SUPPLEMENTAL STATEMENT OF WATER DIVERSION AND USE FOR 2020			
Primary Owner: GUALALA REDWOOD TIMBER LLC			
Statement Number: S021354			
Date Submitted: 06/30/2021			
1. Water is used under		Riparian Claim	
2. Year diversion commenced		1985	
<b>2. Purpose of Use</b>			
Mining		Gravel washing	
<b>3. Special Use Categories</b>			
C1. Are you using any water diverted under this right for the cultivation of cannabis?			No
<b>4. Amount of Water Diverted and Used</b>			
Month	Amount directly diverted (Acre-Feet)	Amount diverted or collected to storage (Acre-Feet)	Amount beneficially used (Acre-Feet)
January	0.0031	0	0.0031
February	0.0031	0	0.0031
March	0.0031	0	0.0031
April	0.0031	0	0.0031
May	0.0153	0	0.0153
June	0.0153	0	0.0153
July	0.0153	0	0.0153
August	0.0153	0	0.0153
September	0.0153	0	0.0153
October	0.0153	0	0.0153
November	0.0031	0	0.0031
December	0.0031	0	0.0031
Total	0.1104	0	0.1104
Type of Diversion	Direct Diversion Only		
Comments			

Figure C-1. Reporting Year 2020 submission for S021354. This report contains water usage data from January 2020 to December 2020.

SUPPLEMENTAL STATEMENT OF WATER DIVERSION AND USE FOR REPORTING PERIOD			
January 1, 2021 to September 30, 2021			
Primary Owner: GUALALA REDWOOD TIMBER LLC			
Statement Number: S021354			
Date Submitted: 03/31/2022			
Water is used under		Riparian Claim	
Year diversion commenced		1985	
<b>Purpose of Use</b>			
Mining		Gravel washing	
<b>Special Use Categories</b>			
Are you using any water diverted under this right for the cultivation of cannabis?			No
<b>Amount of Water Diverted and Used</b>			
Month	Amount directly diverted (Acre-Feet)	Amount diverted or collected to storage (Acre-Feet)	Amount beneficially used (Acre-Feet)
January	0.0031	0	0.0031
February	0.0031	0	0.0031
March	0.0031	0	0.0031
April	0.0031	0	0.0031
May	0.0153	0	0.0153
June	0.0153	0	0.0153
July	0.0153	0	0.0153
August	0.0153	0	0.0153
September	0.0153	0	0.0153
Total	0.0889	0	0.0889
Type of Diversion	Direct Diversion Only		
Comments			

Figure C-2. Reporting Year 2021 submission for S021354. Because of the change from calendar years to water years, these reports only cover January 2021 to September 2021.

SUPPLEMENTAL STATEMENT OF WATER DIVERSION AND USE FOR REPORTING PERIOD			
October 1, 2021 to September 30, 2022			
Primary Owner: GUALALA REDWOOD TIMBER LLC			
Statement Number: S021354			
Date Submitted: 10/30/2023			
Water is used under		Riparian Claim	
Year diversion commenced		1985	
<b>Purpose of Use</b>			
Mining	Gravel washing		
<b>Special Use Categories</b>			
Are you using any water diverted under this right for the cultivation of cannabis?			No
<b>Amount of Water Diverted and Used</b>			
Month	Amount directly diverted (Acre-Feet)	Amount diverted or collected to storage (Acre-Feet)	Amount beneficially used (Acre-Feet)
<b>2021</b>			
October	0.0153	0	0.0153
November	0.0031	0	0.0031
December	0.0031	0	0.0031
<b>2022</b>			
January	0.0031	0	0.0031
February	0.0031	0	0.0031
March	0.0031	0	0.0031
April	0.0031	0	0.0031
May	0.0153	0	0.0153
June	0.0153	0	0.0153
July	0.0153	0	0.0153
August	0.0153	0	0.0153
September	0.0153	0	0.0153
Total	0.1104	0	0.1104
Type of Diversion	Direct Diversion Only		
Comments			

Figure C-3. Reporting Year 2022 submission for S021354. It contains monthly volumes from October 2021 to September 2022.

APPLIC.	YEAR	JAN_TO	FEB_TO	MAR_TC	APR_TC	MAY_TC	JUN_TO	JUL_TO	AUG_TC	SEP_TO	OCT_TC	NOV_TC	DEC_TC	L_DIVERSION
A021202	2017	0	0	0	0	0	0	0	0	0	0	0	0	
A021202	2018	0	0	0	0	0	0	0	0	0	0	0	0	
A021202	2019	0	0	0	0	0	0	0	0	0	0	0	0	
A021202	2020	0	0	0	0	0.036827	0.036827	0.039896	0.046033	0.049102	0	0	0	
A021202	2021	0	0	0	0	0.033758	0.036827	0.036827	0.042964	0.046033	NA	NA	NA	
A021202	2022	0	0	0	0	0.033758	0.036827	0.039896	0.042964	0.046033	0.049102	0	0	
A021202	2023	0	0	0	0	0.033758	0.033758	0.036827	0.039896	0.042964	0.042964	0	0	

Figure C-4. Diversion data for water right A021202 as shown in the file “GL\_2017\_2023\_DemandDataset\_MonthlyValues.csv”. In the row corresponding to Reporting Year 2021, the last three months’ columns contain “NA” (i.e., they are treated as “missing values”).