2   GETTING STARTED

2.1   Terminology

2.1.1   Sample

An environmental sample is the portion of a media (water, sediment, tissue, etc.) collected from a site for a specific date and time for analysis. A sample in the database is uniquely identified by the key variables, analyzing agency code, station number, begin date, begin time, end date, end time and medium code. The results of the chemical analyses and physical determinations for a sample are generally stored as one sample. These key variables are sometimes adjusted to separate samples that are hydrologically linked, but cannot be stored as one sample. Examples: Quality-control samples are stored as separate samples using one of the designated QC (Quality Control) medium codes. Samples from different depths or cross-sections may use time to 'off-set' samples that were collected at the same site on the same date.

2.1.2   Result

Values derived from chemical analyses or physical determinations of a sample are stored in the database as results. Results are uniquely identified in each sample by a parameter code that defines the property or constituent measured.

2.1.3   Parameter Code

Parameter Codes are five-digit codes used to identify the constituent measured and the units of measure. Some parameter code definitions include information about the methods used to measure the constituent, but this level of information is not currently consistent in the naming system.

2.1.4   Record

This is equivalent in the database to a single, uniquely identified sample.

2.1.5   Record Number

A record number is an eight-digit number assigned to a sample when the sample is logged into the database. The number is composed of a three-digit water year followed by a five-digit sequential number (Example: 00200075 is the 75th sample entered in water year 2002.) The record number is equivalent to the key variables that uniquely identify a sample, and can be used instead of the key variables in most of the NWIS QWDATA programs. These numbers are only unique within a database and must be used cautiously in an NWIS installation using multiple databases. When retrieving record numbers from the database, the file will contain the 8-digit record number followed by a two-digit database number.

2.1.6 Site

A site is the physical location where a sample is collected or a measurement is made. Information about a site must be stored in the NWIS SITEFILE before any data can be stored for that site. Equivalent terms: station.

2.1.7 Station Number

Station number is typically an eight-digit or 15-digit number assigned to a sampling site when it is established in the NWIS system (Novak, 1985). The eight-digit numbers are generally downstream order numbers used for surface-water locations on streams or rivers. Additional digits can be added to the eight-digit downstream order station numbers (up to a maximum of 15-digits) to achieve some sequencing. The 15-digit numbers are generally a combination of the latitude, longitude and sequence number of the location. A sequence number is used to separate locations that are very close to each other. Equivalent terms: site ID or station ID.

2.1.8 WATLIST

WATLIST is an output file produced by batch input programs that contain (a) a listing of the records that were updated, (b) a cation-anion balance table if the balance can be computed, and (c) a listing of any generated error messages.

2.1.9 Time definitions

In NWIS 4.3, QWDATA will require the entry of time datum during login of a sample. In an attempt to simplify the discussion within this documentation and between sub-systems of NWIS, the following definitions have been established:

Time zone - a geographic polygon where time is observed using a particular offset (in hours) from UTC (universal time).

Time datum - a variable that combines the time zone and the daylight-saving time usage associated with a time measurement.

Watch time - the time-of-day used by the person making and recording a measurement, or setting up an instrument to record measurements. The time-of-day recorded on field notes is converted to and from UTC within the database using a watch-time datum.

Site-default time datum - the time datum normally used by a USGS District office when recording time at a data collection site. The date of the time measurement must be taken into account to determine the site-default time datum, since daylight-saving time is (typically) only in effect during the summer. Note: the time-zone component of the site-default time datum may, or may not, be the same as the time zone in which the data-collection site is physically located. Similarly, the daylight-saving-time component of

the site-default time datum may not agree with the daylight-saving policy of the region surrounding the site.

Locally-observed time datum - the time datum that is commonly used by the populace in the vicinity of a data-collection site. (One might conceptualize this as the time datum used in the local television listings.) Conversion of a time measurement (in UTC) using the locally-observed time datum generates a measurement time-of-day that would be readily intelligible to a non-technical reader in the vicinity of the data-collection site.

An example is included to clarify the distinction between the site-default and the locally-observed time datums. A District in the Eastern time zone chooses to operate all gages in standard time year-round, even though the state where the district is located employs daylight-saving time. The site-default time datum in this case is EST, and the locally-observed time datum is EDT for summer-time dates.

2.2 Using the Interactive Programs

All interactive programs can be accessed from the main QWDATA menu. To start QWDATA type '**qwdata'** at the Unix prompt and the main menu will appear. Detailed information about each of the menu options within QWDATA is available in *Section 3* of this documentation.

