

6.4 RDB Format

by Glenn B. Engel

There are man pages describing the RDB format and separate man pages that give information on each operator used in the RDB system that the user can access from the UNIX command line by typing “man rdb.” Each man page has detailed instructions for use, and examples are given for developing RDB tables.

This document will describe RDB in a general way and describe the data outputs from NWIS for the various data types including the NWIS comments at the top of the RDB table. RDB is a simple format for tabular data and a collection of data management programs using the format. RDB is a fast, portable, relational database management system without arbitrary limits, other than memory and processor speed that runs under and interacts with the UNIX Operating System.

6.4.1 General Data Tables

The data is contained in regular UNIX ASCII files, and can be manipulated by regular UNIX utilities such as ls, wc, mv, cp, cat, more, text editors, head, RCS, etc. Each file of data is in the form of a relation, or table, with rows and columns of information. The relation, or table structure, is achieved by separating the columns with ASCII TAB characters, and terminating the rows with an ASCII NEWLINE character. All rows in a file of data contain data separated by TAB characters and terminated with NEWLINE characters; therefore the data must not contain TAB characters.

Each data file consists of zero or more comment lines starting with a sharp character, “#,” followed by a header consisting of two lines containing structure definition information used by the operators, and zero or more lines of data which are referred to as the body. A file of data, structured as defined above, is said to be an RDB table.

6.4.2 NWIS Data Tables

Any comment that starts with “# //,” is an NWIS comment. Right after the “//” is a keyword that identifies what kind of NWIS data this line is. If it takes more than one line to hold all the data, then the keyword is repeated on the next line. Everything else on the line is a set of keyword=value pairs, separated by blanks. The values have to be double quoted if they contain embedded blanks, otherwise quoting is optional. The following is a sample comment block for DV data:

```

# //FILE TYPE="NWIS-I DAILY-VALUES" EDITABLE=NO
# //DATABASE NUMBER=1 DESCRIPTION=" Montana District NWIS Data"
# //STATION AGENCY="USGS " NUMBER="88110553 " TIME_ZONE="MST"
  DST_FLAG=Y
# //STATION NAME="Scott Bartholoma Test Site at Magna, Utah"
# //DD DDID=" 2" RNDARY="022233332" DVABORT=120
# //DD LABEL="DISCHARGE, IN CFS"
# //PARAMETER CODE="00060" SNAME="DISCHARGE"
# //PARAMETER LNAME="DISCHARGE, CUBIC FEET PER SECOND"
# //STATISTIC CODE="00003" SNAME="MEAN"
# //STATISTIC LNAME="MEAN VALUES"
# //RANGE START="20001001" END="20010331"

```

Daily-Values RDB table

- The FILE record documents the RDB file.
- The DATABASE record documents the number and description of the database from which the data came.
- The STATION record contains information about the site the data came from and takes up two lines. The value pairs are for: agency, station number, time-zone code, daylight-savings time flag, and station name.
- The DD record contains DD information and also takes up two lines. The value pairs are for: data-descriptor ID, DD rounding array, Daily-Values abort limit, and DD label.
- The PARAMETER record contains the parameter information and also takes up two lines. The value pairs are for: parameter code, parameter short name, parameter long name.
- The STATISTIC record contains information about the statistics and also takes up two lines. The value pairs are for: statistic code, statistic short name, and statistic long name.
- The RANGE record documents the start and end dates (in YYYYMMDD format) requested for the retrieval. This date range may be larger than the actual period covered by the data lines.
- After the comments and NWIS “header” described above, the table has the RDB header lines and data as indicated in the example below:

DATE	TIME	VALUE	PRECISION	REMARK	FLAGS	TYPE	QA
8D	6S	16N	1S	1S	32S	1S	1S
20001001		5220	3			C	W
20001002		4870	3			C	W
20001003		4630	3			C	W
20001004		4370	3			C	W

The fields in the first row of the header contain the names of each column. The column names define the data in that column. The fields in the second row contain the data definitions and optional documentation for each column. The actual data starts in the third line.

