

3.8 Option 8 -- Batch Processing

<p style="text-align: center;">YOU ARE USING WATER-QUALITY DATABASE NUMBER 01</p> <p style="text-align: center;">Batch Processing</p> <p>1 -- Retrieve National Water-Quality Laboratory Data</p> <p>2 -- Process Batch File for All Logged In Samples</p> <p>3 -- Process Batch File for All Samples</p> <p>4 -- Process Batch File for All Logged In QA Samples</p> <p>5 -- Process Batch File for All QA Samples</p> <p>6 -- Produce Batch Output</p> <p>7 -- Review/Edit Batch files</p> <p>8 -- Produce 1- and *-Card Output</p> <p>9 -- Reload QW data from batch file, overriding DQI</p> <p>10 -- Reload QA data from batch file, overriding DQI</p> <p>98 -- Exit menu</p> <p>99 -- Exit system</p> <p>Please enter a number from the above list or a Unix command:</p>

Batch Processing Menu

The options in this menu work with batch files of water-quality data. Batch files are used for many purposes such as entering data from laboratories, making large-scale changes to the water-quality data, or output of entire records by specifying only the record information. The most common application is entering data from laboratories.

Batch File Format

The NWIS software accepts two batch file formats:

(1) A fixed-format of ‘1’ and ‘*’ records (also referred to as ‘cards’). All information is in a file named either *qwcards* or *qacards*. The 1-card contains all the *sample level information* including station ID, date, time, and medium code. The *-cards follow the 1-card and contain the *result level information* including parameter codes, parameter values, and value remark codes. This format was the batch format designed for versions of NWIS prior to NWIS 4_1. Although the format is still accepted by the software, it does not contain any of the sample and result information added in release 4_1 or later. The format of the 1- and *-cards for this type of batch file is defined in [Appendix F](#).

(2) A tab-delimited format uses a pair of files (*qwsample* and *qwresult* or *qasample* and *qaresult*) to hold all the sample and result information. This batch file pair will contain all sample and result level information currently stored in the water-quality database. The related data in the file pair is connected by a ‘sample integer’ generated when the two files are created. Although the sample integer does not have any meaning beyond the

batch file pair, it is critical to keeping the sample and result information properly connected. The sample and result file formats are defined in [Appendix F](#).

The batch input programs look for the batch files in the directory where the program is initiated. Both the 1- and *-card and tab-delimited batch file formats are accepted by the batch input programs in **Options 1-5, 9, and 10** of the Batch Processing Menu, and both will be processed if present in the working directory simultaneously. The 1- and *-card files will be processed first.

Environmental and QC Data

Environmental and quality-control data are entered with the batch entry programs available in this menu. A process is initiated to separate the environmental data from the QC data (*qa_sep*) so that slightly different protocols can be applied to the entry of each and the data can be written to separate databases, if they exist. Samples with a medium code of Q-Z are written to the file *qacards* or the file pair *qasample/qaresult*, depending on the format of the input file(s). (If those filenames exist, the existing filenames are changed to include an extension of ‘.old.yymmdd.hhmm’) Environmental samples are entered into the default environmental database and *QA* samples are entered into the default quality-assurance database. The only difference in the data entered in each database, is that the District Processing Status code (DSTAT) is set to ‘Local Only’ for the QC data. The separate processes produce separate output files with different filenames (a ‘.qa’ is added to the filenames) to distinguish the environmental output files from the QC output files.

The association of environmental and quality assurance water-quality databases and related data entry behavior can be a little complicated. The program used, the database the user is ‘pointing to’, the type of data contained in the batch files (environmental, QC, or a combination), and the existence or non-existence of the environmental database and an associated QC database all affect the actions taken by the software. The typical conditions and expected actions are described for each program. Although the programs will accept other conditions, the actions may be both unexpected and less than desirable.

Data Entry Behavior

The behavior of the batch input processes for the tab-delimited files with respect to overwriting existing data, treatment of nulls, and the use of default values can be found in the format table in [Appendix F](#).