2.2.1 Answering Prompts

When YES/NO questions are asked throughout the programs, an answer of "Y", "y", "YES", "yes", "N", "n", "NO", "no", or a blank will be accepted; any other answer should receive an error prompt and a repeat of the question. Similarly, where the user may answer a prompt with "QUIT", "Q", "q", or "quit" also will be accepted. Several programs request a numeric response to select an option (1, 2, or 3). When numeric data are requested, it is not necessary to enter final decimal points; however, imbedded decimal points must be entered.

2.2.2 Cursor Control

Cursor control screen movement allows the user to navigate from item to item, screen to screen, and to other parameters in the interactive programs. The valid cursor control characters and options are:

Enter	Action
^D	(Ctrl-D) Skip to next block
#	Delete (clear) value
/	Move back one field
/x	Continue at item number x
/+x	Move forward (x) items, default x is 1
/-x	Move back (x) items, default x is 1
/@	Continue at item with string @ in label
/p	Back up to previous page (screen)
/n	Advance to next page (screen)
/d	Delete current parameter
/a	Insert new parameter
/c	Cancel editing of current record
/q	Skip remaining items

To exit from the QWDATA system, the user can use Control-C. This will result in a query to be sure that the user wants to exit the software.

2.2.3 Default Values

A carriage return can be used to accept a default answer to a query in the interactive programs. The default answer is identified by `<CR> =' in parentheses at the end of the query. Default values for data entry are shown in the entry forms and will be used if the user does not specifically change them.

2.2.4 Mandatory Fields

Information that is required to store or retrieve data from the database is highlighted in the interactive screens.

2.2.5 Help with Valid Codes

A `**?**' may be entered in a field to obtain help on the information expected for that field. Where help is not available, a message stating that help is not available will appear.

2.2.6 Boolean Logic

In programs where criteria can be set to select sites, samples, or results, the default logic is for multiple criteria to all be 'true' for a site, sample or result to be selected. In some programs an option is provided to the user to specify that if any of the criteria are 'true' the site, sample, or result should be selected. In these programs the user selects either AND or OR as the logic. AND Boolean logic expects all criteria to be true. OR Boolean logic expects any criteria to be true.

2.2.7 Command Line Processing

A user can enter UNIX commands from most queries in the NWIS QWDATA programs and see the results of the UNIX command without leaving the NWIS software. The UNIX commands cannot be used if the user is positioned in a data entry field in the software. *NOTE: Some optional settings may not work when used from within QWDATA. An example is the 'rm' command. If this command requires a confirmation of the action outside of QWDATA, it will not require one from within QWDATA.*

2.3 Site Level Information

Site-level information includes data associated with the site being sampled. Most information for a site is stored in the NWIS SITEFILE that can be accessed from the Ground-Water Site Inventory (GWSI) System of NWIS. All documentation for GWSI is available at: *http://wwwnwis.er.usgs.gov/nwisdocs4_3/gw/GW.user.book.html*.

A site must be established in the SITEFILE before any water-quality information can be added. A site can be established using the GWSI entry programs as discussed in the GWSI user documentation.

Most basic site-level information can be retrieved from the SITEFILE through QWDATA. For example, if parameter codes 81024 (drainage area), 72000 (datum), or 72008 (well depth) are included in output and the values are not stored with the water-quality record, the software will retrieve the stored results for these parameter codes from the SITEFILE. Also some alphabetic parameter codes are available to retrieve site-level information from the SITEFILE through QWDATA. Alphabetic parameter codes are listed in *Appendix A.*

2.4 Sample Level Information

A sample in the database is uniquely identified by the key variables, analyzing agency code, site identifier code, begin date, begin time, end date, end time and medium code. The mandatory sample information that must be entered and the optional sample information that may be entered are described below.

2.4.1 Record Number (Automatic)

This is an eight-digit number assigned to a sample when the sample is logged in to the database. The number is composed of a three-digit water year followed by a five-digit sequential number (Example: 00200075 is the 75th sample entered in water year 2002.) The record number is equivalent to the key variables that uniquely identify a sample, and can be used instead of the key variables in most of the NWIS QWDATA programs. These numbers are only unique within a database and must be used cautiously in an NWIS installation using multiple databases.

2.4.2 Agency Code (Mandatory, Key Variable)

Agency Code is a five-character code that defines the agency responsible for the data stored for a sample. The default value is USGS_ (where '_' is a blank). The other options for agency code can be found in the NWIS file: /usr/opt/nwis/support/aanwdx.all.agency.

2.4.3 Station number (Mandatory, Key Variable)

Station number is typically an eight-digit or 15-digit number assigned to a sampling site when it is established in the NWIS system. The eight-digit numbers are generally downstream order numbers used for surface water locations on streams or rivers. The 15-digit numbers are generally a combination of the latitude, longitude and sequence number of the location. A sequence number is used to separate locations that are very close to each other.