The data definitions in the second line include column width, data type, and justification. The column width must be specified; the others are optional. The data definitions take the form of adjacent characters in a single word.

The width of each field is specified by a numeric count. The type of data is “string,” “numeric,” or “month.” The types are specified by an “S,” “N,” or “M,” respectively. Default is type string. If not specified, types string and month will be left justified and type numeric will be right justified. It is important to note that only actual data is stored in the data fields, with no leading or trailing space characters. This fact can have a major effect on the size of the resulting files.

In the example above, the first column in the data line contains the date, using eight character spaces and a specific type designation “D” for the date that is used so that graphics programs can recognize dates. The second column contains the time with six spaces in the string, but there is no time designated, as these are Daily-Values. The third column contains the discharge value, with up to 16 numerical characters. The fourth column contains the precision code of three-place rounding. The fifth column contains a remark code, if present. The sixth column contains a definition of a flag on the data, which can be a 32-character string. The seventh column contains the data type, which is “computed” in this case and is a single character string. The eighth column contains the quality-assurance flag, which shows that the data is in “Working” status.

6.4.3 Unit Values RDB Table

The following is an example of an RDB table retrieved for Unit-Values. Differences from the Daily-Values table are that there is an NWIS comment line “TYPE” with the CODE=C for computed Unit-Values. In the data lines, times are listed and a column headed TZCD, which indicates the time zone code, in this case EST (Eastern Standard Time).

```

# //FILE TYPE="NWIS-I UNIT-VALUES" EDITABLE=NO
# //DATABASE NUMBER=1 DESCRIPTION=" Montana District NWIS Data"
# //STATION AGENCY="USGS " NUMBER="01010000  "
TIME_ZONE="EST" DST_FLAG=Y
# //STATION NAME="St. John River at Ninemile Bridge, Maine"
# //DD DDID=" 1" RNDARY="022223332" DVABORT=120
# //DD LABEL="DISCHARGE (well-DCP), in CFS"
# //PARAMETER CODE="00060" SNAME="DISCHARGE"
# //PARAMETER LNAME="DISCHARGE, CUBIC FEET PER SECOND"
# //TYPE CODE=C NAME=COMPUTED
# //RANGE START="20010101000000" END="20010110235959"
ZONE="LOC"
DATE TIME TZCD VALUE PRECISION REMARK FLAGS QA
8D 6S 6S 16N 1S 1S 32S 1S
20010101 000000 EST 5390 3 W
20010101 001500 EST 5390 3 W
20010101 003000 EST 5390 3 W
20010101 004500 EST 5390 3 W

```

6.4.4 Measurement RDB Table

The following is an example of an RDB table retrieved for discharge measurements. The column names and data are the same as in the ADAPS measurement summary tables.

```

# //FILE TYPE="NWIS-I DISCHARGE MEASUREMENTS"
# //DATABASE NUMBER=1 DESCRIPTION=" Montana District NWIS Data"
# //STATION AGENCY="USGS " NUMBER="01010000  "
TIME_ZONE="EST" DST_FLAG=Y
# //STATION NAME="St. John River at Ninemile Bridge, Maine"
# //RANGE START="20001001" END="99999999"
NUMBER DATE TIME PARTY WIDTH AREA VELOCITY IGH
OGH Q1 QCODE1 Q2 QCODE2 RATING SHIFT PCTDIF NSE
CT GHCHGF GHCHGT RATED AIRT WATERT CONTROL MSTYPE
BASEFLOW REMARKS
8S 8D 4S 8S 12N 12N 12N 12N 12N 12N 10S 12N
10S 12N 12N 12N 6N 12N
12N 2S 12N 12N 10S 10S 2S 100S
318 20010108 1710 TCS/JMC 385 670 1.63 1090
MEASURED 0.00
39 0.00 1.2 F ICE COVER ICE 100% ice cover on control; GH = 3.75
319 20010319 1716 TCS/GRS 415 419 1.01 425 MEASURED 0.00
37 0.00 1.0 F ICE COVER ICE 100% ice cover; cfsm=.32
320 20010604 1150 TCS 422 1880 2.23 4.00 4.00 4190 MEASURED
5.0 0.00 0.0 32 0.00 1.8 G CLEAR CABLEWAY

