USGS
science for a changing world

Preserving Water-Quality Records from the Early 20th Century

Robert Swanson[1], Robert Baskin[2], and Melanie Clark[3]

[1]USGS Nebraska Water Science Center

[2]USGS Utah Water Science Center

[3]USGS Wyoming-Montana Water Science Center

9th National Monitoring Conference, Cincinnati, Ohio, April 29, 2014

U.S. Department of the Interior
U.S. Geological Survey

Background - 1903-1973

1903 - U.S. Geological Survey (USGS) first laboratory opened in Washington, DC

1970 - 22 regional water-quality labs

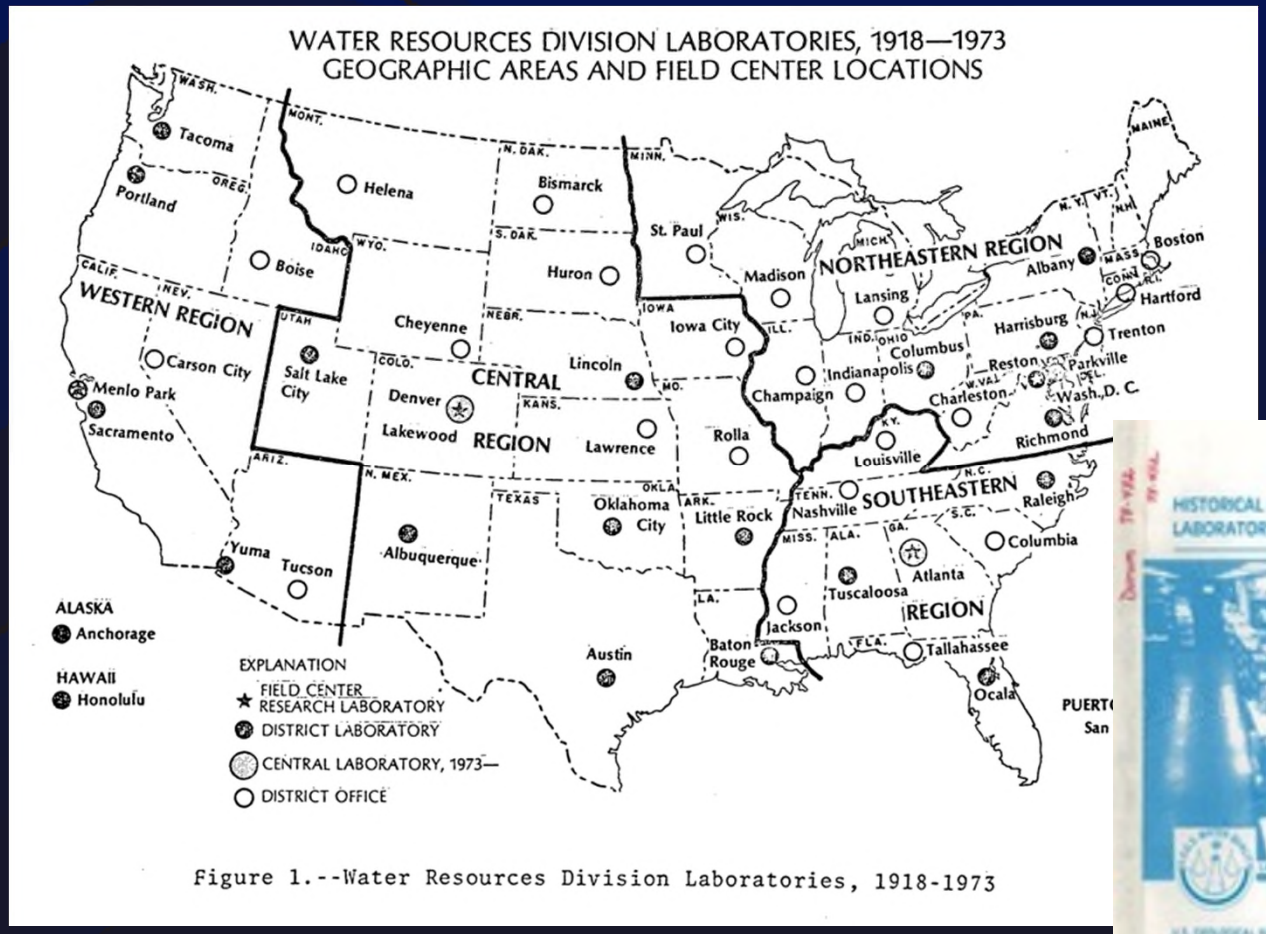
1972 – Consolidation of laboratory facilities

