Metadata Recovery as an Important Step in Trusting and Using Historic Water Quality Data

Presented by Ariel Reed
with acknowledgements to Donna Myers, Ted Stets, and Rob Baskin, Office of Water Quality

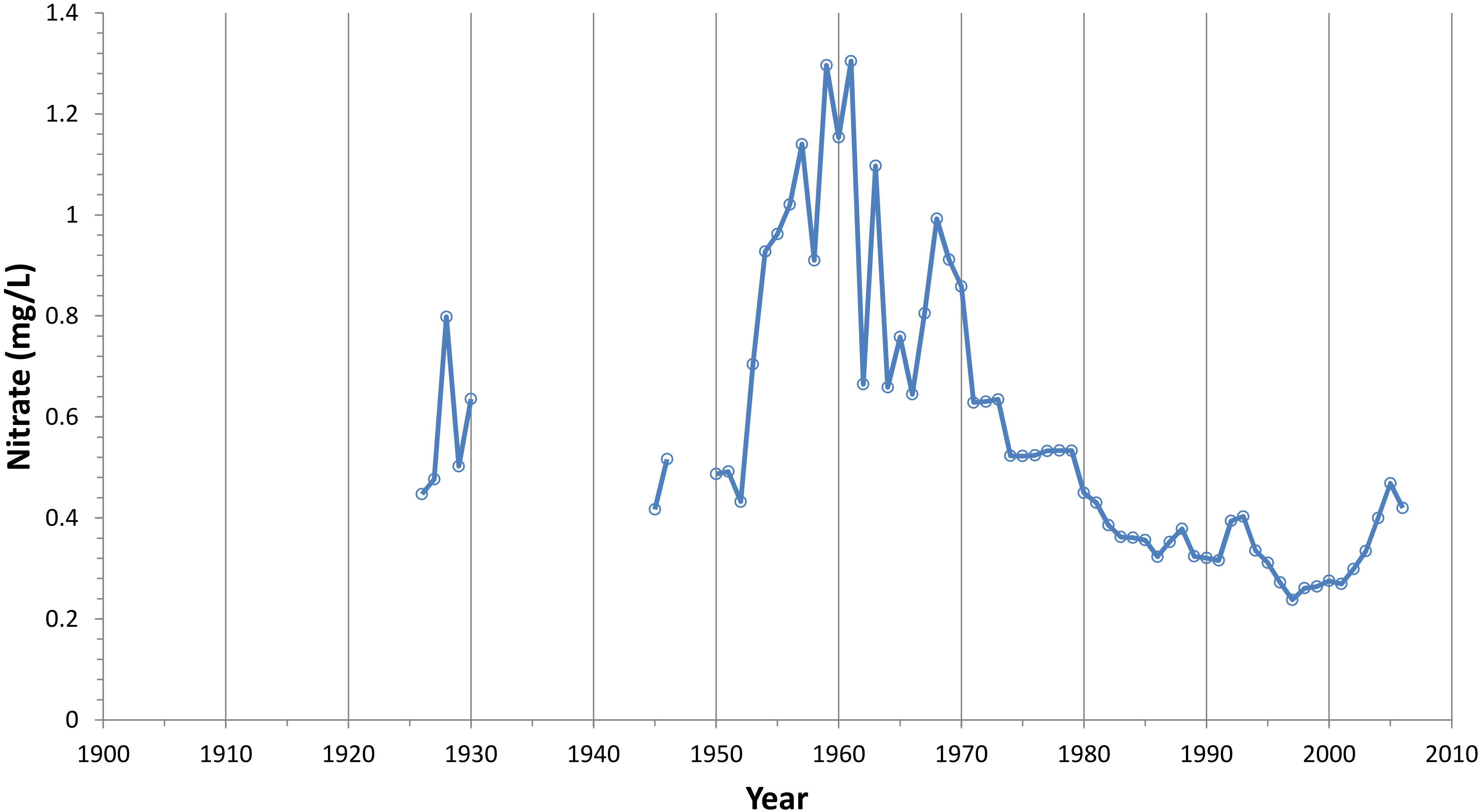
10th National Monitoring Conference
Tampa, FL
May 2 – 6, 2016

USGS
science for a changing world

Recovering metadata for long-term water quality interpretation

- Long-term data can show dramatic changes in water quality, trends otherwise gone unnoticed in short-term studies.
- Interpreting the data correctly requires an understanding of how the data were collected and analyzed.
- This is expressed in the metadata, which is a critical part of data rescue.
- Thus, the water quality archive project.

Colorado River example



Data provided by Edward Stets (USGS)

USGS science for a changing world

Overview of the Water Quality Archive Project (QWAP)

- Boxes, filing cabinets, warehouses full of old laboratory data
- Goal:
  - Digitize these old lab records, make them available through NWIS, QWDX
  - …Provides us with the ability to establish a robust baseline of the chemical, physical, fluvial sediment, and biological condition of US waters = long term trend analysis
- How:
  - Crowdsourcing
  - Develop an online transcription service
  - Make the data NWIS-ready

- What is metadata? Where does metadata fit into this project and its goal?