Most of the entries allowed in the batch file fields are self explanatory or defined in the domain lists for the coded fields. There are some ‘tricks’ that might be useful:

1- and *-card files:

1. To delete a result value, enter a ',' (comma) after the '=' sign in the *-card. It would look like this to delete the result value for parameter 00095: *00095=,
2. Set a remark code to 'X' to delete a result and all of the associated attributes.
3. A #-card can be inserted after the 1-card to add project codes to the input. The '#' should be in the first column, and the project code should be in columns 2-10.
4. To delete an entire sample, enter 'DELETE' for GUNIT.
5. To enter a null value with a remark code, enter a '#' in the value field and a valid null-value remark code in the remark field.

Tab-delimited files:

1. Set a remark code to 'X' to delete a result and all of the associated attributes.
2. To delete a value qualifier code, set the value qualifier code to '#(code to be deleted)'. For example, to delete a value qualifier code of 'a', the value qualifier code should be '#a' in the batch file.
3. If a time datum is included in a batch file, the time datum will be updated in the database, but the UTC sample time will not be changed. The result may not be desired, so consider carefully before including a time datum in a batch file. If you want to change the stored time, consider changing the time using [Modify Sample or Results \(section 3.2\)](#) or a separate program [qwtimeshift \(section 3.9.4\)](#).
4. Batch files may be prepared using computer systems running Microsoft (TM) operating systems. Often such files need additional processing to be interpreted properly on a computer running the Unix operating system (such as your NWIS system). You can test one of these files using the following Unix command to determine if additional processing is needed (replace *filename* with the name of the file you want to test).

```
perl -ne 'if (/^r\n$/) {print "MS-DOS file\n";exit;} < filename
```

You may safely ignore messages regarding “keyboard type.” All the files in the current directory can be rewritten with their Unix-format equivalents, using the following Unix command:

```
ksh -c 'for x in * ; do dos2unix $x $x ; done'
```

Explanation: MS-DOS text files are formatted with lines that end with a carriage return, followed by a linefeed. Unix text files are formatted with lines that end with a linefeed only. Data-processing programs in Unix interpret the carriage return (a non-printing control character) as data, which often causes confusing errors. In some instances, the QWDATA batch processing software might print error messages regarding apparently valid codes as being invalid when MS-DOS formatted batch files are supplied. The “dos2unix” command removes the unwanted carriage return.

If the batch files are obtained using “ftp,” then the files should be transferred using “type ascii.” Use of this option will cause the ftp operation to perform the conversion from MS-DOS format to Unix format.

Record of Actions Taken (WATLIST)

A record of the actions taken when a batch file is processed with a batch file program is written to the files named *watlist.yymmdd.hhmm* and *watlist.qa.yymmdd.hhmm*. The records created or modified, the cation/anion balance (if it can be computed), and any error information generated for samples and results are recorded in these files. Samples that are not processed are shown in this file and are differentiated from processed samples by a blank record number field and an error message stating why the sample wasn't processed. . The WATLIST is automatically sorted by project code. A sample WATLIST output can be found in [Appendix C](#).

The parameter codes for the parameters listed in the WATLIST output are preceded by a one-letter code that informs the user if any changes were made to the stored result for that value. The column header for this code is '*' in the WATLIST output.

Code	Definition
N	New parameter
U	Updated parameter
C	Calculated parameter

In the error section, messages for those results that are greater than defined limits ([see appendix E](#)) will appear. As of NWIS 4.3, a result with any remark code other than a less-than symbol (<) will be compared to the list of defined limits. If the result is greater than the defined limit, an error message will appear.

Rejected Samples and Results

Samples will not be processed for the following reasons: (a) invalid formats, (b) invalid site ID's, dates, times, or medium codes, (c) results for samples that do not already exist in the water-quality file (except menu options 3 and 4), and (d) samples that contain results that are overwrite protected with a DQI value (except menu options 9 and 10). All results are rejected for a sample if a rejection error is found for one or more results. After the necessary steps are taken to correct any problems, the corrected samples can be entered by renaming the files to an appropriate batch filename and initiating one of the batch input programs.