2.4.4 Begin Date (Mandatory, Key Variable) and Begin Time (Optional, Key Variable)

Begin date is the date the sample was collected (or the begin date for a sample that is collected over multiple days). The format is YYYYMMDD.

Begin time is the time the sample was collected (or the begin time for a sample that is collected over a date or time range). The format is HHMM. Although this is an optional entry, it is one of the key variables to identify a sample. If the field is not populated, then a null will be used on retrieval.

2.4.5 End Date (Optional, Key Variable) and End Time (Optional, Key Variable)

The end date is the end date of a sample that is collected over multiple days. The format is YYYYMMDD.

The end time is the end time of a sample that is collected over a date or time range. The format is HHMM.

2.4.6 Time Datum

Time datum *(see section 2.1.9)* is a required code that represents the time zone and daylight saving time indicators that define the times entered in the begin date and end date fields. This field will be automatically populated with the settings in the SITEFILE for the sampling site entered, but can be changed to any valid time datum. A complete list of valid time datums is available in *Appendix J – Table 1*.

2.4.7 Time-Datum Reliability Code

Time-datum reliability code is a required code that represents the reliability of the time datum entered for a sample. This field will be automatically populated with 'K' (Known), but can be changed to 'E' (Estimated) if desired. Descriptions of valid time-datum reliability codes are available in *Appendix J – Table 2*.

The time-datum reliability code is used by QWDATA to determine when the time datum appears on some output formats. If the code is set to 'K', the time datum will appear on all output formats. If the code is set to 'T' or 'E', the time datum will not automatically appear. Review *section 3.4.3.4* for additional details.

2.4.8 Medium Code (Mandatory, Key Variable)

Medium code is the medium from which the sample was collected (i.e. - surface water, ground water, sediment, tissue). Medium codes 0-9 and A-P are used for environmental samples. Medium codes Q-Z are used for quality-control samples. A complete list of the valid medium codes can be found in *Appendix A, Table 1*.

2.4.9 Sample Type (Mandatory)

Sample type is the type of sample collected (i.e. - regular, composite, blank, spike). The default value for this code is `**9**' (Regular). The sample type codes can be found in *Appendix A, Table 5*.

2.4.10 Analysis Status (Mandatory)

Analysis status is the review status of the sample results (i.e. - initial entry, retrieved and in review, proprietary). The default value for this code is `**H**' (Initial Entry). The analysis status codes can be found in *Appendix A, Table 7*.

2.4.11 Hydrologic Condition (Mandatory)

Hydrologic condition is the hydrologic condition that is represented by the sample collected (i.e. - falling, rising, or peak stage). The default value for this code is `**A**' (Not Determined). The hydrologic condition codes can be found in *Appendix A, Table 3*.

2.4.12 Hydrologic Event (Mandatory)

A hydrologic event that is represented by the sample collected (i.e. - storm, drought, snowmelt). The default value for this code is `**9**' (Routine Sample). The hydrologic event codes can be found in *Appendix A*.

2.4.13 Analysis Source (Mandatory)

Analysis source identifies the source of the laboratory analyses and field data in the sample as USGS or non-USGS. The default value for this code is `**A**' (not determined). The analysis source codes can be found in *Appendix A, Table 8*.

2.4.14 Analysis Types (Optional)

Analysis types identify the types of analyses that were run for a sample (i.e. - Chemical, biological, nutrients). Zero to eight of the codes may be entered. The analysis type codes can be found in *Appendix A, Table 6*.

2.4.15 District Processing Status Code (Mandatory)

District Processing Status Code identifies the review status of a sample. Historically, this flag was used to identify samples that could be transferred to the national database. It still plays a role in the samples selected for the NWIS Web database. The district processing status codes can be found in *Appendix A, Table 9*.

2.4.16 Geologic Unit Code (Optional)

Interchangeably referred to as the aquifer code, this is an eight-character code that designates the aquifer associated with ground-water samples. The codes are defined in the "Catalog of Aquifer Names and Geologic Unit Codes." Option 4 of the Support Files sub-menu can be used to help find the appropriate code.

2.4.17 Lab Number (Optional)

Lab number is the analytical laboratory identification number given to a group of bottles from one field sample that are received together in a shipment.

2.4.18 Project Number (Optional)

Project Number is the nine-character project code associated with the sample.

2.4.19 Organism Code (Optional)

Organism code is the code to identify the organism of a tissue sample.

2.4.20 Body Part Code (Optional)

Body part code is the code to further qualify the tissue analyzed for a tissue sample. The body part codes can be found in *Appendix A, Table 19*.