QWAP and Metadata

- What is metadata?
  - Metadata provides information *about* the data
  - In this context, metadata includes:
    - Lab methods
    - Detection limits
    - Monitoring station location
    - Sampling conditions (i.e. discrete vs. composite)
    - Sampling motivation (i.e. high flow events, base flow sampling, winter sampling, etc.)

QWAP and Metadata

- Where does metadata fit into this project and its goal?

  - Metadata is an essential part of the data rescue project because:
    - The final, reported values are not enough to perform a trend analysis
    - **Metadata provides a complete, thorough, reliable, and trusted account of old data**

↓

  - Enhance older data already in NWIS
  - Ability to establish a robust baseline

Discerning Laboratory Methods and Key Texts

- Key texts to recovering and interpreting laboratory methods:

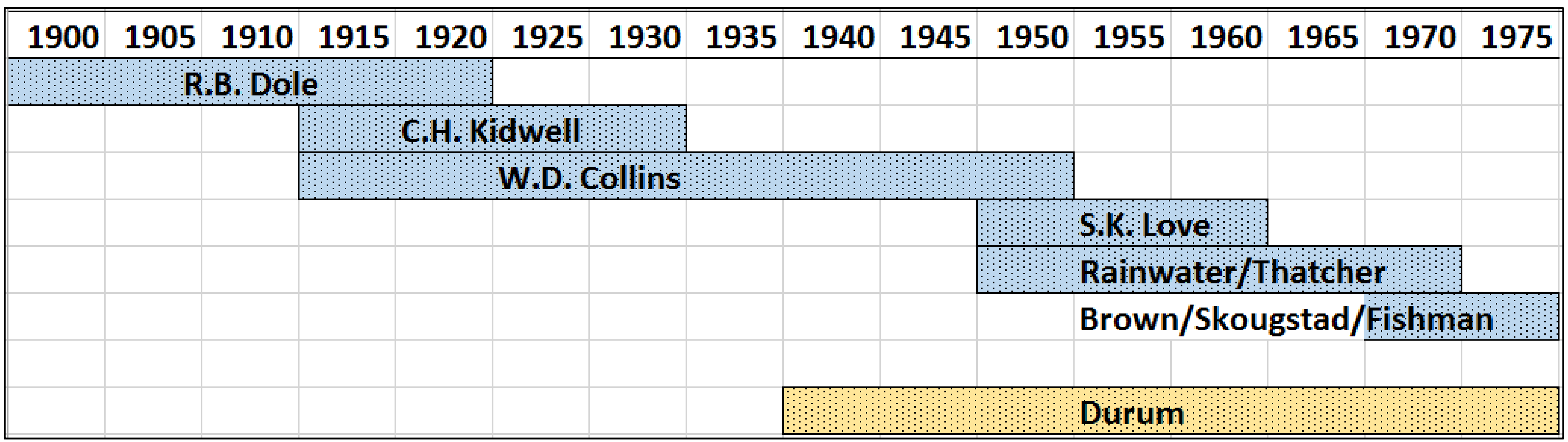
  - **Durum Text:**
    - Durum, W.H. (1978). “Historical Profile of Quality of Water Laboratories and Activities, 1879 – 1973.” U.S. Geological Survey.
  - **Water Supply Paper 236:**
    - Dole, R.B. (1909). “The Quality of Surface Waters in the United States, Part 1. – Analyses of Waters East of the One Hundredth Meridian.” U.S. Geological Survey.
  - **Unpublished Water Analyses Paper (1920):**
    - Kidwell, C.H. (1920). “Methods of Analysis Employed in Water Resource Laboratory.” U.S. Geological Survey.

Discerning Laboratory Methods and Key Texts

- Key texts to recovering and interpreting laboratory methods:
  - **Water Supply Paper 596H:**
    - Collins, W.D. (1923). “Notes on Practical Water Analysis.” U.S. Geological Survey.
  - **Water Supply Paper 1454**
    - Rainwater, F. H.. Thatcher, L. L.. (1960). “Methods for Collection and Analysis of Water Samples – Water Supply Paper 1454.” U. S. Geological Survey.
  - **TWRI Book 5 Chapter A1 (Techniques in Water Resources Investigations):**
    - Brown, Eugene. Skougstad, M.W.. Fishman, M. J.. (1970). “Methods for Collections and Analysis of Water Samples for Dissolved Minerals and Gases.” U. S. Geological Survey.