```

6.4.5 Expanded Rating RDB Table

The following is an example of an RDB table retrieval of an expanded rating table. The NWIS comment lines unique to rating table retrievals are the RATING ID and TYPE and NAME, RATING REMARKS, RATING EXPANSION, both INDEPENDENT and DEPENDENT PARAMETER rounding arrays, and RATING_DATETIME with BEGIN and END times with corresponding time zone codes BZONE and EZONE.

```

# //FILE TYPE="NWIS RATING"
# //STATION AGENCY="USGS " NUMBER="01010000    "
  TIME_ZONE="EST"  DST_FLAG=Y
# //STATION NAME="St. John River at Ninemile Bridge, Maine"
# //DD NUMBER=" 1" LABEL="DISCHARGE (well-DCP), in CFS"
# //PARAMETER CODE=""
# //RATING ID=" 5.0" TYPE="STGQ" NAME="stage-discharge"
# //RATING REMARKS="New low end and refinement of high end of rating 4"
# //RATING EXPANSION="logarithmic"
# //RATING_INDEP ROUNDING="0223456782" PARAMETER="GAGE
  HEIGHT in (FEET)"
# //RATING_DEP ROUNDING="0222233332" PARAMETER="DISCHARGE
  in CFS"
# //RATING_DATETIME BEGIN=19931001010000 BZONE=EDT
  END=23821230190000 EZONE=EST
INDEP DEP  STOR
16N  16N  1S
0.50  80   *
0.51  82
0.52  85
0.53  87
0.54  90
0.55  92
0.56  95
0.57  97
0.58  100
0.59  102
0.60  105
0.61  107
0.62  110   *

```

The data lines consist of the independent parameter, gage height, and the dependent parameter, discharge, with the stored points indicated with an asterisk. An RDB table consisting of only the stored rating points can be retrieved from ADAPS also.

6.4.6 Peak-Flows RDB Table

The following is an example of an RDB table retrieved for peak flows. The values for RANGE start, 00000000, and end dates, 99999999, indicate that the entire period of the record was requested.

```
# //FILE TYPE="NWIS-I PEAK FLOWS"
# //DATABASE NUMBER=1 DESCRIPTION=" Montana District NWIS
Data"
# //STATION AGENCY="USGS " NUMBER="01010000    "
TIME_ZONE="EST" DST_FLAG=Y
# //STATION NAME="St. John River at Ninemile Bridge, Maine"
# //RANGE START="00000000" END="99999999"
QDATE  QTIME  QVALUE  QQUALS  QGH   QGHQUALS
YEAR_LAST_PEAK  GHDATE  GHTIME  GHVALUE  GHQUALS
8N   6N   8N   12S   8N   4S   4N   8N   6N   8N   4S
```

The column names for the data lines are: QDATE, peak discharge date; QTIME, peak discharge time; QVALUE, peak discharge; QQUALS, peak discharge qualification codes; QGH, gage height of peak discharge; QGHQUALS, gage height qualification codes; YEAR_LAST_PEAK, recorded peak is highest since listed year; GHDATE, date of maximum gage height if different than gage height of peak discharge, GHTIME; time of maximum gage height; GHVALUE, maximum gage height; GHQUALS, maximum gage height qualification codes.