2.4.21 Sample Field Comment (Optional)

Sample field comment is a text field to hold information from the field about a sample that cannot be defined by the coded information in the system.

2.4.22 Sample Lab Comment (Optional)

Sample lab comment is a text field to hold information from the lab about a sample that cannot be defined by the coded information in the system.

2.5 Result Level Information

2.5.1 Remark Code

Remark codes provide additional information about the magnitude (or absence) of a value. The remark code is almost always viewed with the value in the software to avoid misinterpretation of the value. The remark codes can be found in *Appendix A, Table 10*.

2.5.2 Value Qualifier Codes

Value Qualifier Codes provide information about the process used to determine an analytical value and, often, the remark code associated with the value. Up to three value qualifiers can be stored with any single result. The value qualifier codes can be found in *Appendix A, Table 16*.

2.5.3 Data Quality Indicator Code

Data Quality Indicator Codes indicate the review status of a result, controls the ability of a batch input program to overwrite a value, and affects the inclusion of a result in output. The data quality indicator codes can be found in *Appendix A, Table 14*.

2.5.4 Null Value Qualifier Code

Null Value Qualifier Codes identify a failed measurement due to field, lab, or shipment problem. The null value qualifier codes can be found in *Appendix A, Table 15*.

2.5.5 Laboratory Standard Deviation

Laboratory standard deviations are usually determined by the laboratory as an explanation of the uncertainty associated with a result value. This field was introduced in

QWDATA to provide the ability to round values on output using this information.

2.5.6 Method Code

Method codes identify the analytical method used to determine a value. In QWDATA 4.3, a table containing parameter codes and associated method codes was established and used to validate a data entry. More information about the development of the method code list can be obtained at:
http://wwwok.cr.usgs.gov/nawqa/phoenix/www/references.html.

2.5.7 Quality Assurance Code

Quality assurance code identifies the type of agency that did the field and laboratory work and the review status of the result. This code has only been sparsely used in the system (the majority of the entries are left at the default value) because the domain is confusing and there are other fields that can be set to convey this information. It will most likely be discontinued in a future revision. The quality assurance codes can be found in *Appendix A, Table 2*.

2.5.8 Preparatory and Analysis Dates

Preparatory and analysis dates are two fields used to identify the dates (YYYYMMDD) of the preparatory step and analysis at the laboratory.

2.5.9 Preparatory and Analysis Set Identifiers

Preparatory and analysis set identifiers are two fields used to store the set identification code (up to 12 characters) of the preparatory set and analysis set at the laboratory.

2.5.10 Report Level and Report Level Type Code

The report level is the numeric value associated with the analytical method when the result is determined. The report level type is a code to identify the type of report level used for the method. The report level type codes can be found in *Appendix A, Table 17*.

2.6 Parameter Codes

The parameter codes are five-digit codes used to identify the type of result stored. The codes are documented in the NWIS Parameter Code Dictionary (PCD). The PCD can be searched or listed from options in the Support Files sub-menu of the QWDATA menu. Selecting the appropriate parameter code may require some verification with the laboratory. The schedules run by the NWQL will contain the parameter code for the data that will be generated; therefore it may be useful to check their catalogue of schedules for help identifying the parameter code(s) needed. Only numeric values can be entered for any parameter code in the NWIS system.

2.6.1 Calculated Value Parameter Codes

Calculated values are results that may be stored in the database, or may be determined on output using algorithms stored in a reference file. The parameter codes for these calculated values are requested during retrieval. If the value is not stored; the parameter codes needed for calculation must be present. All calculated parameter codes available in an analysis can be retrieved by using of the alpha code – CALCV. A list of available calculated parameters, parameters required for calculation, and descriptions of special cases can be found in *Appendix D*. Algorithms for the calculated parameter codes can be viewed in QWDATA by use of option 8 in the *Support Files* menu.

2.6.2 District Specific Parameter Codes

Twenty-six parameter codes (99900-99925) are available in the PCD for use by individual Districts. The use of a District-specific code is to store District results for constituents that do not have a valid parameter code. Being stored in QWDATA allows the results to be stored, tabled, graphed, or exported for use in statistical programs. The first five codes (99900-99904) have default rounding codes of two, the second five (99905-99909) have default rounding codes of three, and the remaining codes (99910-99925) have default rounding codes of five. For more information about rounding codes, refer to *sections 2.7.1* and *3.6.7*. If these parameter codes are included in a publication-style table, the heading for each parameter contains seven characters per line and seven lines of text. The heading may be modified with an editor to accurately describe the constituent. Following is an example of District-specific code table output:

```
WATER-QUALITY DATA, WATER YEAR OCTOBER 1997 TO SEPTEMBER 1998


                                   9990001   9990201
                SPE-               9990002   9990202
                CIFIC              9990003   9990203
                CON-      TEMPER-  9990004   9990204
                DUCT-      ATURE   9990005   9990205
   DATE         ANCE       WATER   9990006   9990206
               (US/CM)    (DEG C)  9990007   9990207
               (00095)    (00010)  (99900)   (99902)


   SEP
    11...          109      10.5     70           .01
```

Because the definitions for these parameter codes are district specific and there is no storage of these definitions in QWDATA, it is recommended that a log of the definitions applied to these codes be kept by the database manager.