Discerning Laboratory Methods and Key Texts

Time period	Method Paper	Author
1903-1918	WSP 236	R.B. Dole
1918-1928	Methods of Water Analysis	C.H. Kidwell
1918-1950	WSP 596H	W.D. Collins
1950-1960	Methods of Water Analysis	S.K. Love
1950-1970	WSP 1454	Rainwater/ Thatcher
1970-1975	TWRI Book 5 Chapter A1	Brown/Skougstad/Fishman



Official Record vs. Lab Notes

9-260
(March 1948)

UNITED STATES DEPARTMENT OF THE INTERIOR, GEOLOGICAL SURVEY
WATER RESOURCES BRANCH

2715A

ANALYTICAL STATEMENT
[Parts per million]

Location at Ouray, Utah	Date of collection Sept. 29, 1948		
Source Green River	Use	SiO_2	4.6
Time: 6:30 pm	Temperature	Fe	
— c.f.s.	Color / pH	Ca	60
	Suspended matter	Mg	31
	Hardness as $CaCO_3$ 277	Na	74
	N.C. (133) / Total	K	
	Ignition loss	CO_3	0
	Dissolved solids	HCO_3	176
	Specific conductance at 25°C (micromhos) 809	SO_4	234
		Cl	35
		F	
Chemist E. F. Williams	T/a. ft. 0.72	NO_3	0.5
Lab. No. 1927	ORIGINAL RECORD		
Collector SKL-RTK		Sum	526

16—55248-1

Lab. No. 1927

rCa	2.99	
rMg	2.55	
rNa, rK	3.2N	
	5.54	8.75
rCO_3		-5.54
$rHCO_3$	2.88	3.21
rSO_4	4.87	x 23
rCl	.99	74 Na+K
rF		
rNO_3	.01	37% Na
	8.75	SAR 1.9

Date completed EFW 10/21/48
Checked by WMW 10/21/48
Project Colorado River
Transmitted
Remarks 1947-48 QW REPORT

GPO 16—55248-1

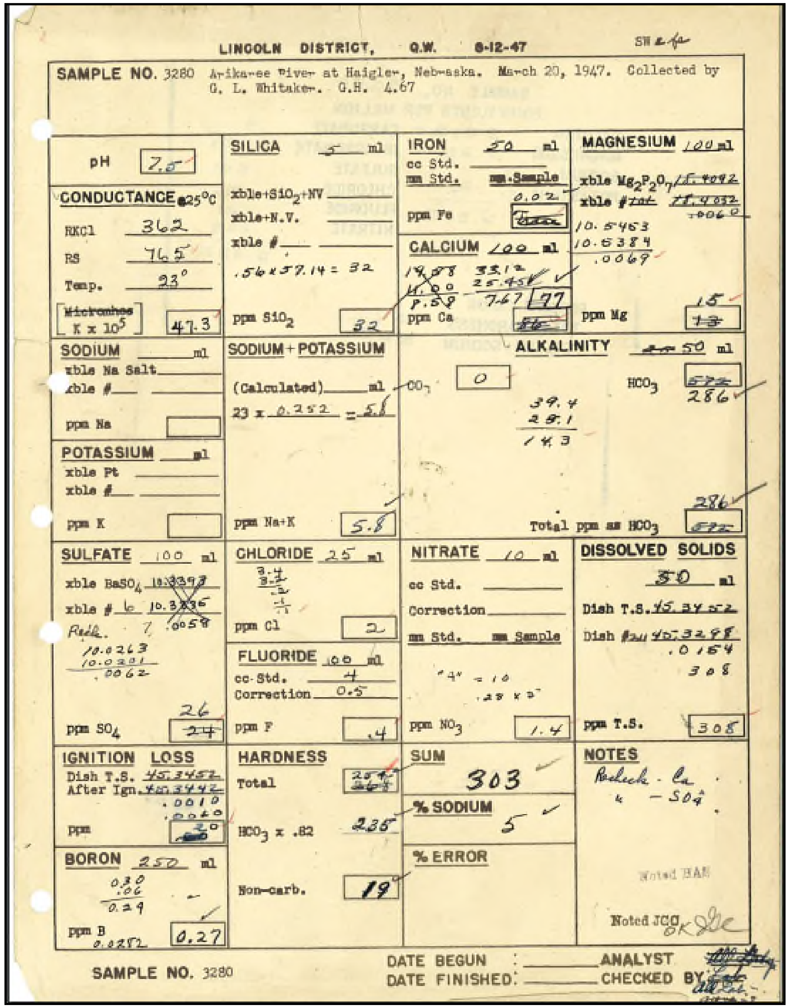
LINCOLN DISTRICT, Q.W. 8-12-47

SAMPLE NO. 3280 Arikaree River at Haigler, Nebraska. March 20, 1947. Collected by G. L. Whitaker. G.H. 4.67

pH 7.5	SILICA 5 ml	IRON 50 ml	MAGNESIUM 100 ml
CONDUCTANCE @25°C: RKCl 362; RS 765; Temp. 23°; Micromhos K x 10^5 47.3	ppm SiO_2 32	CALCIUM 100 ml; ppm Ca 77	ppm Mg 15
SODIUM ml; ppm Na	SODIUM + POTASSIUM (Calculated) ml; 23 x 0.252 = 5.8; ppm Na+K 5.8	ALKALINITY 50 ml; CO_3 0; HCO_3 286; Total ppm as HCO_3 286	
POTASSIUM ml; ppm K			
SULFATE 100 ml; ppm SO_4 26	CHLORIDE 25 ml; ppm Cl 2	NITRATE 10 ml; ppm NO_3 1.4	DISSOLVED SOLIDS 50 ml; ppm T.S. 308
	FLUORIDE 100 ml; cc Std. 4; Correction 0.5; ppm F .4		
IGNITION LOSS	HARDNESS Total 254; HCO_3 x .82 235; Non-carb. 19	SUM 303	NOTES: Recheck - Ca, - SO_4
		% SODIUM 5	
BORON 250 ml; ppm B 0.27		% ERROR	Noted HAS; Noted JCG