Rejected samples from 1-and *-card files:

The files, *badqw.yymmdd.hhmm* and/or *badqa.yymmdd.hhmm* are created when samples are rejected from a *qwcards* batch file.

An E-card is inserted in the *badqw* files to describe an error after the card causing the error. The E-card is treated as a comment and ignored by the input programs.

Rejected samples from tab-delimited files:

The files, *rejected.sample.yyyymmdd.hhmmss* and *rejected.result.yyyymmdd.hhmmss* and/or *rejected.sample.qa.yyyymmdd.hhmmss* and *rejected.result.qa.yyyymmdd.hhmmss* are created when samples are rejected from a *qwsample/qwresult* batch file pair.

A '#' followed by a sample integer is used to insert error messages in the rejected sample and result files. The error message is included before the record causing the error. The format of the error messages in the sample file is:

#SINT<tab>error message

The format of the error messages in the result file is:

#SINT<tab>parameter code<tab>error message

These error messages are treated as comments and ignored by the batch input programs. However, the program in Option 7 to review or edit the tab-delimited batch files will show the error messages.

Additional Output from the Batch Programs

A command output file, *program.como.yymmdd.hhmmss*, is produced to capture runtime information and errors encountered during the run of the program. This file should be checked to verify successful processing.

3.8.1 Option 1 -- Retrieve National Water-Quality Laboratory Data

The *qwgetlab_cmd* program is used by the person(s) responsible for water-quality data management to process data that has been analyzed by the NWQL (National Water-Quality Laboratory). The NWQL sends completed laboratory data to the District server and it is placed in the */var/ftp/incoming/wrd_only/nwis.nwql* directory.

The NWIS 4_1 tab-delimited formatted files are named:

SAMPLE.uc.111.111.111.11.222.222.222.2.yymmdd.hhmmss.checksum.t.a
RESULT.uc.111.111.111.11.222.222.222.2.yymmdd.hhmmss.checksum.t.a

Where:

uc = user code

111.111.111.11 = IP address of the machine where the data came from

222.222.222.2 = IP address of the machine where the data is transferred to

yymmdd = date

hhmmss = time

checksum = the checksum value used to make sure no data were lost in the transfer

The NWIS 3_2 card-image formatted files (1 and * cards) are named:

ADATA.uc.111.111.111.11.222.222.222.2.yymmdd.hhmmss.checksum.t.a

Where:

uc = user code

111.111.111.11 = IP address of the machine where the data came from
222.222.222.2 = IP address of the machine where the data is transferred to
yymmdd = date
hhmmss = time
checksum = the checksum value used to make sure no data were lost in the transfer

Generally, the user will not need to know the names of the files being transferred from the *nwis.nwql* directory. The *qwgetlab_cmd* program is invoked as an interactive process and searches for the lab data in the */var/ftp/incoming/. wrd_only/nwis.nwql* directory. A file named *qwgetlab.como.yymmdd.hhmmss* is produced showing the results of the *qwgetlab_cmd* program.

If the files *qwcards*, *qwsample*, or *qwresult* already exist in the working directory, the names of the existing files are changed to *qwcards.old.yymmdd.hhmmss*, *qwsample.old.yymmdd.hhmmss*, and *qwresult.old.yymmdd.hhmmss*.

The following messages are displayed on the screen when no lab data are found:

```
Enter user/district code if required:

qwgetlab initiated: Thu Mar 15 13:47:35 EST 2001

Working in directory: /home/tester/qwgetlab_dir

Searching for any lab data in: /var/ftp/incoming/. wrd_only/nwis.nwql

No lab data found.

Output will be logged to: qwgetlab.como.010315.134733
```