2.6.3 Fixed Value Codes

There are many parameter codes in QWDATA for which the result is not the outcome of an environmental measurement but rather a subjective observation or description. These types of parameter codes are referred to as Fixed Value Codes and have a small number of valid results that are related to a description. For example, parameter code 01035 – the severity of detergent suds – has the following valid results:

0	none
1	mild
2	moderate
3	serious
4	extreme

In other cases for which the result is not the outcome of an environmental measurement, valid results refer to a list that further defines the sample. For example:

Parameter code 00027 – collecting agency has a domain ranging from 300 to 99999 and each valid discrete result refers to a specific agency.

A complete list of parameter codes that have fixed values and the corresponding valid results can be found in *Appendix B*.

2.6.4 Alpha Parameter Codes

Alpha parameter codes, which may include both alpha and numeric characters, are used in QWDATA to retrieve sample and result-level descriptions as well as groupings of other codes. The descriptions include information stored with the sample such as date, time, and sample record number. An example of an alpha parameter code that represents a grouping of parameter codes is 'ADDPC' which adds all numeric codes to a retrieval. Another example of an alpha parameter code that represents a grouping of parameter codes is 'CALCV' which adds all calculated parameter codes to a retrieval. A complete list of alpha codes can be found in Table 18 of *Appendix A*.

For output in a by-result format, only alpha parameter codes are used to define the columns in the output. For more information on this type of retrieval refer to *section 3.4.4.*

2.7   Numeric Information

Values in QWDATA are stored in the database using double-precision floating-point representation, which yields 15-17 decimal digits of precision. The retrieval programs available through the QWDATA menus provide more limited precision. Precision of results produced through QWDATA is limited by the internally defined width of the output field, the rounding method, and rounding criteria selected. Even when unrounded

results are requested, precision is limited. In general, no more than eight digits are provided with the retrieval programs in QWDATA. Results used in calculations or retrieved through other avenues, such as Sequel Query Language (SQL) or Open Database Connectivity (ODBC) will be double precision.

2.7.1 Rounding

Water-quality values from NWIS may be output as rounded or unrounded. The unrounded choice will output the values as they were received from the lab and stored in the database. There are two choices for rounding values: (1) allow the software to determine how to round the result by using an existing *laboratory standard deviation (section 2.5.5)* stored with the result or the value in the rounding array for the specific parameter and method or (2) use a rounding code stored with each value.

Only one of these rounding options can be selected per retrieval. The first one is default rounding and the second one is called user-defined rounding. Some user-defined rounding codes have been set to odd values, sometimes by input from the lab, and may produce unexpected results. The most conservative and consistent option for rounding values at present is default rounding.

User-defined rounding can be used if a rounding code is stored with the result. If a rounding code is not stored with a result, the software will use the value from the parameter-method reference table. To add a rounding code, you can use *Modify Samples or Results* from the main QWDATA menu. Adding rounding codes can be done on a large scale using batch files and the QWFIX program as described in *section 3.9.3.*

Uncensored, non-zero concentration values that round to zero are in general, converted to a null value on output, with the remark set to "**M**" (constituent identified in sample, but not quantified). However, many uncensored measurements may be reported as zero and those values will remain zero in the output. An example of an acceptable rounded-to-zero result is water temperature. Censored, non-zero concentration values that round to zero are reported with the same number of significant figures that was originally entered with the value.

A rounding code of zero is not accepted by the system. New interactive entries with rounding codes of zero, (usually generated by the software) require the user to select a different rounding code before the value will be stored. New batch entries with rounding codes of zero will replace the zero with the number of significant figures in the actual value.

