

Water Balance Automation
Distributed Data Production
&
Atlas of Water Balance

29 May 2025

John Helly ¹, Jennifer Stricklin ²,
George Valente³, James Common⁴, Melika Mani²

¹Helly Associates

²Headquarters, California Department of Water Resources, Sacramento, California

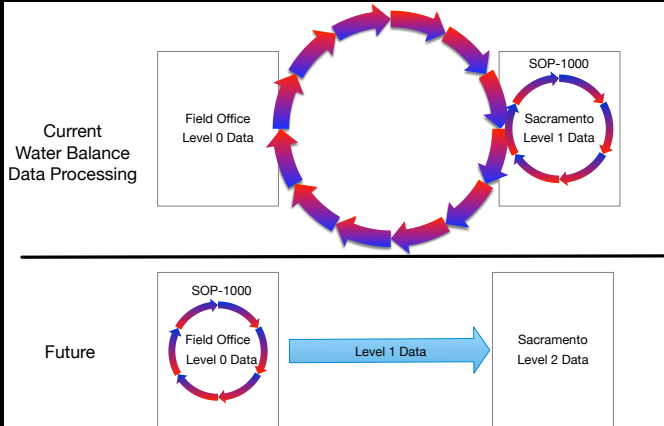
³Regional Office, California Department of Water Resources, Northern, California

⁴Regional Office, California Department of Water Resources, North Central, California

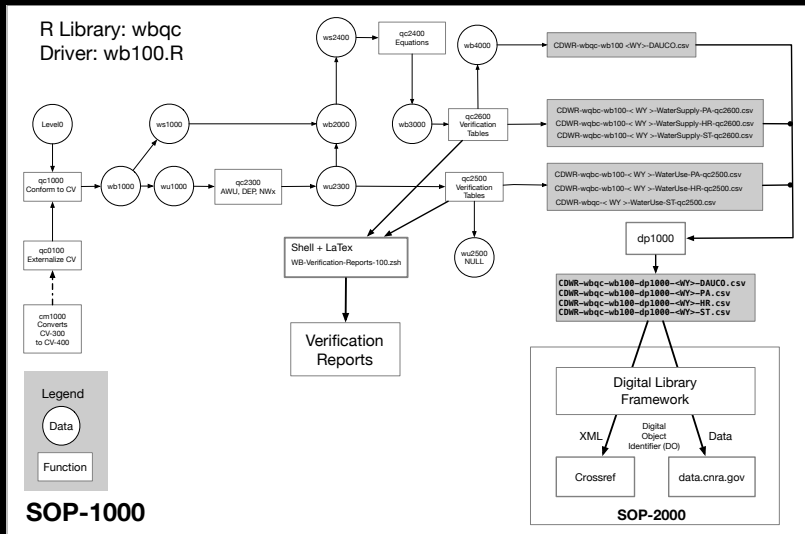
- ① Goals of Water Balance Automation
- ② Software and Training
- ③ Extension to Atlas of Water Balance

Goals of Water Balance Automation

High Quality, Less Time, Less Cost



- Automated, Centralized Level 1 Data Production
- Data Standardization
- < Time for WY Completion
- > Error Detection & Correction
- < Time Level 1 Data Production
- < Time for WY Completion