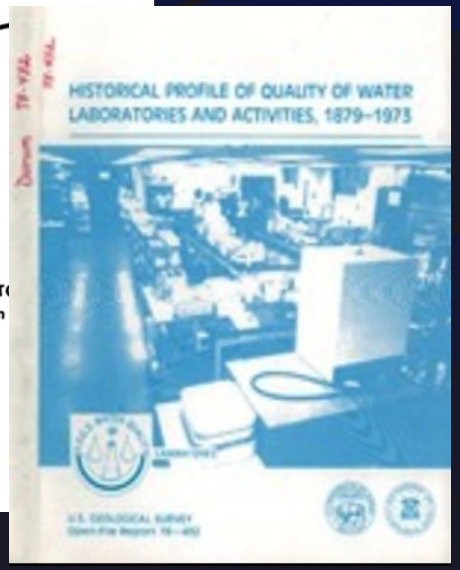
1918 to 1928 - 480 samples/year

1946 - 7,800 samples

1970 - 50,000 samples

Locations of Laboratories





From… Durum, W.H., 1978, Historical profile of quality of water laboratories and activities, 1879-1973: USGS OFR 78-432

Background - 1918-1973

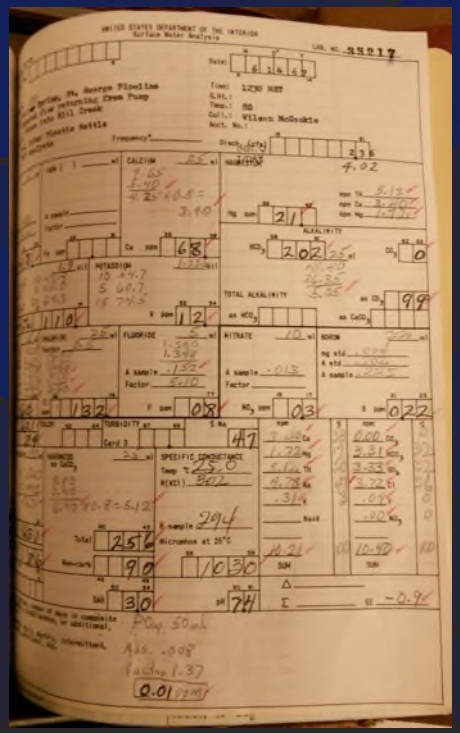
- Paper records were checked and results uploaded into the USGS Water Storage and Retrieval System (WATSTORE)
- WATSTORE migrated to the National Water Information System (NWIS) in 1972
- Laboratory original records were transferred to ????

Background - 1918-1973

- Most of the available metadata available for the analysis was not stored
  - Holding times
  - Composite samples
  - Analytical methods
  - Minimum reporting levels
  - Accuracy and bias
  - Sampling method

Background - 1918-1973

- Legacy water-quality data represented the “best available” methods of the time
  - Used current accepted methods
  - Documented processes
- Original analyses not stored with the Centers maintaining the data



Early Water-Quality Analyses

When legacy analyses results were loaded into WATSTORE by hand…

- Some results set to public access
  - ✓ Varying levels of scrutiny
  - ✓ Quality of data checking unknown
- Some results set to internal access only

Early Water-Quality Analyses

Why the difference?

Was it based on the quality of the analyses or left to an individual reviewer interpretation?

USGS wants a consistent and logical resolution

Problem

Provenance of the pre-1970 water-quality data is not in the database

Important for long-term trend analysis, in understanding changes in natural water systems, and the effects of differing analysis methods on the quality of the analyses