Programs within QWDATA present result values using the various rounding options. The table below summarizes the rounding options used by the various QWDATA programs:

QWDATA main menu option (option number)	Sub-menu option (option number)	Rounding Behavior
Login Sample (1)	---	result
Modify Samples or Results (2)	Enter Field Results (1)	result
	Enter Laboratory Results (2)	result
	Edit Sample or Results (3)	result
Data Review (3)	List Samples and Results (3)	unrounded
	Sample list and/or Cation-Anion Balance (4)	default
	Chemical Validation Checks (5)	default
Data Output (4)	Water-Quality Table by Sample (Publication Format) (3)	choice
	Water-Quality Table by Result (4)	choice
	Flat File by Sample (5)	choice
	Flat File by Result (6)	choice
	Make a P-STAT Data Set (7)	choice
Graphs (5)	Summary statistics table (7)	choice
	Detection Limits Table (8)	choice
Batch Processing (8)	Process batch file for all logged in samples (2)	unrounded
	Process batch file for all samples (3)	unrounded
	Process batch file for all logged in QA samples (4)	unrounded
	Process batch file for all QA samples (5)	unrounded
	Produce batch output (6)	unrounded

	Review/Edit batch files (7)
	    Review tab-delimited batch input files (1)
	    Edit tab-delimited batch input files (2)
		unrounded
		unrounded

	Produce 1- and *-card output	unrounded
	Reload QWDATA from batch file, overriding DQI (9)	unrounded
	Reload QADATA from batch file, overriding DQI (10)	unrounded

[Rounding Behavior: **result**, user-rounding based on rounding code stored with the result; **unrounded**, the value displayed is the stored value in the database; **default**, rounding is based on the rounding array stored in the parameter-method table; **choice**, user can choose what type of rounding will be applied.]

2.7.2 Scientific Notation

During input, the field for entering values in QWDATA is limited to eight digits and a decimal place. Scientific notation can be used to store values that won't fit in the field because the values are very small or very big numbers. The format for storing scientific notation is: the mantissa number, E, and the logarithm value. For example, .000000000023 would be stored as 2.3E-11, and 48,900,000,000 would be stored as 4.89E10. In some cases the PCD rounding specifications cannot accommodate very small and very large numbers on output. In this case, user-defined rounding can be selected for output.

2.7.3   Negative Values

Negative values may be stored in QWDATA for selected parameters. The types of parameters that can have negative values generally are field constituents (for example, air temperature, in degrees Celsius, parameter code 00020) and chemical constituents that are isotopes (for example, radium 226 dissolved, in pci/l, parameter code 09503). Some of the district-specific codes (99900-99909) also will allow negative values to be stored. A complete listing of the parameters that allow negative values is contained in *Appendix H*.

2.7.4   Zero Values

Zero values in the database are processed in different ways. The software accepts zeros during processing for all parameters. However, a zero may or may not be a valid value from a scientific point of view. For example, zero for a temperature is a valid measurement. Whereas, for chemical constituents, laboratory methods do not actually measure zero, but rather values are censored as being less than some laboratory reporting level.

For some chemical constituents, zero may have been incorrectly stored for historical data values. Users may censor these zero values on output. If the user selects "User Specified" for "Censoring of Zero Values" in the tabling options menu, a zero value will be converted to a null value and a remark code of U (material analyzed for, but not detected) will be reported. *Section 3.4.3.4* describes the tabling options in further detail. A reference list of parameters that are known to have zeros reported for some historical data are contained in *Appendix I*.

In some cases, the rounding specifications for a parameter in the PCD do not go to a decadal unit that is small enough for some result values and may result in the value rounding to zero. This is because laboratory methods are not accurate to that level, even though a laboratory measurement was made and reported. In this case, during output processing, a small number would incorrectly round to zero, even though a number may be stored. Values that belong to constituents that have reporting units that are used to quantify the abundance of an analytical constituent within some unit of area, volume, mass, or weight cannot be zero. Where the rounding procedures cause a non-zero value to round to zero, an 'M' (presence of material verified, but not quantified) will be displayed on output for parameters with the units contained in the table below. Values that belong to constituents that have reporting units other than those shown in this table can be reported as zero.

PG/KG	MG/ML	G/SQ M
PG/ML	MG/L	G/ SQ M
NANOGRAMS/LITER	MG/SQ M	G/KG
NG/G	MG / SQ M	UEQ/L
NG/L	MG/M2	MEQ/100 G
MICROGRAMS/LITER	MG/G	MEQ/L
UG/L	MG/KG	ML/L
UG/G	G/M2	PICOCURIES/GRAM
UG/KG	G/CU CM	PCI/G
UG/M2	G/CU M	PCI/L
MILLIGRAMS/LITER	G/SQ METER	PCI/ML

Reporting Units Where Zero Values Are Unacceptable

2.8 Text Fields for Comments

Text comments can be included within QWDATA records at the sample level and at the result level. Each of these two levels has two categories of comments, those from the laboratory and those from the field. In general, messages from the laboratory are sent directly from the laboratory and should not be edited or altered, although there are circumstances where this might be necessary. Messages from the field can be used to include information relative to the entire sample or to a specific result that cannot be characterized in another field. The recommended maximum length of these comment fields is 300 characters.