Single R-Script Encapsulates Entire Data Production Workflow

```
44 - #
45 - for ( WY in c(2016:2021) ){
46   #for ( WY in c(2016:2016) ){
47     # WY = 2016
48     time1 = Sys.time()
49     cat(sprintf('wb100: Start: %d at %s\n', WY, format(time1, format = "%X %R %Z")))
50   #
51   wb0100 = qc0100A(WY) # Ingest and integrate regional data
52   wb1000 = qc1000(wb0100) # Make conformal with CV
53 - #
54 - # Later years (>2015) require full processing starting here. We
55 - # are re-processing level1 files for WY <= 2015 to bring them
56 - # into conformance with current CV and variable names
57 - #
58   wu1000 = subset(wb1000, !grepl('SPL', wb1000$CategoryD))
59   ws1000 = subset(wb1000, grepl('SPL', wb1000$CategoryD))
60 - #
61 - # Water Use
62 - #
63   wu2300 = qc2300(wu1000) # Water Use: Compute AMU, DEP, NMx # TRAs come from here
64   wu2500 = qc2500(wu2300) # Tables: df02 returned with CategoryE
65 - #
66 - # Re-integrate since WS needs WU to complete computations
67 - # but needs to have SPL023-6 included
68 - #
69   wb2000 = rbind(wu2300, ws1000) # After WU adjustments are made
70 - #
71 - # Water Supply: ws2400 Needs complete WB dataset since some values need USE
72 - # as part of the equations (e.g., TDS)
73 - #
74   wb3000 = qc2400(wb2000) # Add water supply equations computed terms
75   ws3000 = subset(wb3000, grepl('SPL', wb3000$CategoryD)) # Avoid double counting WU
76   ws4000 = qc2600(wb3000) # Compute water supply tables and level data
77 - #
78 - # Create level2 files: Re-integrate WS and WU
79 - # Only write-out DAU00 after all the additional parameters are computed
80 - #
81   wb3000 = wb3000[order(wb3000$CategoryD),]
82   # Write Level1 files
83   write.table(wb3000, paste(DATA_LEVEL1,
84                             '/', WY, '/CDNR-wbqc-', WY, '-DAU00-wb100.csv',
85                             sep=''), sep=',', row.names=FALSE)
86 - #
87 - # Write shapefiles by WY and HR.Name values
88 - #
89   rc2000 = sf1000(wb3000)
90 - #
91   rc1000 = dp1000(WY) # Generate Level2 files
92   cat(sprintf('Duration %d: %S.2f minutes\n', WY, difftime(Sys.time(), time1, units='mins')))
93 - }
```

Stand-Alone Application Running on Windows: wbqc (R package)

The screenshot shows the RStudio interface running the `wbqc` package. The console window displays the R version (4.4.3) and the package path. The script in the editor installs the package and sets up data processing for water use (WU) and water quality (WQ) data.

Console Output:

```
R version 4.4.3 (2025-02-28 ucrt) -- "Trophy Case"
Copyright (C) 2025 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```

Script Content (wb100.R):

```
17 rm(list = ls()) # Clear workspace
18 install.packages('./src/Library/wbqc-1.0.zip', repos = NULL, type="bin")
19 library(wbqc)
20 wbenv('')
21 # =====
22 # Run for WY >= 2016
23 # =====
24 #for ( WY in c(2016:2021) ){
25 for ( WY in c(2020:2020) ){
26   time1 = Sys.time()
27   cat(sprintf('wb100: Start: %d at %s\n', WY, format(time1, format = "%
28 #
29   wb0100 = qc0100A(WY) # Ingest and integrate regional data
30   wb1000 = qc1000(wb0100) # Make conformal with CV
31 # =====
32 # Later years (>2015) require full processing starting here. We
33 # are re-processing level1 files for WY <= 2015 to bring them
34 # into conformance with current CV and variable names
35 # =====
36   wu1000 = subset(wb1000, !grepl('SPL', wb1000$CategoryD))
37   ws1000 = subset(wb1000, grepl('SPL', wb1000$CategoryD))
38 # =====
39 # Water Use
40 # =====
41   wu2300 = qc2300(wu1000) # Water Use: Compute AWU, DEP, NWX
42   wu2500 = qc2500(wu2300) # Tables: df02 returned with CategoryE
43 # =====
```

(1) Edit Data in Inflow/Outflow Sheets then
(2) Run thru wbqc

The screenshot displays a Windows desktop environment with several open applications:

- File Explorer:** Open to the 'Project-WaterBalance-CA' folder, showing a list of files and folders. The 'data' folder is selected, showing a list of years (2016-2021) and folders like 'Controlled-Vocabulary', '.DS_Store', and 'CA-DWR-WaterBalance-GS-1000-Spatial-...'. The 'wbqc' folder is also visible in the left sidebar.
- Command Prompt:** Open to the 'Project-WaterBalance-CA' directory. The command 'tree' has been executed, showing the directory structure of the project, including 'data', 'logs', 'shapefiles', 'doc', and 'src' folders.
- RStudio:** Open to the 'Project-WaterBalance-CA' project. The 'Environment' pane shows the 'data' folder selected, with a list of files and folders. The 'Files' pane shows the 'data' folder selected, with a list of files and folders. The 'wbqc' folder is also visible in the left sidebar.

The desktop background is a scenic image of a river and trees. The taskbar at the bottom shows the Windows logo, a search bar, and several application icons. The system tray on the right shows the date and time as 4:44 PM on 5/23/2025.

(1) Detect and Correct Errors (Rinse and Repeat) (2) Send to HQ

The screenshot displays the RStudio interface with the following components:

- Top Bar:** Shows the file path `C:\Users\helly\Desktop\Project-WaterBalance-CA\` and the R version `R 4.4.3`.
- Source Editor:** Contains R code for processing water balance data. The code includes comments and functions for writing level 1 files, generating level 2 files, and processing data for various years (2020, 2021, 2022).
- Console:** Displays the output of the R code, showing the start time (2020 at 2025-05-23 16:47 PDT) and the processing status for various years (2020, 2021, 2022).
- Environment:** Shows the current environment with variables like `wb100`, `qc0100A`, `qc0100B`, `qc0100C`, `qc0100D`, `qc0100E`, `qc0100F`, `qc0100G`, `qc0100H`, `qc0100I`, `qc0100J`, `qc0100K`, `qc0100L`, `qc0100M`, `qc0100N`, `qc0100O`, `qc0100P`, `qc0100Q`, `qc0100R`, `qc0100S`, `qc0100T`, `qc0100U`, `qc0100V`, `qc0100W`, `qc0100X`, `qc0100Y`, `qc0100Z`, `qc0100A`, `qc0100B`, `qc0100C`, `qc0100D`, `qc0100E`, `qc0100F`, `qc0100G`, `qc0100H`, `qc0100I`, `qc0100J`, `qc0100K`, `qc0100L`, `qc0100M`, `qc0100N`, `qc0100O`, `qc0100P`, `qc0100Q`, `qc0100R`, `qc0100S`, `qc0100T`, `qc0100U`, `qc0100V`, `qc0100W`, `qc0100X`, `qc0100Y`, `qc0100Z`.
- Terminal:** Shows the output of the R code, including the start time and the processing status for various years.

