Photo credit: John Hannon, Reclamation

We Are What We Eat

- Taming Data Management

Jeff Schuyler, Eyasco, Inc.

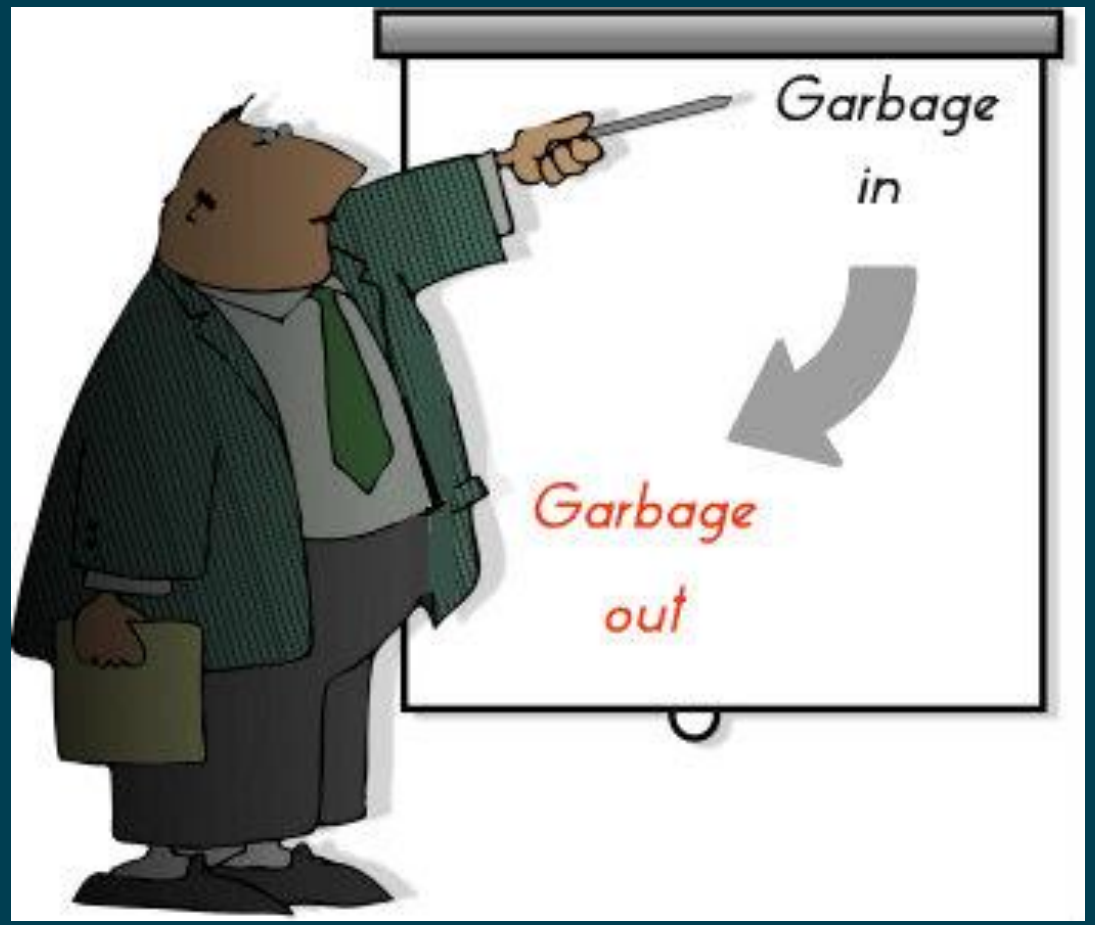
Mike Deas, PhD, PE, Watercourse Engineering, Inc.

Purpose of Data Management System (DMS)

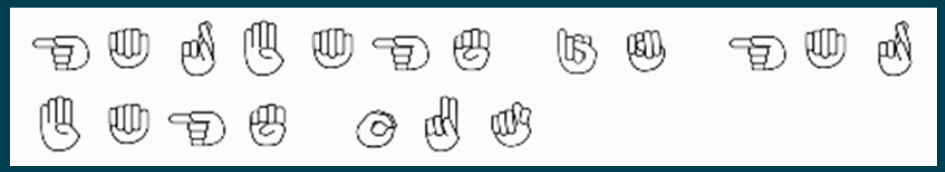
- **Goal:** Provide higher quality data for modelers to confidently and effectively:
  - Assess data
  - Develop models
  - Apply models

- **Intended Outcome:**
  - Produce more reliable data management to support decision making