To retrieve these text fields, the following alpha parameter codes can be used when retrieving records:

Code	Meaning
SCMLB	Sample-level comment from lab
SCMFL	Sample-level comment from field
RCMLB	Result-level comment from lab
RCMFL	Result-level comment from field

2.9 Field Forms

Field forms are used by interactive entry of data and edit programs in QWDATA to provide a list of parameters that are routinely collected for a particular project, field trip, a District water-quality field sheet, or data from a non-USGS laboratory. Field forms are a list of parameter codes and associated information that are accessed by various programs so that a user can enter efficiently field or laboratory data. Some national field forms are provided with the NWIS release, but any number of District-specific field forms can be added to the system. Each field form is identified by a unique number that

is part of the field form name. Entry programs for field and laboratory data are discussed in *sections 3.2.1 – Enter Field Results, 3.2.2 – Enter Laboratory Results,* and *Tip Sheet 5.2*.

2.9.1 List of Field Forms Available

A list of field forms that are available can be viewed either by going to the directory where the forms reside or by listing them within QWDATA. To view the available forms in the directory where they reside, change directories to */usr/opt/nwis/data/auxdata/qw_field_forms/*. In this directory there are several files with the format *field.parmsnn.* The *'nn'* represents a two-digit number that identifies the form (e.g., field.parms01). Any editor can be used to view the contents of these ASCII files.

Another way to list the available field forms is within QWDATA during data entry. The prompt shown below appears during data entry after the login of a sample and when *Enter Field Results* or *Enter Laboratory Results* has been selected:

Enter field form nn, ?nn for detail of form nn, ? for list of forms available:

At this prompt, the user can enter a '?' and a list of available field forms will appear that includes the form number followed by either a form title or a list of parameter codes in that field form. To view the details of a specific form, enter a '?nn' at the prompt above, where 'nn' is the field form number of interest. A list of parameter codes and associated information will be displayed for that field form.

2.9.2 Adding a New Field Form

New field forms can be added to be used in QWDATA by creating a form using a specific format and storing it in the */usr/opt/nwis/data/auxdata/qw_field_forms* directory. New field forms should be created if none of the available field forms fit the data to be entered. Available field forms should be reviewed prior to creating a new field form so that duplication of field forms is avoided.

To make a new field form, the user should change directories to */usr/opt/nwis/data/auxdata/qw_field_forms.* Choose what field form number you are going to use for the new field form by listing the field forms in the directory. Field form names are in the format *field.parmsnn* where 'nn' is the two-digit number that identifies the form. Choose a field form number that is not already used. Use any editor that produces an ASCII output file to create or edit a field form file. The first line of the field form should be used to document the purpose of the field form by placing a "*#*" in the first column. Additional lines can be used for comments as long as a "*#*" is in the first column of the line. The format and an example of a field form are shown below:

Format of the field form

Line 1	Begins with a "#" in column 1 and is used to describe the purpose of the field form, name of the person who designed the form, and date created. (This line is not required but is strongly recommended because the line is displayed if a list of field forms is requested.)
Columns 1-5	Parameter code [five digits, use leading zeros].
Column 6	Method code [optional].
Column 7	Not used.
Columns 8-32	Parameter names or descriptions [could match District water-quality field sheet]
Column 36	Y indicates parameter is mandatory.

Example of field form

```
#NAWQA SW field form for Biological project was created
#  by John Smith on 12-25-1999
#2345678901234567890123456789012345678901234567890 Column numbers
00061  Streamflow
00065  Gage height
00010  Temp water
00020  Temp air
00400  pH
00095  Specific Conductance
00300  Dissolved Oxygen
00025  Barometric pressure
00452B CO3 Carbonate M=B
00453B HCO3 Bicarbonate M=B
39086B Alkalinity Inc. Titration
31625  Fecal Coliform
31673  Fecal strep
84164  SAMPLER TYPE              Y
71999  PURPOSE                   Y
99105  REPLICATE
99111  QA DATA TYPE              Y
```

Any numeric parameter can be entered into a field form. Alpha parameters (such as GUNIT for geologic unit code) cannot be used in the field form. Be sure that the final line of the field form contains a carriage return at the end. If this is not included, the last parameter will not be included to enter data. A maximum of 100 parameters can be included in any field form.