If laboratory data exist in the */var/ftp/incoming/. wrd_only/nwis.nwql*, the files are moved to the working directory where the user was attached when the *qwgetlab_cmd* was invoked, and renamed to *qwcards*, *qwsample*, or *qwresult*, depending on what file formats were found. The following types of messages are displayed on the screen:

```
Enter user/district code if required:

qwgetlab initiated: Thu Mar 15 18:48:01 EST 2001

Working in directory: /home/tester/qwgetlab_dir

Renaming old copy of qwsample to qwsample.old.010315.184801

Renaming old copy of qwresult to qwresult.old.010315.184801

Searching for any lab data in: /var/ftp/incoming/. wrd_only/nwis.nwql

Checksum OK for: ADATA.08.130.118.109.93.136.177.8.2.20000310.083438.4949.t.a
Moving to qwcards
```

found lab SAMPLE file: SAMPLE.30.130.118.109.93.136.177.224.5.20000306.140951.t.a
Moving to qwsample

found lab RESULT file: RESULT.30.130.118.109.93.136.177.224.5.20000306.140951.t.a
Moving to qwresult

Found 1 files containing lab ADATA data.
Found 1 files containing lab SAMPLE data.
Found 1 files containing lab RESULT data.

Output will be logged to: qwgetlab.como.010315.184800

If more than one batch file is found in the *nwis.nwql* directory, the files of the same format are appended to each other. The total number of files found is displayed to the screen and included in the *qwgetlab.como* file.

This program also may be invoked by entering *qwgetlab_cmd* from the Unix prompt. There are four arguments that may be supplied to the command:

- **-u** user/district code, if any is required.
- **-w** working directory (generally a directory under */usr/opt/nwis/data/auxdata*)
- **-p** processing program (*qw_cardsin* / *qw_enter*) if processing is desired
- **-d** database number

Application of the **-u** option allows the user to specify the user/district code, located in the batch filename, so that only those files are retrieved. If no code is specified, all user/district codes will be merged into one set of batch files. Application of the **-w** option will allow output from the *qwgetlab_cmd* program and from the processing programs to be written to the directory specified. If no directory is specified, the default is the current working directory. Application of the **-p** option will allow the *qwgetlab_cmd* program to move the data to the working directory and then process the data using one of the batch entry programs. Application of the **-d** option allows the user to specify a database number (the default database is 01).

The processing of laboratory-transmitted data can also be set up as an operation that is completed in off-hours. An example of a cron job file is available in */usr/opt/nwis/data/auxdata/qwgetlab_cronjob.master* and *qwgetlab_crontab.master*. To set up this retrieval to run in off-hours (as a cron job), please contact the NWIS help (email: GS-W Help NWIS) for instructions.

3.8.2 Option 2-- Process Batch File for All Logged-In Samples

This option runs *qw_cardsin*, a program that *updates* sample records with sample and result information from batch files. Thus, an entry for each sample (environmental and QC) must already exist in NWIS.

Generally, the NWQL or other laboratories transfer data to a District and *qw_cardsin* is used to enter the data into NWIS. The user will ‘point’ to the appropriate environmental database number. QC data will be separated and written to an associated QA database.

3.8.3 Option 3-- Process Batch File for All Samples

This option runs *qw_enter*, a program that *enters or updates* analytical data into NWIS. This program is different from *qw_cardsin* in that a record for a sample does not have to be in the database for the data to be successfully processed.

This program may be used to update existing records with USGS laboratory analytical data, create new records, add analytical data from non-USGS sources, and perform limited editing functions. Using *qw_enter* to create records is not recommended. Erroneous information entered in key identification fields (site ID, date, time, and medium code) can create unintended records instead of updating the intended records. These erroneous records are sometimes difficult to detect and are time-consuming to repair.

To appropriately use *qw_enter*, the user will ‘point’ to the appropriate environmental database number. QC data will be separated and written to an associated QA database.

3.8.4 Option 4 -- Process Batch File for all Logged in QA Samples

This option runs *qa_cardsin*, a program that *updates* QA sample records with sample and result information from batch files. Thus, an entry for each sample must already exist in NWIS.

This program is identical to *qw_cardsin*, but is tailored to process a batch file containing only QC data. The user should ‘point’ to the appropriate QC database. The program to separate the environmental data and the QC data (*qa_sep*) is run, so all data in the batch file is written to the QC database and flagged as ‘local only’ with the District.

3.8.5 Option 5 -- Process Batch File for All QA Samples

This option runs *qa_enter*, a program that *enters and updates* QA sample records with sample and result information from batch files. A record for a sample does not have to be in the database for the data to be successfully processed.