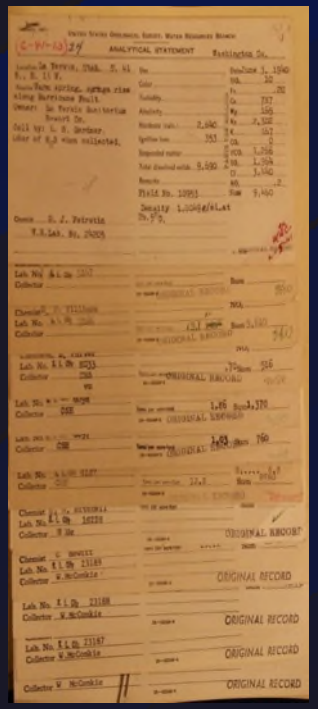
Issue

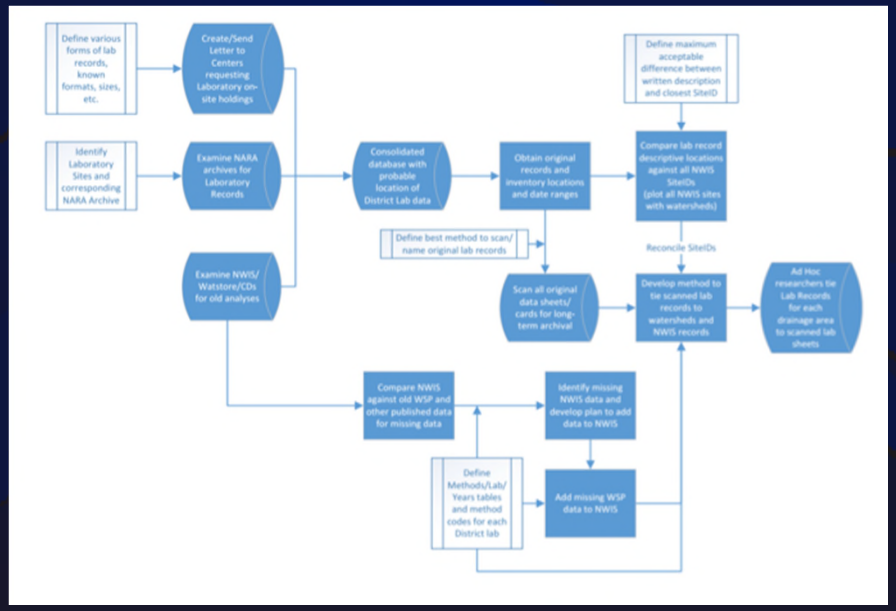
Q. How do we reconnect the data in the original records with methods documentation to the database so that researchers have the advantage of knowing the provenance of the analyses?

Goal

“Water Quality Records Rescue Made Easy”

(Unfortunately, it is not as easy as it sounds)





Nebraska – Lincoln (Missouri River Basin) Laboratory

~18,000 records 1910 -1971 for Nebraska alone

These were set with flag ‘i’ meaning Internal Only and not available to public

Location of the majority of original analyses/records unknown

Mostly dissolved ions, sediment, and nutrients

Discovered some analytical records not in NWIS

Utah – Salt Lake City Laboratory

>63,000 records available through 1971

All? original records have been located

Majority were archived in Denver Federal Records Center - some still in Center archive room

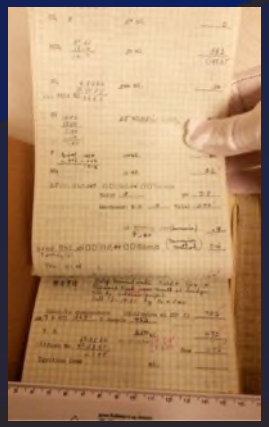
Other States - in discovery process

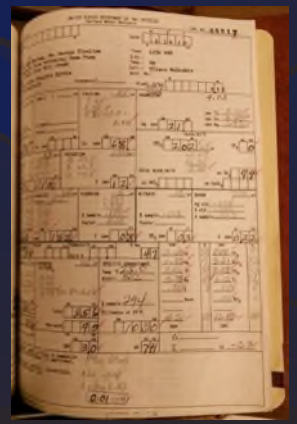
USGS Water Quality Rescue

Utah – Salt Lake City Laboratory

Records in a variety of formats and organization (LABNO, STAID, BASIN, YEAR)

Analyses from AK, AZ, CA, CO, ID, NV, OR,UT, WA, WY, ??

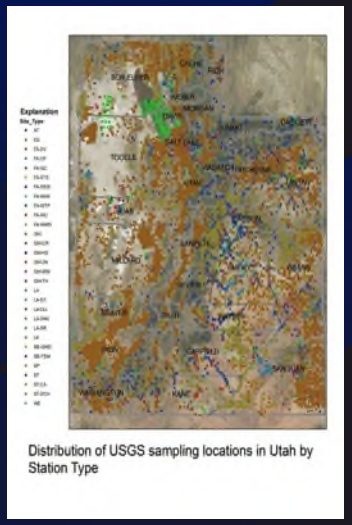


One Step at a Time...

- Step One: Find the original records
  - National Archives and Records Administration (NARA) Search
    - Determine records location
    - Recover records for inventory
    - Scan records for preservation
    - Logically organize all records
    - Scanned records available through WWW

One Step at a Time...