GIGO – No Matter How You Say It



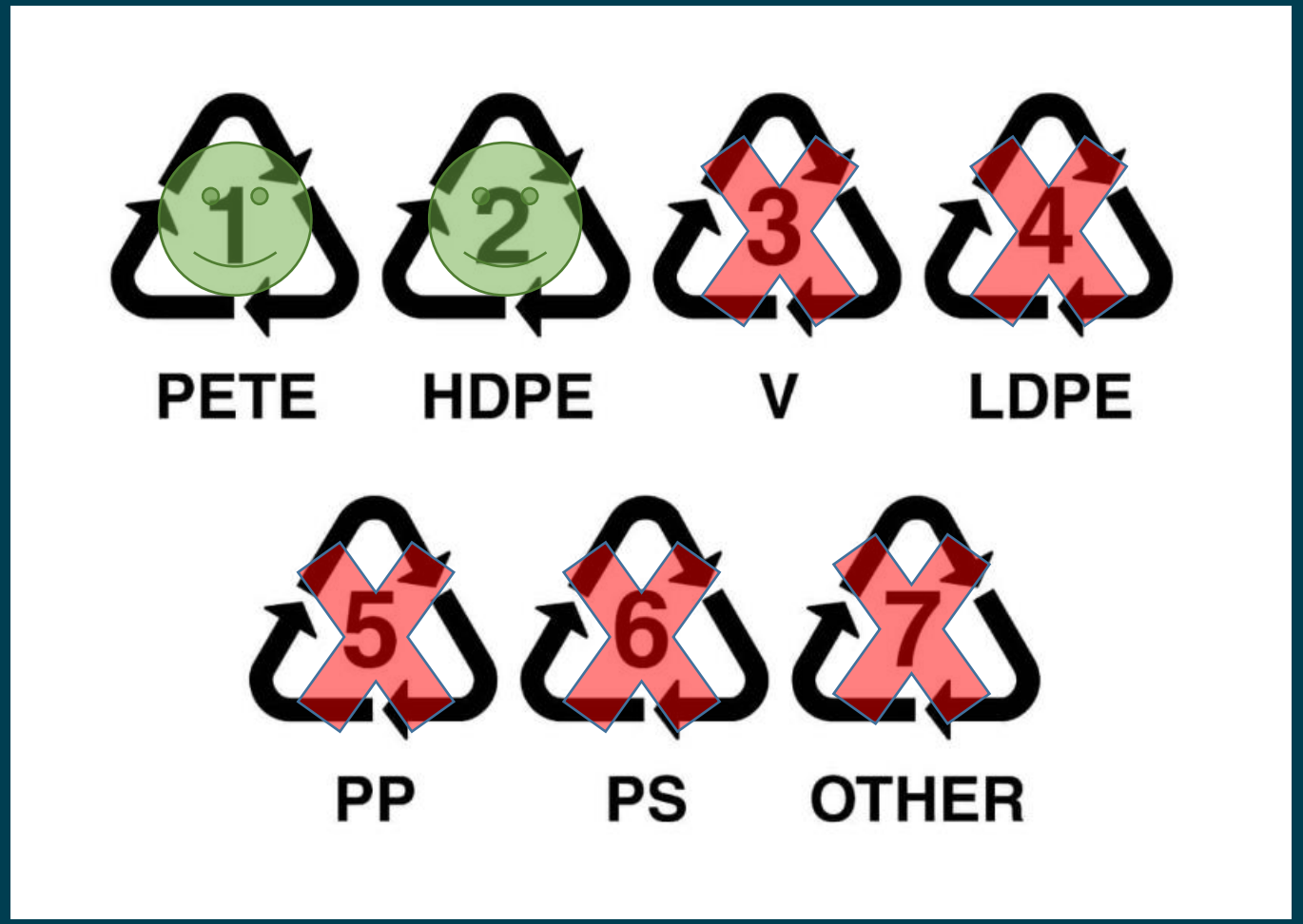
القمامة في ، القمامة خارج	Arabic
basura dentro basura fuera	Spanish
garbage in, garbage out	English
कचरा अंदर कचरा बाहर	Hindi
աղբը, աղբը	Armenian
sampah masuk sampah keluar	Indonesian
ゴミ出し、ゴミ出し	Japanese
쓰레기통, 쓰레기통	Korean
syf na wejściu, syf na wyjściu	Polish
мусор в, мусор в	Russian
skräp in skräp ut	Swedish
குப்பை உள்ளே குப்பை வெளியே	Tamil
מיסט אין, מיסט אויס	Yiddish
垃圾进垃圾出	Chinese



Impacts of Poor-Quality Data

- Poor-Quality Data
  - Incomplete
  - Inaccurate
  - Inconsistent
  - Invalid
  - Redundant
  - Non-standard
- Impacts
  - Inefficiency/lost time
  - Added costs
  - Missed opportunities
  - Lack of confidence
  - Poor quality analysis and model results

GIGO – But What is in the Garbage Really?



GIGO – But What is in the Garbage Really?