6.4.7 Shift RDB table

The following is an example of an RDB table retrieval for shifts. The DD used for retrieval is the discharge DD because the rating is tied to the discharge but the parameter that the shifts are applied to is the gage height.

```

# //FILE TYPE="NWIS-I SHIFTS"
# //DATABASE NUMBER=1 DESCRIPTION=" Montana District NWIS Data"
# //STATION AGENCY="USGS " NUMBER="01010500  "
TIME_ZONE="EST" DST_FLAG=Y
# //STATION NAME="St. John River at Dickey, Maine"
# //DD DDID=" 5" RNDARY="0223456782"
# //DD LABEL="DISCHARGE, IN CFS"
# //PARAMETER CODE="00065" SNAME="GAGE HEIGHT"
# //PARAMETER LNAME="GAGE HEIGHT, FEET"
# //RANGE START="19991001" END="20011030"
RATCD RATNO SEQ BEGDATE BEGTIME BEGTZCD ENDDATE
ENDTIME ENDTZCD INDEP SHIFT
      3N  5S  5S  8D  6S  6S  8D  6S  6S  16N  16N
      STGQ 14.0 1  20000426  234500 EDT  0.00
0.00
      STGQ 14.0 2  20000426  234500 EDT  4.00
0.00
      STGQ 14.0 3  20000426  234500 EDT  5.00
0.00
      STGQ 14.0 1  20000427  000000 EDT  0.00 -
0.06
      STGQ 14.0 2  20000427  000000 EDT  4.00 -
0.06
      STGQ 14.0 3  20000427  000000 EDT  5.00
0.00

```

The data lines are in sets of three for each variable shift entry indicating the three points of the shift diagram. The first entry is the rating code STGQ for stage-discharge and then the rating number. The variable shifts are in three consecutive lines with sequence numbers 1 through 3, begin date and time, time-zone code, end date and time with time-zone code (there are no entries for end dates in this example), independent variable, gage height, for the three diagram points, and the three shifts at the gage height points.

6.4.8 Data Corrections RDB table

The following is an example of an RDB table retrieved for data corrections. The NWIS comments have been explained in earlier examples. The data corrections are applied to the input DD such as the gage height in NWIS.

```

# //FILE TYPE="NWIS-I DATA CORRECTIONS"
# //DATABASE NUMBER=1 DESCRIPTION=" Montana District NWIS Data"
# //STATION AGENCY="USGS " NUMBER="01010500 "
TIME_ZONE="EST" DST_FLAG=Y
# //STATION NAME="St. John River at Dickey, Maine"
# //DD DDID=" 2" RNDARY="0223456782"
# //DD LABEL="GAGE HEIGHT Cross-section: 590.0, in (FEET)"
# //PARAMETER CODE="00065" SNAME="GAGE HEIGHT"
# //PARAMETER LNAME="GAGE HEIGHT, FEET"
# //RANGE START="19991001" END="20011030"
SET SEQ BEGDATE BEGTIME BEGTZCD ENDDATE ENDTIME
ENDTZCD INDEP CORR
3N 3N 8D 6S 6S 8D 6S 6S 16N 16N
1 1 20000320 130800 EST 0.00 0.00
1 1 20000815 080500 EDT 20000815 080600 EDT 0.00
-0.01
1 1 20010416 120000 EDT 0.00 -0.19
1 1 20010516 163000 EDT 0.00 -0.19
1 1 20010601 073000 EDT 0.00 0.00
1 2 20010601 073000 EDT 2.00 -0.02
1 3 20010601 073000 EDT 5.00 -0.04
2 1 20010601 073000 EDT 0.00 0.00
2 2 20010601 073000 EDT 2.00 0.02
2 3 20010601 073000 EDT 5.00 0.10
3 1 20010601 073000 EDT 0.00 0.00
3 2 20010601 073000 EDT 2.00 0.00
3 3 20010601 073000 EDT 5.00 -0.15

```

The data lines include the set number, sequence number, begin date and time and time-zone code, end date and time and time-zone code (if available), and the stage and data correction pair. In the example, on June 1, 2001, there are three data correction diagrams of three points each, hence the three sets labeled 1 through 3 and the three data-correction pairs for each set with sequence numbers 1 through 3. In ADAPS, set 1 is designated Gage Height Corrections, set 2, Datum Corrections from Levels; and set 3, Other Corrections.