- Step Two: Georeference the sites
  - Link scanned records to NWIS STAIDs
    - Historical records predate USGS Site ID assignment
    - Identify locations by description
    - Compare locations in GIS to identify matches or probable matches
    - Assign missing STAIDs



Distribution of USGS sampling locations in Utah by Station Type

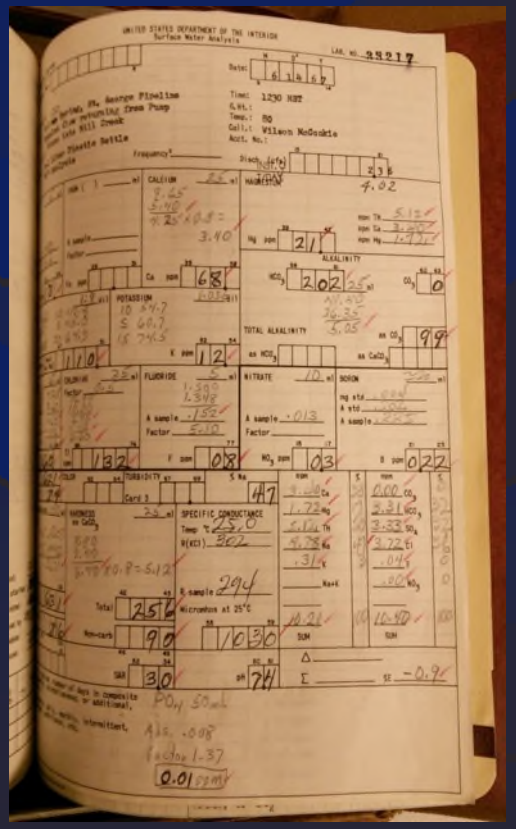
One Step at a Time...

- Step Three: Check the data and compare records to NWIS
  - Automate as much as possible
  - Determine what is correct and fix the rest

One Step at a Time...

Develop water-quality tools to examine NWIS database

- Determine what can be released
  - Filter for other types (sediment)
  - Filter for typograpghic errors ;-)
- Reconcile the remaining analyses using current USGS “rules”
  - Calculated values
  - “Zero” values
- Automate process using scripts



Status and Future Efforts

Complete test procedures on Salt Lake City laboratory records

- Laboratory original records located and retrieved
- Basic inventory completed
- Scan quality test on records completed
- Records numbering systems identified
- Scan control cards being generated for directories
- Records being packaged and sent out for scanning