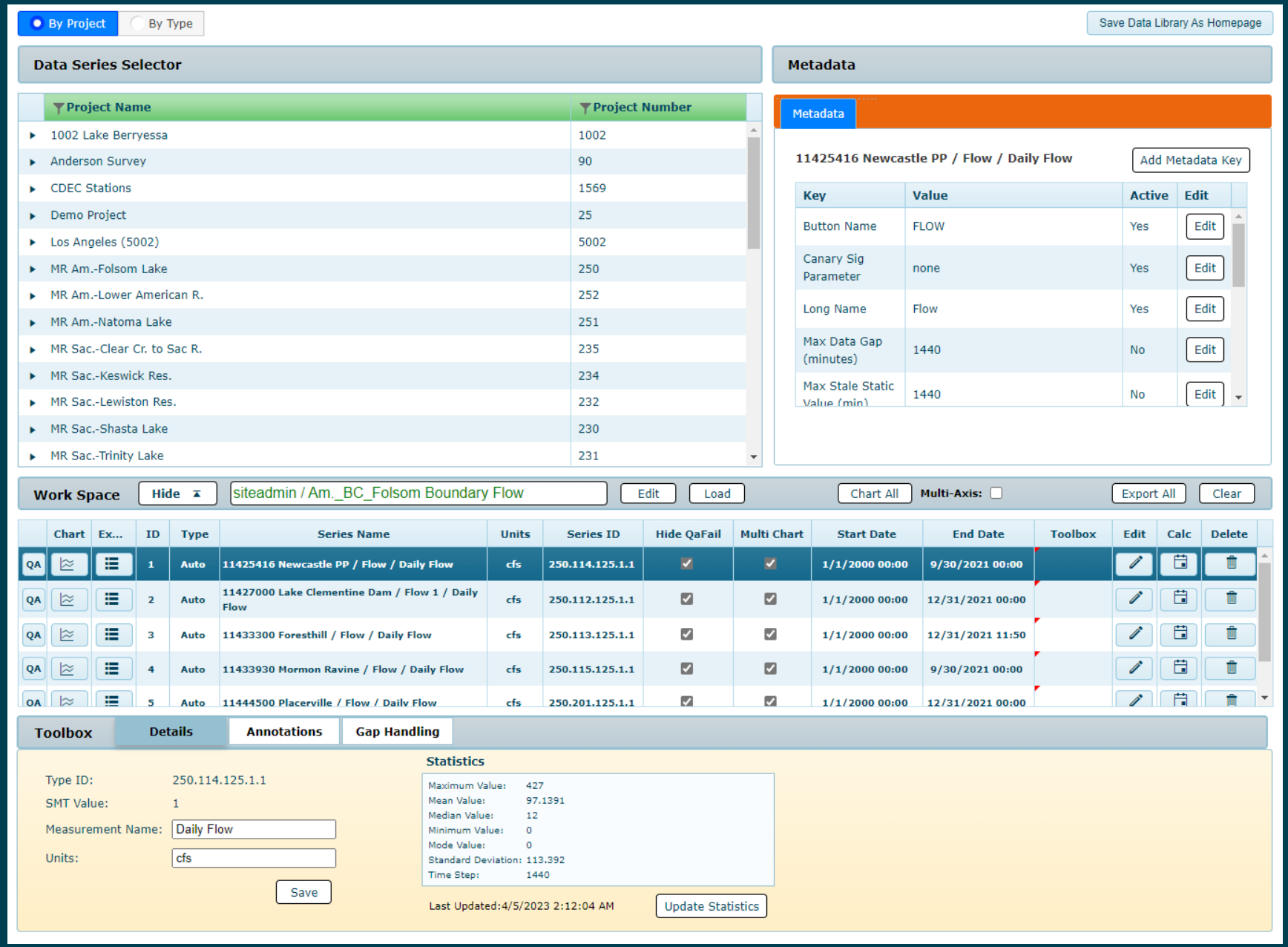
	A	B	C	D	E	F	G	H	I
1	Datetime (PST)	Raw Data	Raw Quality Code	Gap Filled Data	Gap Filled Quality Code	Quality Code Stn Type ID	Corrected Data	Corrected Quality Code	Quality Code Stn Type ID
2	07/26/2004 02:26 PM	340.8	Provisional				345.8	Corrected	18
3	07/27/2004 02:29 PM	340.8	Provisional				345.8	Corrected	18
4	07/28/2004 02:20 PM	340.8	Provisional				345.8	Corrected	18
5	07/29/2004 02:16 PM	336.8	Provisional				340.8	Corrected	18
6	07/30/2004 02:20 PM			336.8	Gap Filled	18			
7	07/31/2004 02:31 PM			336.8	Gap Filled	18			
8	08/01/2004 02:25 PM			336.8	Gap Filled	18			
9	08/02/2004 02:30 PM			336.8	Gap Filled	18			
10	08/03/2004 02:25 PM			336.8	Gap Filled	18			
11	08/04/2004 02:33 PM	333.9	Provisional						
12	08/05/2004 02:25 PM	333.9	Provisional						
13	08/06/2004 02:25 PM	333.9	Provisional						
14	08/07/2004 02:25 PM	333.9	Provisional						
15	08/08/2004 02:32 PM	333.9	Provisional						
16	08/09/2004 02:30 PM	333.9	Provisional						
17	08/10/2004 02:30 PM	333.9	Provisional						
18	08/11/2004 02:29 PM	333.9	Approved						
19	08/12/2004 02:26 PM	333.9	Approved						
20	08/13/2004 02:30 PM	333.9	Approved						
21	08/14/2004 02:30 PM	333.9	Approved						
22	08/15/2004 02:25 PM			333.3	Gap Filled	18			
23	08/16/2004 02:30 PM			333.3	Gap Filled	18			
24	08/17/2004 02:25 PM			333.3	Gap Filled	18			
25	08/18/2004 02:25 PM			333.3	Gap Filled	18			
26	08/19/2004 02:25 PM			333.3	Gap Filled	18			
27	08/20/2004 12:00 AM			333.3	Gap Filled	18			
28	08/21/2004 02:25 PM	333.3	Approved						
29	08/22/2004 02:25 PM	333.3	Approved						
30	08/23/2004 02:40 PM	333.3	Approved						
31	08/24/2004 02:18 PM	333.3	Approved						
32	08/25/2004 02:36 PM	333.3	Approved						
33	08/26/2004 02:32 PM	333.3	Approved						

What is in the DMS for WTMP?

- Functions of a DMS with database structure
  - Store and organize data (rules-based organization)
  - Streamline data collection
  - Visualize data
  - Create and track metadata
  - Be able to adapt
- WTMP key topics
  - Quality Codes – metadata
  - Model Ready Data – common formats, no gaps
  - Web Access – automated data communication