SAMPLE NO. 3280 — DATE BEGUN — DATE FINISHED — ANALYST — CHECKED BY

Lab notes can be just as important as methods texts



LINCOLN DISTRICT, Q.W. 8-12-47

SAMPLE NO. 3280 Arikaree River at Haigler, Nebraska. March 20, 1947. Collected by G. L. Whitaker. G.H. 4.67

pH 7.5

CONDUCTANCE @25°C
RKCl 362
RS 765
Temp. 23°
Micromhos K x 10⁵ 47.3

SILICA 5 ml
xble+SiO₂+NV
xble+N.V.
xble #
.56x57.14 = 32
ppm SiO₂ 32

IRON 50 ml
cc Std.
mm Std. mm Sample
ppm Fe

CALCIUM 100 ml
ppm Ca 77

MAGNESIUM 100 ml
xble Mg₂P₂O₇ 15.4092
ppm Mg 15

SODIUM ml
xble Na Salt
xble #
ppm Na

SODIUM + POTASSIUM
(Calculated) ml
23 x 0.252 = 5.8
ppm Na+K 5.8

ALKALINITY 50 ml
CO₃ 0
HCO₃ 286
Total ppm as HCO₃ 286

POTASSIUM ml
xble Pt
xble #
ppm K

SULFATE 100 ml
xble BaSO₄
xble #
ppm SO₄ 26

CHLORIDE 25 ml
ppm Cl 2

NITRATE 10 ml
cc Std.
Correction
mm Std. mm Sample
ppm NO₃ 1.4

DISSOLVED SOLIDS 50 ml
Dish T.S.
Dish
ppm T.S. 308

FLUORIDE 100 ml
cc Std. 4
Correction 0.5
ppm F .4

IGNITION LOSS
Dish T.S.
After Ign.
ppm 20

HARDNESS
Total 254
HCO₃ x .82 235
Non-carb. 19

SUM 303

% SODIUM 5

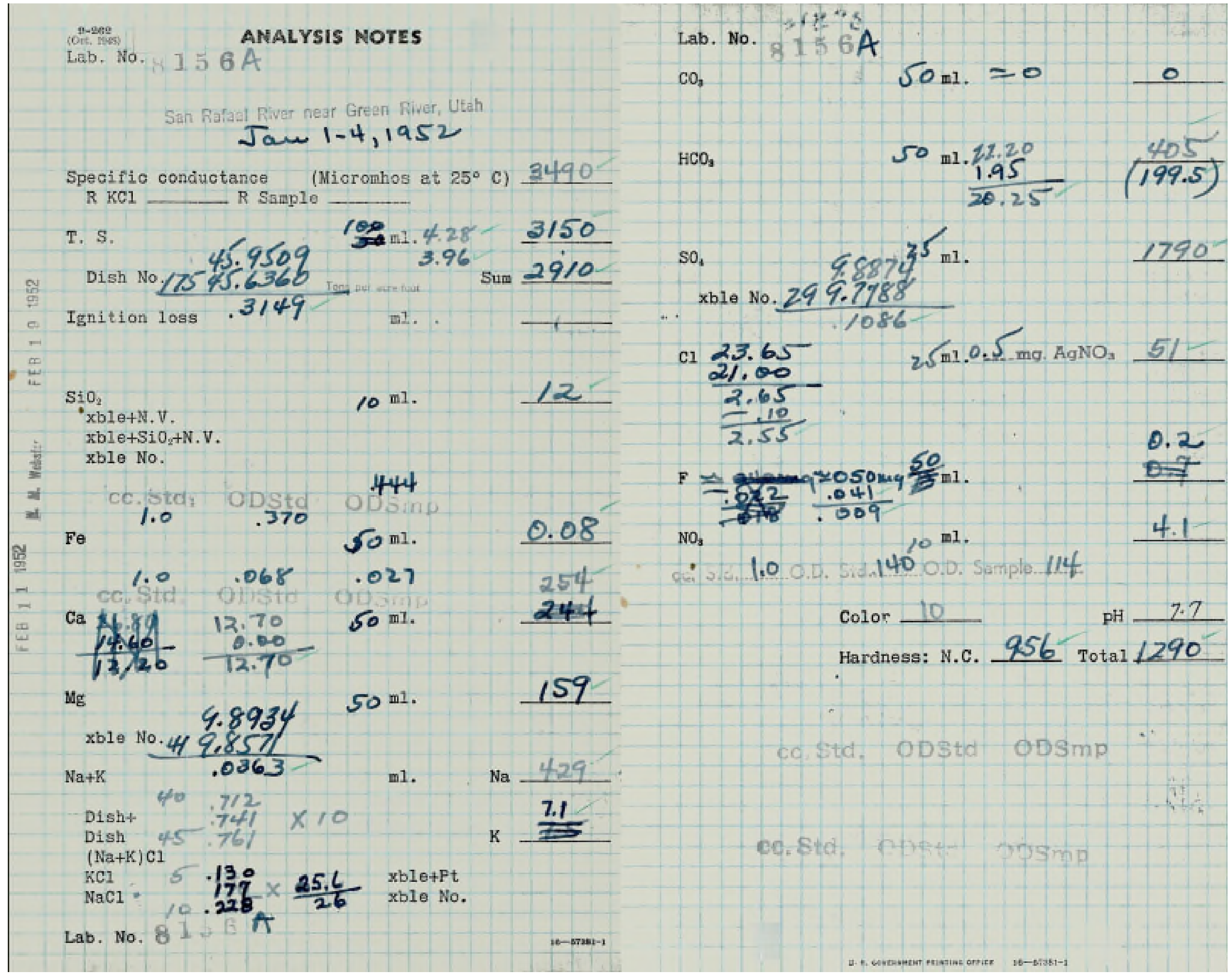
% ERROR

NOTES

BORON 250 ml
ppm B 0.27

Noted JCG

SAMPLE NO. 3280
DATE BEGUN
DATE FINISHED
ANALYST
CHECKED BY



9-262 (Oct. 1945)

ANALYSIS NOTES

Lab. No. 8156A

San Rafael River near Green River, Utah
Jan 1-4, 1952

Specific conductance (Micromhos at 25° C) 3490
R KCl R Sample

T. S. 100 ml. 4.28 3150
Dish No. 175 45.9509 45.6360 Sum 2910
Ignition loss .3149 ml.

SiO₂ 10 ml. 12
xble+N.V.
xble+SiO₂+N.V.
xble No.

Fe 50 ml. 0.08

Ca 50 ml. 254 244

Mg 50 ml. 159
xble No. 41 9.8934 9.8571 .0363

Na+K ml. Na 429
Dish+
Dish
(Na+K)Cl
KCl
NaCl
K 7.1
xble+Pt
xble No.

Lab. No. 8156A

Lab. No. 8156A

CO₃ 50 ml. = 0 0

HCO₃ 50 ml. 405 (199.5)

SO₄ 25 ml. 1790
xble No. 29 9.8874 9.7788 .1086

Cl 25 ml. 0.5 mg. AgNO₃ 51

0.2