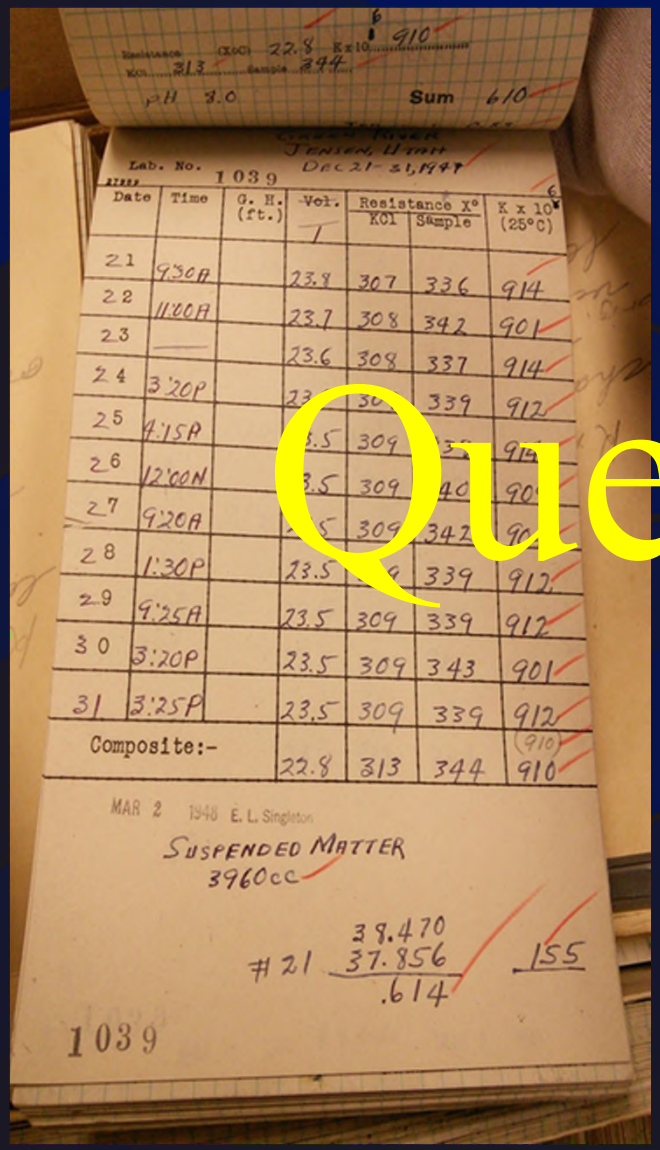
Status and Future Efforts

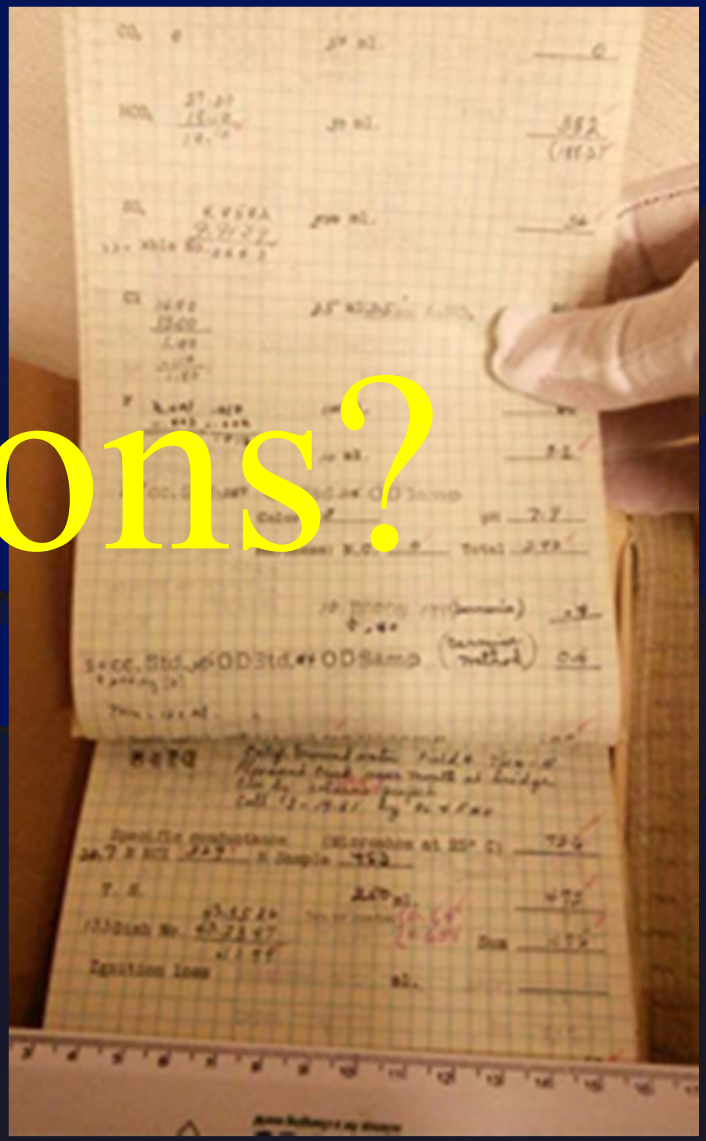
- Locate remaining original records for Lincoln and Salt Lake City laboratories
- Develop tools to assist in checking and classifying Lincoln laboratory records and test on Nebraska NWIS database
- Develop list of analytical methods used for each laboratory and timeframes for use in quantifying analysis quality

Status and Future Efforts

- Link all scanned laboratory original records to Station IDs
- Assign data integrity codes
- Expand the process to all original 22+ laboratories

- Repeat as necessary…


Lab. No. 1039
Jensen, Utah
Dec 21-31, 1947
Date | Time | G. H. (ft.) | Vol. | Resistance X° KCl | Resistance X° Sample | K x 10⁶ (25°C)
21 | 9:30A | 23.8 | 307 | 336 | 914
22 | 11:00A | 23.7 | 308 | 342 | 901
23 | | 23.6 | 308 | 337 | 914
24 | 3:20P | 339 | 912
25 | 4:15A | 309
26 | 12:00N | 309
27 | 9:20A | 309 | 342
28 | 1:30P | 23.5 | 339 | 912
29 | 9:25A | 23.5 | 309 | 339 | 912
30 | 3:20P | 23.5 | 309 | 343 | 901
31 | 3:25P | 23.5 | 309 | 339 | 912
Composite:- | 22.8 | 313 | 344 | 910
MAR 2 1948 E. L. Singleton
SUSPENDED MATTER
3960cc
#21 39.470 37.856 .614 155
1039



Questions?