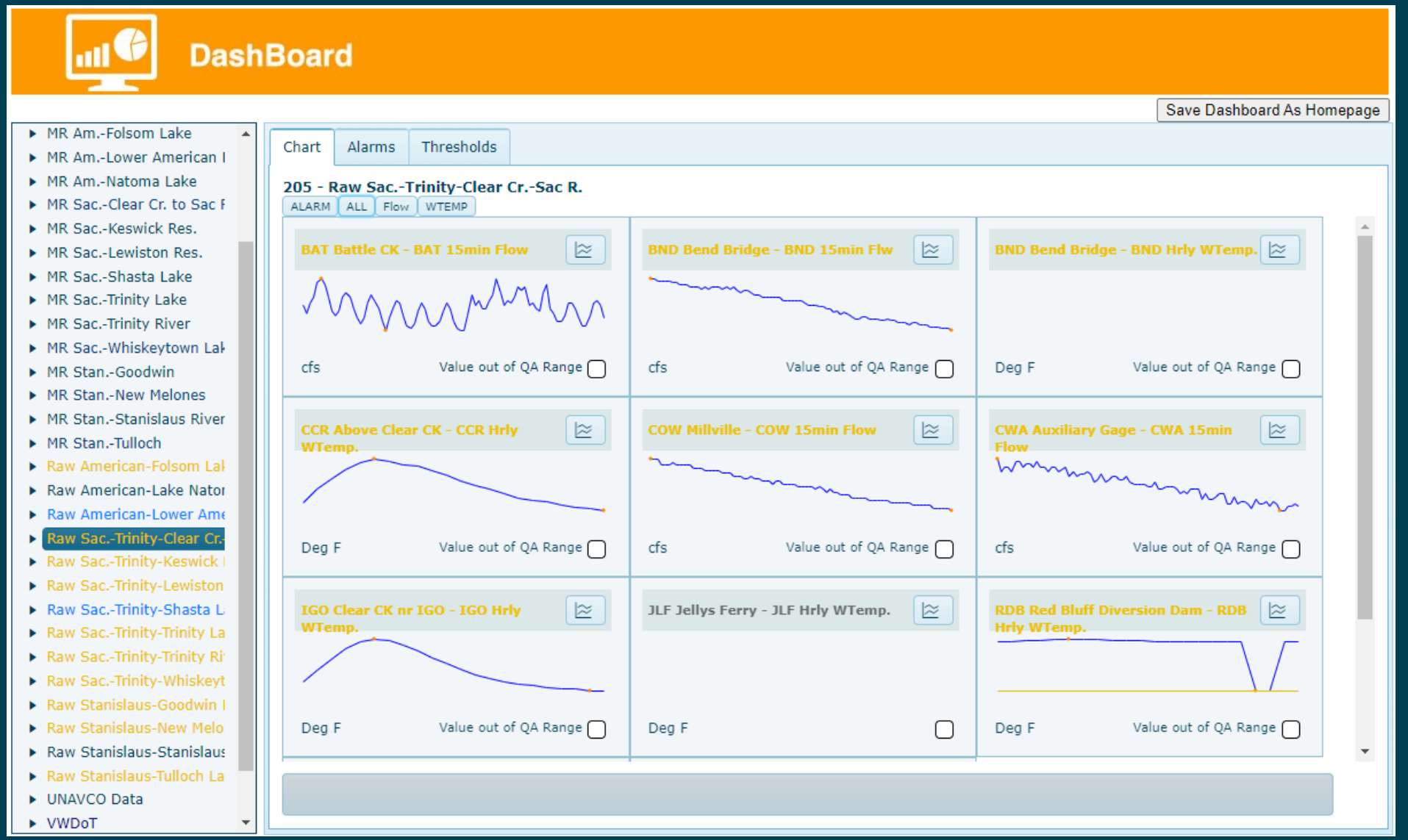
DMS - Store and Organize Data in Data Library

- Time series organized by “Project” that equates to Model Domain.
- Includes metadata and visualization and processing tools.

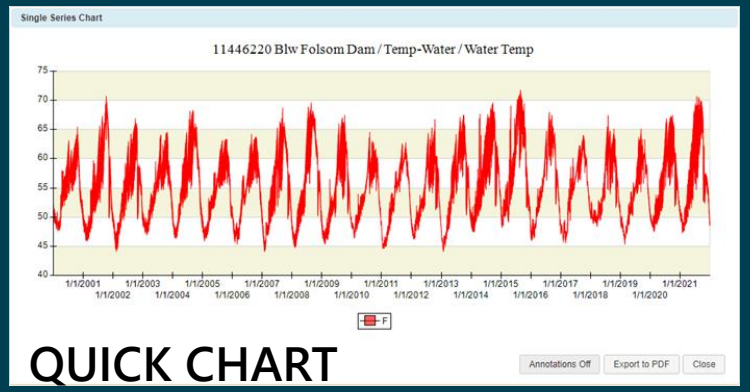


DMS – Streamline Data Collection Shown in the Dashboard

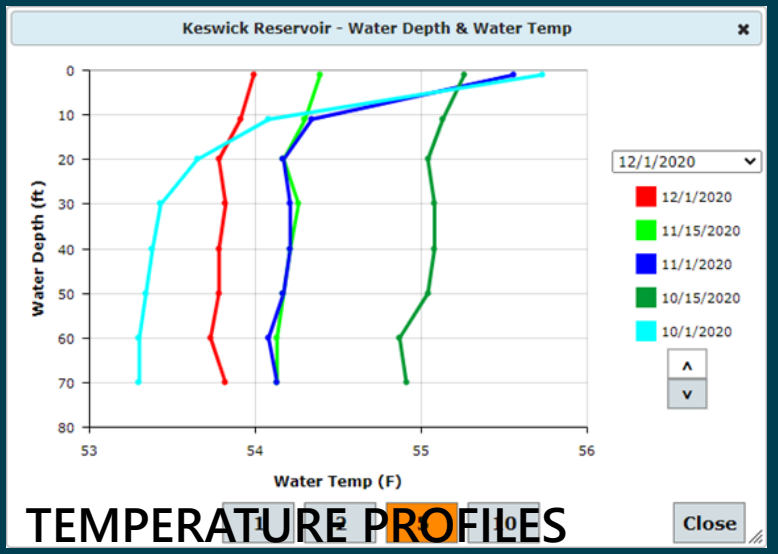
- Most time series data collected in near real time from on-line sources.
- Data is imported with native quality codes if they exist.
- Thresholds can be applied to filter out ‘bad’ data.