```
wb3000 = wb3000[order(wb3000$CategoryD),]
# Write Level 1 files
write.table(wb3000, paste(DATA_LEVEL1,
                           '/', WY, '/CDWR-wbqc-', WY, '-DAUCO-wb100.cs
                           v',
                           sep=''), sep=',', row.names=FALSE)
# Write shapefiles by WY and HR.Name values
# =====
rc2000 = sf1000(wb3000)
# =====
rc1000 = dp1000(WY) # Generate Level 2 files
cat(sprintf('Duration %d: %5.2f minutes\n', WY, difftime(Sys.time
(), timel, units='mins'))))
wb100: Start: 2020 at 2025-05-23 16:47 PDT
qc0100A: Processing 2020 / HR = 01NC
qc0100A: Processing 2020 / HR = 02SF
qc0100A: Processing 2020 / HR = 03CC
qc0100A: Processing 2020 / HR = 04SC
qc0100A: Processing 2020 / HR = 05SR
qc0100A: Processing 2020 / HR = 06SJ
qc0100A: Processing 2020 / HR = 07TL
qc0100A: Processing 2020 / HR = 08NL
qc0100A: Processing 2020 / HR = 09SL
qc0100A: Processing 2020 / HR = 10CR
Errors detected: Check error log and press return to continue ...
```