F 50 ml.

NO₃ 10 ml. 4.1

Color 10 pH 7.7

Hardness: N.C. 956 Total 1290

9-261
(Oct. 1948)

PARTIAL ANALYSIS NOTES

Lab. No.

8139 N.W. Corner of Tunnel Rypatch reservoir
12-10-51 D 11:30 A.M. CSH 57°F

Specific conductance (Micromhos at 25° C) 803

20.5 R KCl 330 R Sample 411

~~Ca~~ T.H (Vers) 50 ml. 30.55 / 24.90 / 5.65 T.H. 113

Ca – 50 ml. 7.05 / 4.00 / 3.05 X 22 = 67.1 $CaCO_3$ X .4 = Ca 27

Mg by diff. 113.0 / – 67.1 / 45.9 X .2432 ml. Mg. 11

Na+K (calc.)

CO_3 HCO_3 50 ml. 29.80 / 14.65 / 15.15 CO_3 0

HCO_3 303

SO_4 9.3416 / 202. 9.3253 / .0163 100 ml. 67

Cl 7.20 / 3.50 / 3.70 / –.10 / 3.60 25 ml. 0.5 mg. $AgNO_3$ 72

F ml.

NO_3 ml.

cc. Std. OD Std OD Samp

Sum

Hardness: N.C. 0 Total 113 Color

pH

Lab. No. 8139 Tons per acre-foot

U. S. GOVERNMENT PRINTING OFFICE 16—57381-1

FEB 8 1952 MILDRED L. PORTER

FEB 8 1952

8139

cc. Std. OD Std OD Samp

Station: NF Big Blue River at Surprise, Nebr. (B) 06879900

River basin ____ Sta. Ident. No. 06880800 — Date of collection: M 02 D 03 Y 66

Location: midflow over dam — County ____ — Discharge (cfs) 0.2

Collected by: P. D. Zabel, D. Lillie — Lab No. Q-1895 — Project No. ____ — Appearance Wc Clear — Wa ____ — Gage height (ft) 1.28 — Time 1455

Constituent	Details	ppm	epm
SiO_2	Sample size 10 ml; A 0.250 mg; A 0.500 mg .361; A sample .366; Factor 39.4	14	
Al	Sample size ____ ml; A 0.0125 mg; A 0.0250 mg; A sample; Factor; PPM Al = ____ -0.12(ppmFe)-0.04(ppm Mn+O)+0.05(ppmF)		
Fe	Sample size 25 ml; A 0.025 mg; A 0.050 mg; A 0.125 mg; A sample .016; Factor 2.06	Total (b) .03; Dissolved (a)	
Mn	.25 ppm; .50 ppm; .75 ppm; 1.00 ppm; A sample	Total (b) .1; Dissolved (a)	
Ca	50 ml	44	2.20
Mg	epm Ca + Mg 5.64; epm Ca 2.20; epm Mg 3.44	42	3.44
Sr	____ ml		
Na	Dilution; Reading 81.0	34	1.48
K	Dilution; Reading 42.5	17	0.43
Li			
% error 0.3	Total cations		7.55