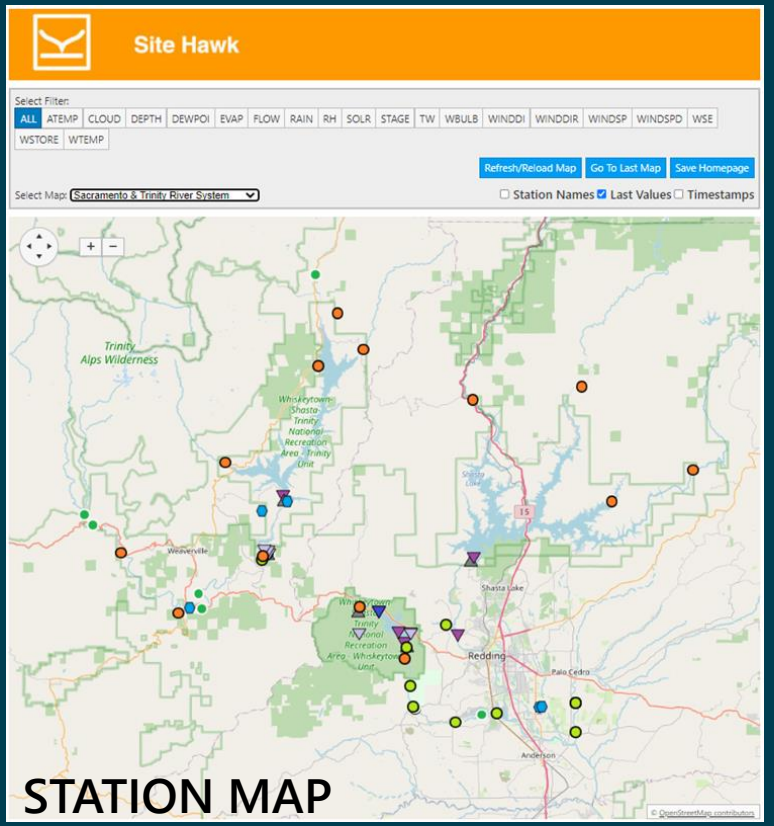
DMS – Visualize Data



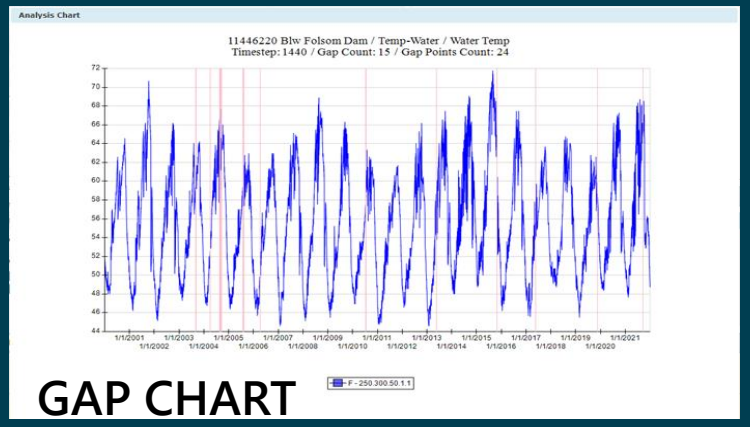
QUICK CHART



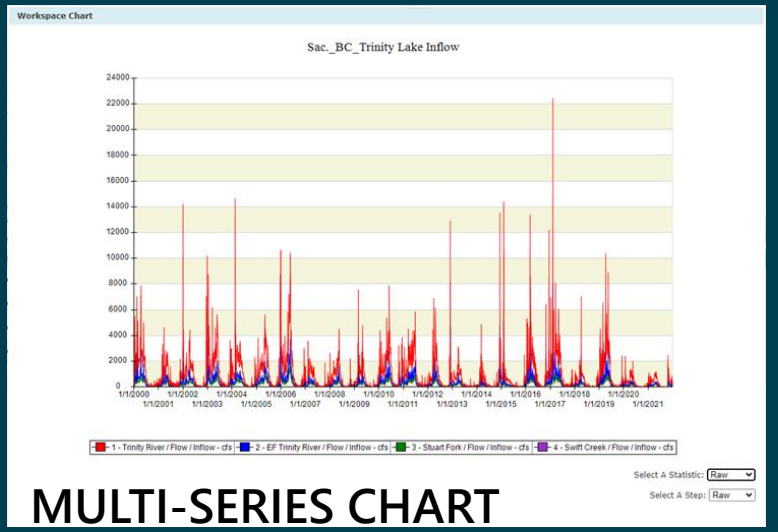
TEMPERATURE PROFILES



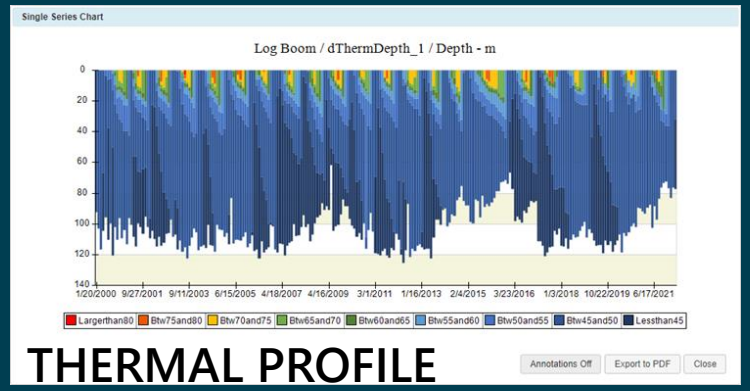
STATION MAP



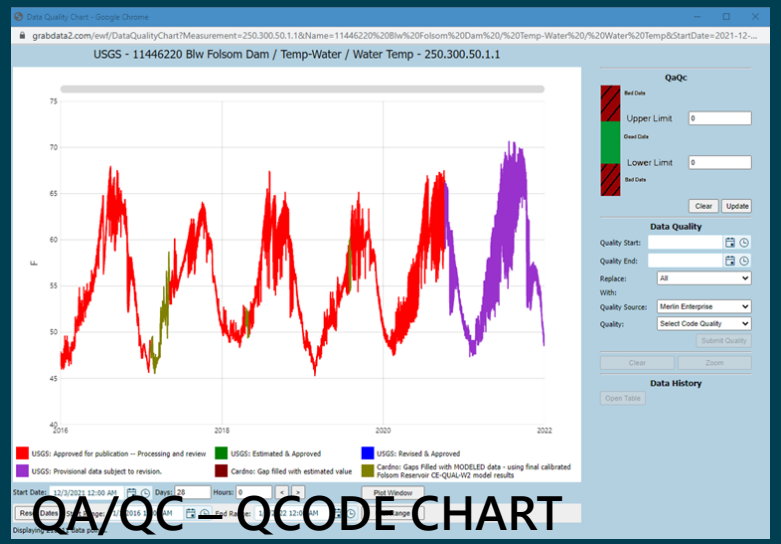
GAP CHART



MULTI-SERIES CHART



THERMAL PROFILE



QA/QC – QCODE CHART

DMS - Create and Track Metadata

MR Am.-Folsom Lake Metadata List

4/5/2023 10:25:50 AM

Name	SHEF	Station Type	Source Website	Data Type	Parameter	Unit	Button	Series ID	Latitude	Longitude	Timestep	Start Date	End Date
11425416 Newcastle PP Daily Flow	n/a	USGS	USGS	BC	Flow	cfs	FLOW	250.114.125.1.1	38.835	-121.091944	1440	2000-01-01 00:00:00	2021-09-30 00:00:00
11427000 Lake Clementine Dam Daily Flow	n/a	USGS	USGS	BC	Flow	cfs	FLOW	250.112.125.1.1	38.93600919	-121.0227778	1440	2000-01-01 00:00:00	2021-12-31 00:00:00
11427000 Lake Clementine Dam 15 min Flow	n/a	USGS	USGS	BC	Flow	cfs	FLOW	250.112.125.2.1	38.93600919	-121.0227778	15	2000-01-01 00:00:00	2021-12-31 23:15:00
11433300 Foresthill Daily Flow	n/a	USGS	USGS	BC	Flow	cfs	FLOW	250.113.125.1.1	39.0060108	-120.760763	15, 1440	2000-01-01 00:00:00	2021-12-31 11:50:00
11433790 Auburn Dam Water Temp	n/a	USGS	USGS	BC	Temperature	F	WTEMP	250.111.50.1.1	38.8829546	-121.0630007	1440, 15	2000-01-01 00:00:00	2021-12-31 23:45:00
11433930 Mormon Ravine Daily Flow	n/a	USGS	USGS	BC	Flow	cfs	FLOW	250.115.125.1.1	38.836667	-121.093333	1440	2000-01-01 00:00:00	2021-09-30 00:00:00
11444500 Placerville Daily Flow	n/a	USGS	USGS	BC	Flow	cfs	FLOW	250.201.125.1.1	38.7710133	-120.8163242	1440	2000-01-01 00:00:00	2021-12-31 00:00:00