This program is identical to *qw_enter*, but is tailored to process a batch file containing only QC data. The user should ‘point’ to the appropriate QC database. The program to separate the environmental data and the QC data (*qa_sep*) is run, so all data in the batch file is written to the QC database and flagged as ‘local only’ with the District.

3.8.6 Option 6 -- Produce Batch Output

This option is used to produce *qwsample* and *qwresult* files from the database. The format of these files is described in [Appendix F](#). The program will use record numbers or agency code, station ID, date, time and medium code to generate the output file. This information can be entered from the terminal or from an existing file. The format of the

optional input files is described in [Appendix G](#). The output files can be specifically named or defaulted to *qwsample* and *qwresult*.

3.8.7 Option 7 -- Review / Edit Batch Files

This option is used to review or edit the *qwsample* and *qwresult* batch input files. These files are paired and both are necessary for the review/edit of the input files. When this option is selected, the following submenu is displayed:

```
QWDATA PROCESSING ROUTINE REV NWIS_4_0+20010328

YOU ARE USING WATER-QUALITY DATABASE NUMBER 01

Review/Edit Batch files

1 -- Review tab-delimited batch input files
2 -- Edit tab-delimited batch input files

98 -- Exit menu

99 -- Exit system

Please enter a number from the above list or a Unix command:
```

Option 1, to review the batch files, combines the sample and result information and displays it to the screen in increments of the number of lines specified. The user moves through the records with carriage returns, and enters a 'Q' to quit the review.

Option 2, to edit the batch files, will let the user edit either the sample information or the result information. The choices for editing the sample information are:

```
Transaction number: 1 Sample identifier: 315

You have several options:

1 -- edit this record
2 -- go to next record
3 -- quit and save
4 -- skip to another sample by transaction number
5 -- skip to another sample by sample identifier
Enter option 1-5, <CR> = 1:
```

The choices for editing result information are:

```
Transaction number: 1 Sample identifier: 315
Parameter code:00681
You have several options:
1 -- edit this record
2 -- go to next record
3 -- quit and save
4 -- skip to another sample by transaction number
5 -- skip to another sample by sample identifier
6 -- skip to another parameter in this sample
Enter option 1-6, <CR> = 1:
```

Carriage returns and/or the *cursor control commands* displayed at the bottom of the screen can be used to move through the various fields in both the sample and result screens.

Once the user has completed edits in the batch files, option 3 should be selected to quit and save the edited file. The screen display states the input has ended, shows the name of the edited file, and queries the user if there are other files to edit. A *Y* response returns the user to the screen for selecting samples or results to edit. An *N* response returns the user to the original submenu for reviewing or editing batch input files.

3.8.8 Option 8 -- Produce 1 - and * Card Output

This option is used to produce *qwcards* files from the database. The format of these files is described in [Appendix F](#). The program will use record numbers to generate the output file. The format of the input file is described in [Appendix G](#).

3.8.9 Reload QWDATA from Batch File, Overriding DQI

This option runs a program that *updates* sample records with sample and result information from batch files and overrides any DQI values ([Appendix A, Table 14](#)) that would have protected existing values from update. This program is a modification of the *qw_cardsin* program, so an entry for each sample (environmental and QC) must already exist in NWIS.

Generally, this program will be used to enter reloaded data from the laboratory or other large-scale changes to the water-quality data. The user will ‘point’ to the appropriate environmental database number. QC data will be separated and written to an associated QA database. This program can only be used by those with sufficient access rights (e.g. database administrators).

3.8.10 Reload QA Data from Batch File, Overriding DQI

This option runs a program that *updates* QA sample records with sample and result information from batch files and overrides any DQI values ([Appendix A, Table 14](#)) that would have protected existing values from update. This program is a modification of the qa_cardsin program, so an entry for each sample must already exist in NWIS.

Generally, this program will be used to enter large-scale changes to QC data only. The user will ‘point’ to the appropriate QA database number. This program can only be used by those with sufficient access rights (e.g. database administrators).