Constituent	Details	ppm	epm
HCO_3	Sample size 25 ml; 8.79 − 0.00 = 8.79; Total CO_3 173	352	5.77
CO_3	Sample size 25 ml	0	0.00
SO_4	Sample size 25 ml; 3.51 − 0.53 = 2.98; 2.95	59	1.23
Cl	25 ml	16	0.45
F	Sample size 10 ml; A .000 mg 1.500; A .005 mg; A .010 mg; A sample 1.368	.3	0.02
Sample No. M; Card No. 1			
NO_3	Sample size 10 ml; A 0.08 mg; A 0.16 mg; A sample .229; Factor 9.58	2.2	0.04
PO_4	ml ____ Ortho ____ Total; A 0.0125 mg; A 0.0250 mg; A 0.0500 mg; A sample; Factor		
	Total anions		7.51

Item	Details	Value
B	Sample size 5 ml; A 0.005 mg; A 0.005 mg; A sample .061; Factor 1.00	.06
Dissolved solids	100 ml; Residue at 180°C 45.0686; Dish no. 103 45.0275; .0411	411
	Calculated	403
Hardness	Sample size 25 ml; Hardness 284 (40.04); Alk. x .82 289 (33.00); Non. carb. hardness 0 (7.04)	282; 0
Acidity H^{+1}	____ ml	
%Na	SAR	
K x 10^6	510 R Sample; KCl 345 R; 21.3 T	676
pH	7.2	Color
Oxygen cons. filtered or DO	unfiltered or BOD	
Organics: Phenol	ABS - MBAS	
Temperature °F		
Sample No. M	Card No. 2	

Chemist: [illegible] — Checked by: [illegible]

Date began: 5/2 — Completed: 6/6

Punched by: CJR Aug 8 1966 — Verified by: MF AUG - 9 1966

Remarks (other side)

7/65

LQW-10c (July 1957) 1960 A R SW

Sample No. N-7864 Cedar River nr. Spalding, Nebr.

Coll. S.C.
App.

Rem: Fe. bottle filtered thru fiber glass but NOT treated.

Date
Time
G. Ht. .82 wt
Temp. (° 52
Disch. 90 cfs

SILICA 10 ml	IRON a/ ml	IRON b/ 25 ml	CALCIUM 50 ml	MAGNESIUM
	A 0.03 mg	A 0.03 mg 0.520 (0.025)	17.89	epm Ca+Mg 1.420
Factor 86.5	A 0.06 mg	A 0.15 mg 1.040 (0.050)	15.50	epm Ca 1.148
A sample 1.116	Factor	Factor 1.92	2.39	epm Mg 0.272
	A sample	A sample 0.027	-0.05	
			2.34	
ppm SiO_2	ppm Fe	ppm Fe 0.05	ppm Ca 23	ppm Mg 3.3

SODIUM dil	POTASSIUM 1/2 dil	ALKALINITY 50 ml
Sample 66.0 %T	Sample 16.5 %T	ppm HCO_3 105 ppm CO_3 0
5 mg/l 55.0 %T	0 mg/l 7.5 %T	46.44
2.5 mg/l 77.0 %T	5 mg/l 27.0 %T	41.20
		5.24
ppm Na 6.2	ppm K 5.7	

SODIUM AND POTASSIUM (calculated) ____ epm x 23 = ppm Na + K (as Na)

Total ppm as HCO_3 105 CO_3 52 $CaCO_3$

SULFATE 20 ml	CHLORIDE 50 ml	FLUORIDE 10 ml	NITRATE 10 ml	BORON 5 ml
75 ml. 0.2 mg. std.	(1 ml = 0.5 mg Cl)	A 0.000 mg 1.500	A 0.08 mg 0.820	A 0.005 mg 0.961
16.38	13.97	A 0.005 mg 1.291	A 0.16 mg 1.633	A 0.005 mg 1.011
14.30	3.80	A 0.010 mg 1.071	Factor 9.58	Factor 1.01
2.08	0.17	A sample 1.405	A sample 0.050	A sample 0.018
+0.00	-0.15		Corr.	
2.08	0.02			
-2.00				
0.08				
ppm SO_4 2.0	ppm Cl 0.2	ppm F 0.2	ppm NO_3 0.5	ppm B 0.02