11446030 Pilot Hill Water Temp	n/a	USGS	USGS	BC	Temperature	F	WTEMP	250.200.50.1.1	38.76295708	-121.0082763	15	1999-12-31 12:00:00	2021-12-31 23:45:00
11446220 Blw Folsom Dam Water Temp	n/a	USGS	USGS	BC	Temperature	F	WTEMP	250.300.50.1.1	38.70452778	-121.1643611	15	2000-01-01 00:00:00	2021-12-31 23:45:00
Dyke 8 Wind Speed	n/a	NWS	CDEC	BC	Wind Speed	m/s	WINDSPD	250.14.133.1.1	38.693054	-121.129723	60	2000-01-01 01:00:00	2022-01-01 00:00:00
Dyke 8 Wind Direction	n/a	NWS	CDEC	BC	Wind Direction	deg	WINDDIR	250.14.133.1.2	38.693054	-121.129723	60	2000-01-01 01:00:00	2022-01-01 00:00:00
EID Folsom Diversion Diversion Flow	n/a	EID	EID	BC	Flow	cfs	FLOW	250.6.125.1.1	0	0	60, 1440	2000-01-01 00:00:00	2021-12-31 00:00:00
FOL Evaporation	ES	USBR	USBR	BC	Evaporation	cfs	EVAP	250.3.155.1.1	38.683	-121.183	1440	1999-10-01 00:00:00	2022-01-19 00:00:00
FOL Elevation	HL	USBR	USBR	BC	Elevation	ft	WSE	250.3.145.1.1	38.683	-121.183	60	1999-10-31 22:00:00	2022-01-19 00:00:00
FOL Reservoir Storage	LS	USBR	USBR	BC	Water Storage	ac-ft	WSTORE	250.3.165.1.1	38.683	-121.183	60	1999-12-22 13:00:00	2022-01-19 00:00:00
FOL Shutter Pos U1	n/a	USBR	USBR	BC	Level	ft	GATE	250.3.145.11.1	38.683	-121.183	1440	2000-01-01 00:00:00	2021-12-31 00:00:00
FOL Shutter Pos U2	n/a	USBR	USBR	BC	Level	ft	GATE	250.3.145.12.1	38.683	-121.183	1440	2000-01-01 00:00:00	2021-12-31 00:00:00
FOL Shutter Pos U3	n/a	USBR	USBR	BC	Level	ft	GATE	250.3.145.13.1	38.683	-121.183	1440	2000-01-01 00:00:00	2021-12-31 00:00:00
FOL MI Gate Elevation	n/a	USBR	USBR	BC	Level	ft	GATE	250.3.145.2.1	38.683	-121.183	1440	2000-01-01 00:00:00	2021-10-19 13:14:00
FOL MI WTemp	n/a	USBR	USBR	Cal/Val	Temperature	F	WTEMP	250.3.50.4.1	38.683	-121.183	1440, 1435, 1445	2004-05-08 14:25:00	2021-10-20 17:55:00