```
rm(list = ls()) # Clear workspace
install.packages('./src/Library/wbqc_1.0.zip', repos = NULL, type="bin")
library(wbqc)
wbenv('')
# =====
# Run for WY >= 2016
# =====
for (WY in c(2016:2021)){
  for (WY in c(2020:2020)){
    timel = Sys.time()
    cat(sprintf('wb100: Start: %d at %s\n', WY, format(timel, format = "%
    # =====
    wb0100 = qc0100A(WY) # Ingest and integrate regional data
    wb1000 = qc1000(wb0100) # Make conformal with CV
    # =====
    # Later years (>2015) require full processing starting here. We
    # are re-processing level 1 files for WY <= 2015 to bring them
    # into conformance with current CV and variable names
    # =====
    wu1000 = subset(wb1000, !grepl('SPL', wb1000$CategoryD))
    ws1000 = subset(wb1000, grepl('SPL', wb1000$CategoryD))
    # =====
    # Water Use
    # =====
    wu2300 = qc2300(wu1000) # Water Use: Compute AWU, DEP, NWx
    wu2500 = qc2500(wu2300) # Tables: df02 returned with CategoryE
    # =====
  }
}
```

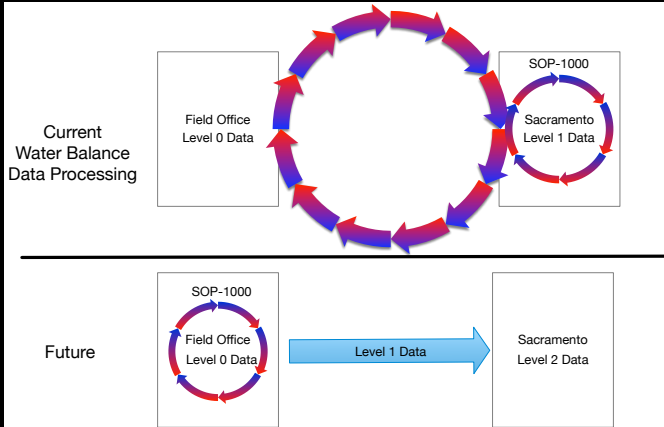

(1) Detect and Correct Errors (Rinse and Repeat) (2) Send to HQ

The screenshot shows a Windows desktop environment. A file explorer window is open, displaying the contents of the 'logs' folder within the 'Project-WaterBalance-CA > data' directory. The file explorer shows a table with columns for Name, Date modified, Type, and Size. The file 'wbqc-wb100-Errors-2020.log' is selected. A Notepad window is open, displaying the contents of this log file. The log file contains the following text:

```
wbqc-wb100-Errors-2020.log - Notepad
File Edit Format View Help
wc0110: Beginning error logging ...
qc0110: c010: Invalid values: special characters ...
Water Supplies 12a State Water Project Deliveries - Agriculture DAU26015 -0.20 07TL 2020
qc0110: c030: Invalid CategoryA value:
```

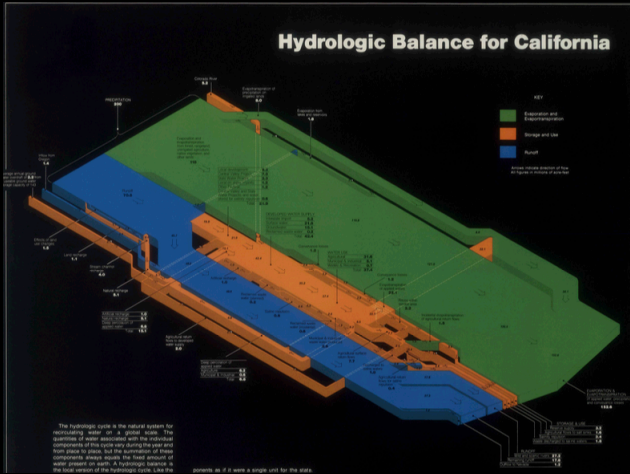
The desktop background is a landscape image. The taskbar at the bottom shows the Start button and several pinned applications, including Microsoft Edge, R 4.4.3, RStudio, and QGIS 3.4.2.0. The system tray at the bottom right shows the date and time as 4:53.

Goals of Water Balance Automation: Less Cost, Better Quality



- Automated, Centralized Level 1 Data Production
- Data Standardization
- < Time for WY Completion
- > Error Detection & Correction
- < Time Level 1 Data Production
- < Time for WY Completion

Atlas of Water Balance: wbqc-Style Methods for Figures, Tables



- Inspired by California Water Atlas, 1979
 - ▶ Based on data from 1972
 - ▶ New digital version could be updated automatically when every new WY is completed
- Catalogue of data, figures and tables for CDWR to draw from
 - ▶ Showcase for new methods of analysis and visualization
 - ▶ Useful for public outreach, research and education
 - ▶ Opportunity for Field Offices to shine
- Developing an outline

- ① Goals of Water Balance Automation
- ② Software and Training
- ③ Extension to Atlas of Water Balance