DISSOLVED SOLIDS	HARDNESS 50 ml	% Na / SAR	pH	epm	%
CALC. ppm 134	32.82	15 / 0.3	7.5	1.148 Ca	62.5
Res. on evap. at 180° C 100 ml	29.29	SPECIFIC CONDUCTANCE	COLOR	0.272 Mg	14.8
45.4292	3.53	R (KCl) 330	9	Ca+Mg (1.420)	77.3
# 145 45.4156	ppm as $CaCO_3$	R (sample) 1881	CO_2 (Calc.)	0.270 Na	14.7
0.0136	Hardness 71	Temp. °C 23.6		0.146 K	8.0
ppm D.S. 136	Alkalinity HCO_3 x .82 86	Micromhos at 25°C 175	% E 1.3	1.836	100.0
T/A ft 0.18	Non. carb. Hardness 0			1.721 | 1.721 HCO_3	96.3
				0.000 CO_3	0.0
				0.042 SO_4	2.3
				0.006 Cl	0.3
				0.011 F	0.6
				0.008 NO_3	0.4
				1.788	99.9
				Ca+Mg	
				Na+K	

Notes: Collected in a 1 liter pyrex bottle & Fe. bottle.
complete + color

Sevier River nr. Lynndyl, Utah 1958-59 Total SFD 62,734

	Oct.-Mar.	Apr.-Sept.	Total	Disch.	ppm
SiO_3	112,49~~2~~ 1	1,207,856	1,320,348 (434)	62,7~~34~~ 41	21
Fe					
Ca	704,1~~68~~ 70	4,541,569	5,24~~5,737~~ 6,147	"	84
Mg	668,44~~3~~ 4	4,462,691	5,131,~~134~~ 634	"	82
Na	2,204,650 (0,854)	16,998,222	19,20~~2,872~~ 0,417	"	306
K	39,~~216.1~~ 163.1	357,085.8	396,~~346.9~~ 280.7	"	6.3
HCO_3	2,264,11~~8~~ 0	17,245,511	19,5~~69,629~~ 11,590	"	311
SO_4	2,867,84~~4~~ 7	21,034,118	23,90~~1,962~~ 3,559	"	381
Cl	3,202,57~~3~~ 0	21,685,206	24,88~~7,779~~ 9,701	"	397
F					
NO_3 (sum)	10,422.~~3~~ 8	292,161.8 (3.9)	302,58~~4.1~~ 6.7	"	4.8
~~T.S.~~	10,927,41~~6~~ 0	79,085,676	90,01~~3,092~~ 9,910	"	1430
Kx106	17,777,1~~71~~ 50	129,504,103	147,28~~1,274~~ 92,990	"	2350
B-	2,325.6~~1~~ 0	21,699.00	24,02~~4.61~~ 6,04	"	0.38

135164

Dolores River at Gateway, Colorado

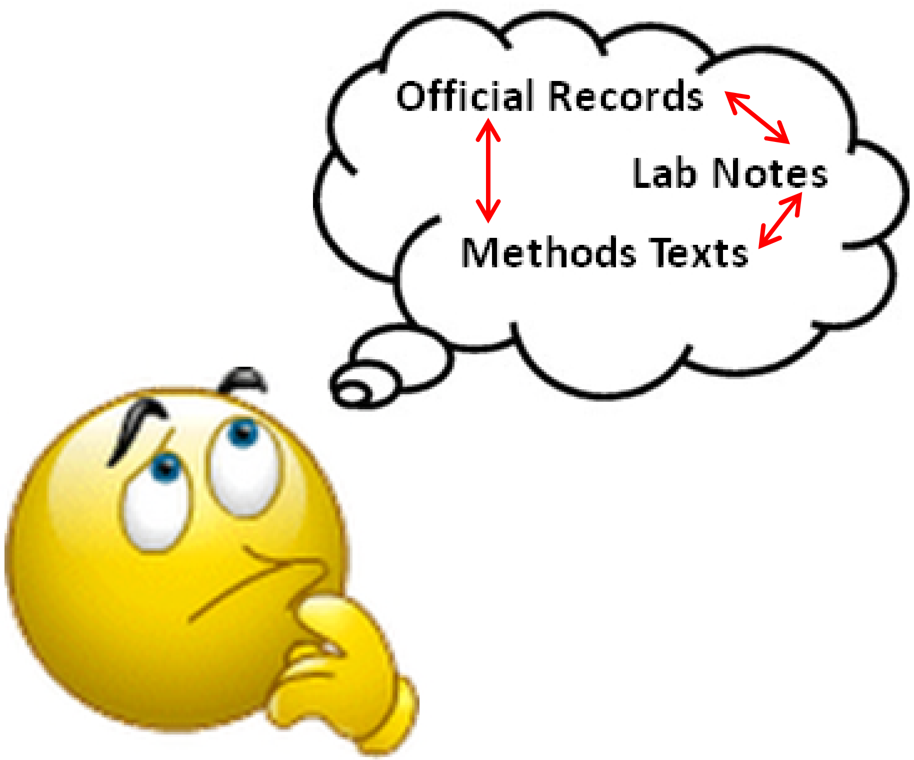
Jan 1-10, 1952

Lab. No. 8143A Max James

Date Jan	Time	~~G. H. (ft.)~~ T°F	T°C	Resistance X° KCl	Resistance X° Sample	(25°C) K x 10^6
1	9:00a	33	21.0	328	107	3070 ✓
2	8:30a	32	21.0	328	115	2850 ✓ A
3	8:10a	33	20.9	328	96.7	3390 ✓
4	9:00a	33	20.9	328	174	1890 ✓ B
5	9:00a	34	20.9	328	108	3040 ✓
6	9:05a	33	20.9	328	81.1	4040 ✓
7	8:45a	33	21.0	328	85.1	3850 ✓ A
8	8:45a	33	20.8	329	77.1	4270 ✓
9	9:20a	33	20.6	330	70.7	4670 ✓
10	8:45a	33	20.2	332	87.5	3790 ✓
					±1180	3660
Composite:-		A	21.3	326	88.5	3680 ✓
		B	20.9	328	174	1890 ✓