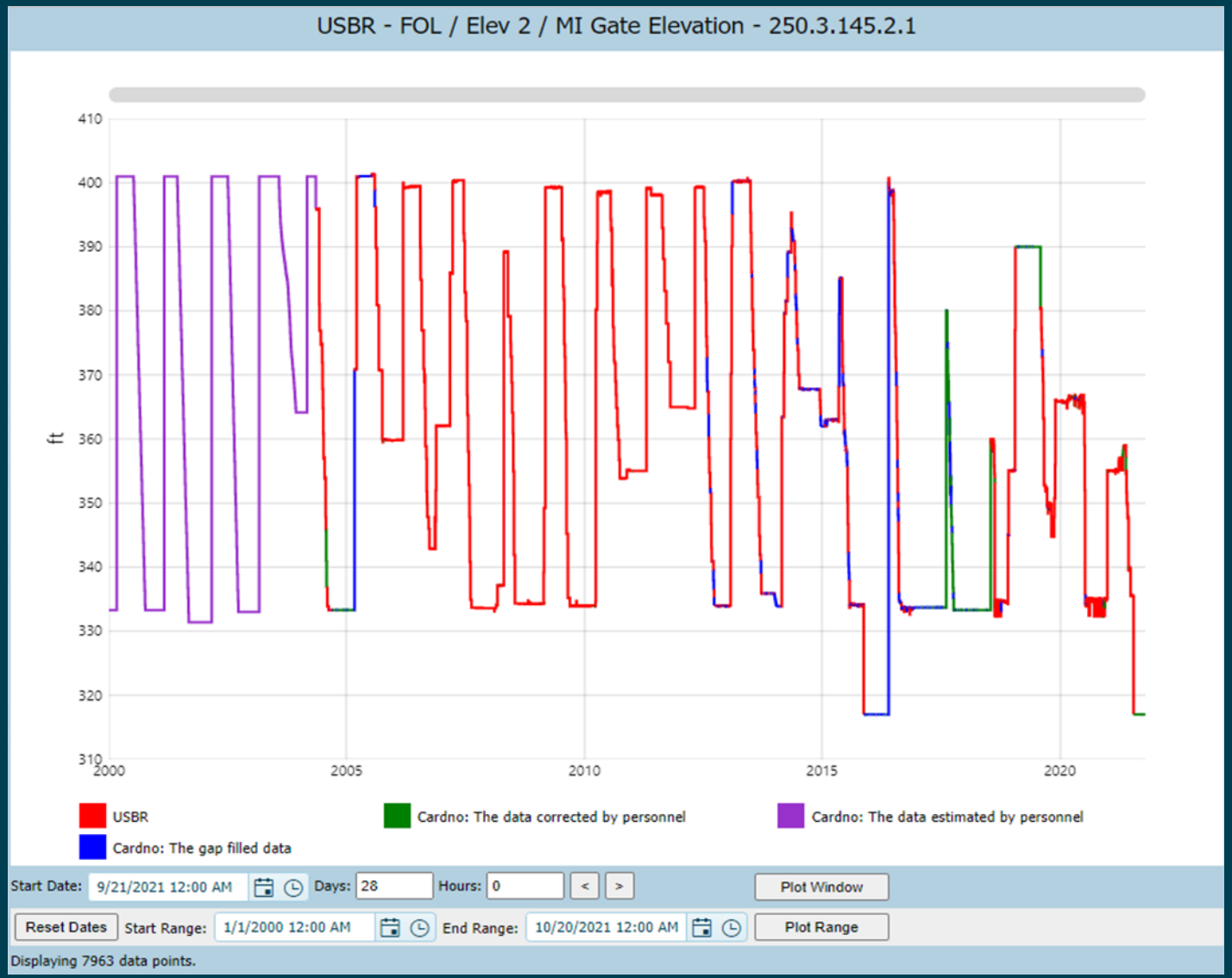
DMS - Create and Track Quality Codes

USGS		
Approved by Source	A	Approved for publication -- Processing and review
	A:[91]	Approved: Daily mean calculated from data on this day matches published daily mean within 1 percent
	A:[92]	Approved: Daily mean calculated from data on this day matches published daily mean within 5 percent
	A:e	Estimated & Approved
	A:R	Revised & Approved
No QCode	0	No data quality code
Provisional	P	Provisional data subject to revision.
	P:[4]	Statistic computed from less than expected number of instantaneous values for the period
	P:e	Estimated & Provisional
Questionable	<	The Value is known to be less than reported value.
	>	The value is known to be greater than reported val
Revised by Source	e	The value has been edited or estimated by USGS personnel

Cardno		
Approved	A_Cardno	Approved data --- Cardno
Approved by Source	Approved	Approved
Calculated	Calc	Calculated data
Estimated	Estimated	The data estimated by personnel --- Cardno
Gap Filled	999	Gap filled with estimated value
	CBR	Gap filled with CDEC CBR daily data --- Cardno
	D8	Gap filled with CDEC Dyke 8 (FLD) --- Cardno
	D8 Reg	Gap filled with CDEC Dyke 8 (FLD) regression data --- Cardno
	FO	Gap filled with CIMIS Fair Oaks Wind Direction --- Cardno
	FO Reg	Gap filled with CIMIS Fair Oaks Wind Speed regression --- Cardno
	Gap Filled	The gap filled data --- Cardno
	Interp	Interpolated & Approved --- Cardno
	Interpolated	Interpolated & Approved
	Last Good Value	Gap filled with last good value --- Cardno
	Modeled	Gaps Filled - using final calibrated Folsom Reservoir CE-QUAL-W2 model results
	NF Daily Regression	NF Daily Regression & Approved
	O	Out of Service
	Previous	Gap filled with data from previous year, same day and time
	R-11 15min	Gap filled with PCWA R-11 15min data --- Cardno
	R-11 Daily	Gap filled with PCWA R-11 daily data --- Cardno
	SF Regression	SF Regression & Approved
	USGS Daily Data	USGS Daily Data & Approved
Modified	Corrected	The data corrected by personnel --- Cardno
	D8 Cor	The data corrected with CDEC Dyke 8 (FLD) data --- Cardno
	Previous Cor	Corrected with data from previous year, same day and time
No QCode	0	No data quality code
Provisional	Provisional	Provisional

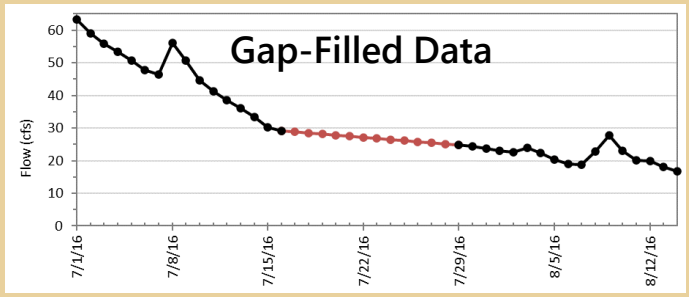
DMS - Create and Track Quality Codes

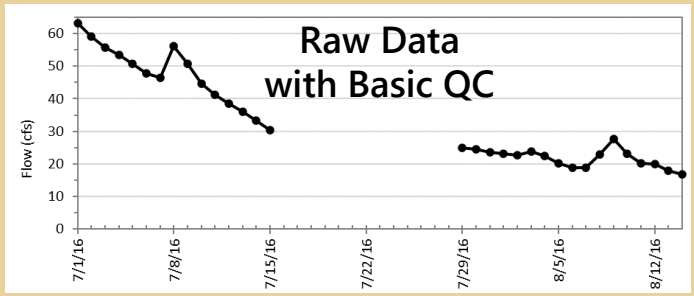
- Quality codes assigned to data set allow tracking of data history
- Allows development of “Model Ready Data”
- Model Ready Data
  - Complete boundary conditions (no gaps) for use in the WTMP models



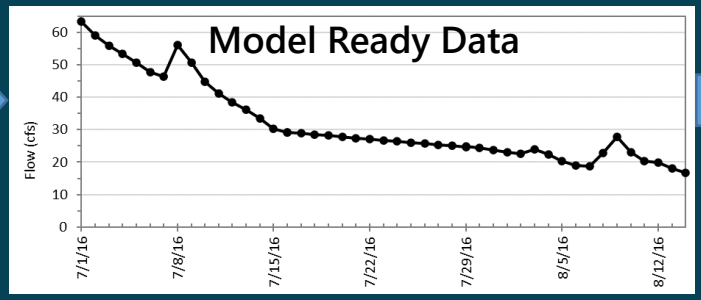
DMS - Provide Model Ready Data and Calibration/Validation Data

DMS





Web Service To WTMP



WTMP

- Build Model or Apply Model (e.g., Forecast)
- Test Model (e.g., Calibration/ Validation)

DMS - Summary

DMS

- **Data Repository (Library)**
  - **Historical Data** — 2000-2021 model ready data (with QA flags)
  - **Appended Data** — Post-2021 data
- **Data Collection Operation**

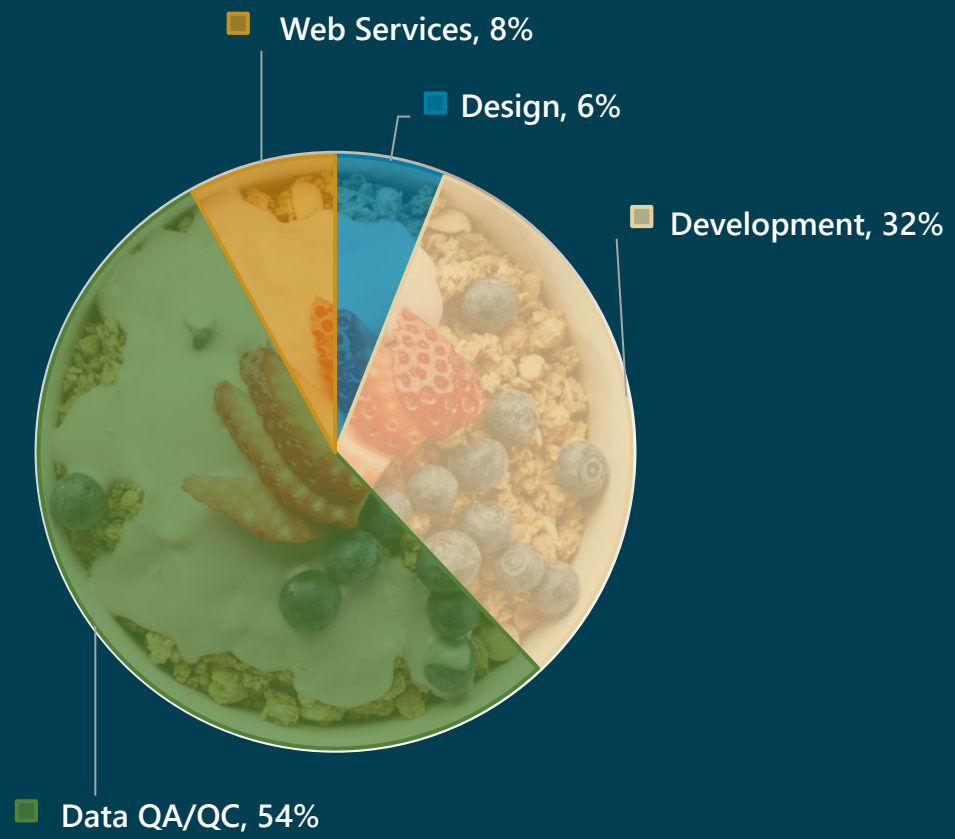
  Data collection at specified intervals (e.g., hourly, every 4 hours, every 6 hours, daily, weekly)
  - Basic data management
  - Automated status report generation to keep data managers updated
- **Data Analysis**

  Improved data access and expanded data analysis
  - Data analysis
  - Model development and application
- **Model Data**
  - **Calibration**
  - **Model Ready**

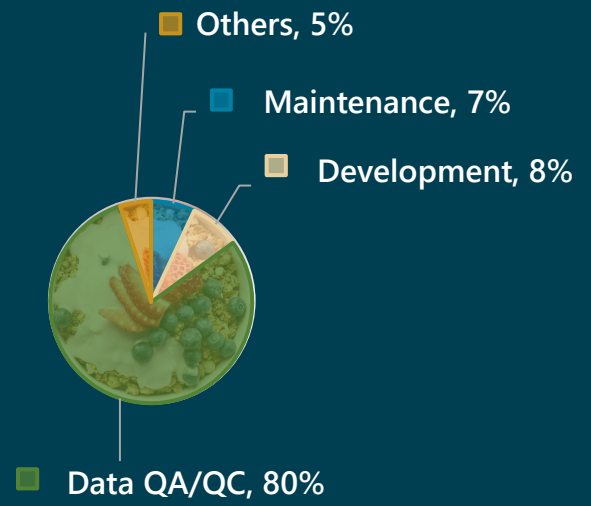
  Robust data management and QA procedures provide good quality data to WTMP
  - Common calibration and boundary condition data
  - Combined with the WTMP reporting, efficient and effective communication of data and model output to data managers and modelers

DMS - Current Investment Will Pay Dividends

Costs for Initial DMS Implementation (NOW with 20 Years of Data)



Costs for Continuous DMS Implementation (Future with Yearly Data)



Future cost of continuous implementation is substantially less due to largely diminished costs for design, web services, and development. Data QA/QC becomes the dominating cost item.

(Not to Scale – For Illustration Only)