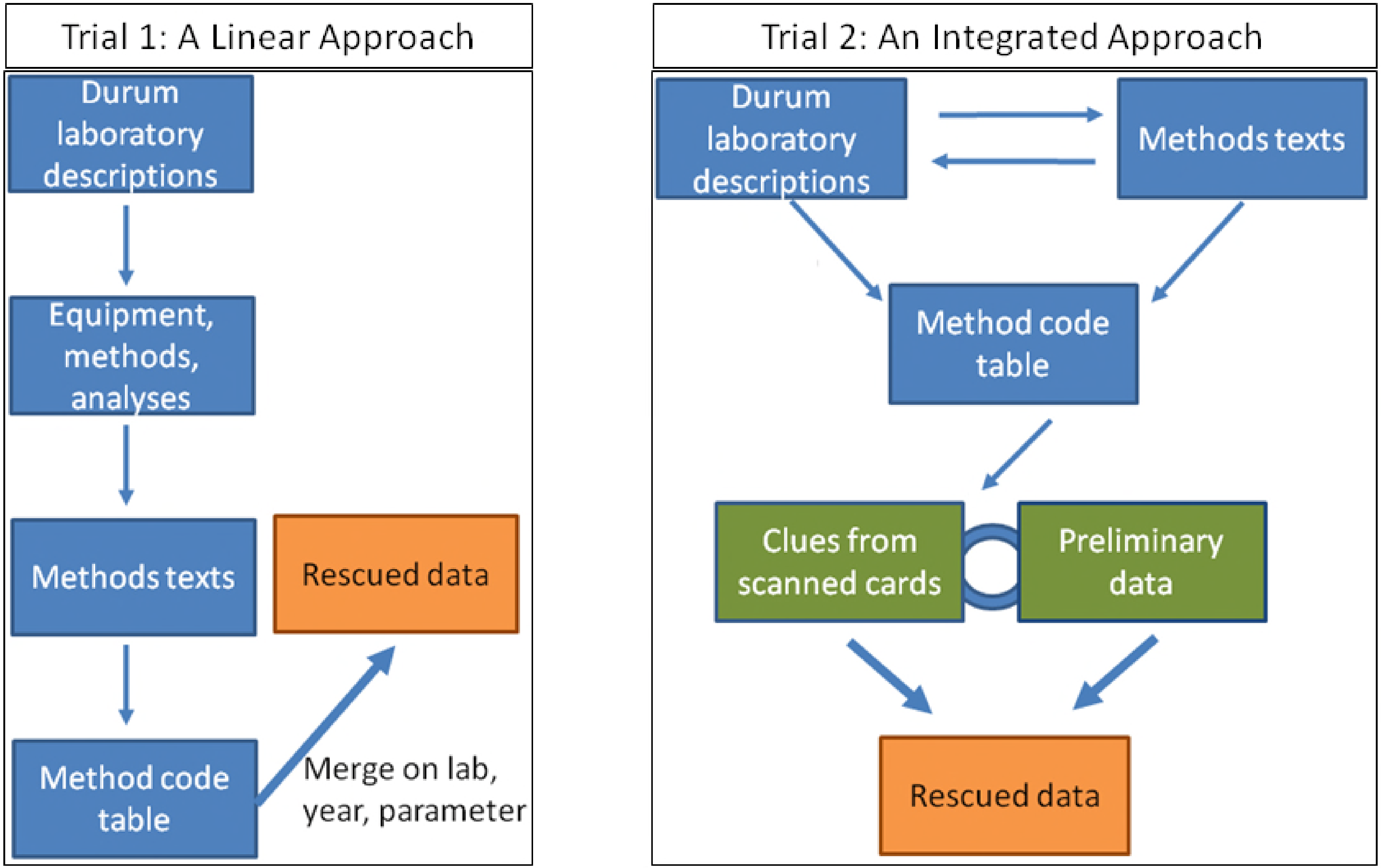
The Result…

Official records, lab notes, and methods texts all work together to provide information about a sample, its results and metadata



Source	Information Provided
Official Records	Location, sampling motive
Lab Notes	Calculated results, methods/techniques, sampling conditions
Methods Texts	Detection limits, instrumentation, methods/techniques

Discerning Laboratory Methods – Not a Linear Process



Concluding Remarks

- Metadata is essential to considering when handling and interpreting old data
- The inclusion of metadata in long-term trend analysis provides a complete, thorough, reliable, and trusted account of old data
- Mining for metadata can be tedious, but is possible
- Metadata provided by this project can have a broader application
- Establishing a robust baseline for current and future scientists