As part of creating new field forms, the user can create descriptions for the parameter codes listed. The user should be careful not to create descriptions that could cause confusion when compared to the definitions in the PCD. For example, a parameter described as 'chloride, dissolved' in the field form and defined as 'chloride, total' in the

PCD could cause incorrect data storage and interpretation, as well as significant future confusion. Creating a new field form is discussed in *Tip Sheet 5.3.*

2.10 Multiple Databases

Users have access to multiple water-quality databases where each database file is associated with a generic file name and a database number. A database number is a two-digit number, ranging from 01-98. (Database number 99 is reserved for use by NWIS maintenance programs and transaction archive files.) Regardless of how many databases are implemented on each system, any user who has not been assigned an alternate database number in the database selection table will be associated with database number 01.

2.10.1 When to Use an Alternate Database

Within NWIS, users have the option to create and use multiple water-quality databases. The decision to create and maintain an alternate database is one that should be carefully considered to ensure that the needs being met by an alternate database are sufficient to outweigh the problems associated with maintaining two or more databases.
Some of the issues involved with having more than one database are:

- duplicate data stored in multiple databases require that updates be applied to more than one database to maintain data integrity;
- multiple databases may lead to confusion about which database is primary when updating NWISWeb;
- an alternate database that is created for a specific project might be abandoned if the project chief leaves;
- alternate databases require increased data management responsibilities;
- interactive and batch entry results require setting the proper database number prior to NWIS use; and
- automated loading of National Water-Quality Laboratory (NWQL) results require setup and use of a separate NWQL user code;

Prior to the release of NWIS 4.1, one of the major reasons for the creation of alternate databases was to prevent the release of data for which there should be no public access; for example: proprietary data -- either permanent or temporary, cooperator data, and QA/QC data. With the release of the 4.1 revisions of QWDATA, there are codes available to restrict access within the same database. However, sample coding must be rigorously correct if proprietary and publicly accessible data are stored in the same database. If proprietary data are segregated into an alternate database, these data can be excluded safely and easily from unauthorized users including NWISWeb.

Other reasons for having an alternate database include testing, training, and organization. For example, much National Water Quality Assessment Program (NAWQA) data is kept in an alternate database because NAWQA study units generally are not aligned with District boundaries, and the primary responsibility for serving data to the public lies with the District that has jurisdictional responsibility for the site where the data were collected.

The NWIS Program office recommends that an alternate QWDATA database be used for storage of QA/QC and proprietary data. This will protect these data from accidental release to unauthorized sources in response to a public inquiry, or the inadvertent inclusion of the data in a report or on NWISWeb.

Districts should avoid the situation where multiple copies of the same sample data are stored in separate databases on the same NWIS installation, due to the difficulty of keeping the multiple copies consistently up-to-date.

2.11 Quality-Control Samples

The recommended procedures for storing quality-control (QC) samples in NWIS are described in this section. QC data provide information about the bias and variability in the environmental data, therefore the QC samples need to be identified in a manner that relates them to corresponding environmental samples. However, the QC samples need to be stored separately from the environmental samples to prevent any unintentional retrieval of this type of data. The recommended procedures address both of these requirements.

2.11.1 Definitions of Quality-Control Samples

Blank samples are taken to ensure that environmental samples have not been contaminated by the data collection process. Any measured value/signal in a blank sample for an analyte (a specific component measured in a chemical analysis) that was absent in the blank solution is believed to be due to contamination. There are many types of blank samples possible, each designed to segregate a different part of the overall data-collection process including sampling, filtering, preserving, storing, transporting, and analyzing:

> *Blank solution* - solution that is free of the analyte(s) of interest. Such a solution would be used to develop specific types of blank samples as described below.

> *Shelf (or hold) blank* - a blank solution put in the same type of bottle used for an environmental sample and stored adjacent to an environmental sample in a storage area.

> *Refrigerator blank* - a blank solution put in the same type of bottle used for an environmental sample and stored adjacent to an environmental sample in a refrigerated storage area.

> *Trip blank* - a blank solution put in the same type of bottle used for an environmental sample and kept with the set of sample bottles both before and after sample collection.

> *Sampler blank* - a blank solution poured or pumped through the same field sampler used for the collection of an environmental sample.

> *Filter blank* - a blank solution filtered in the same manner and through the same filter apparatus used for an environmental sample.

Splitter blank - a blank solution mixed and separated using a field splitter in the same manner and through the same apparatus used for an environmental sample.

Preservation blank - a blank solution treated with the same preservatives used for an environmental sample.

Field blank - a blank solution subjected to all aspects of sample collection, field processing, preservation, transportation, and laboratory handling as an environmental sample.

Equipment blank - a blank solution processed through all equipment used for collecting and processing an environmental sample (similar to a field blank but normally done in the more controlled conditions of the office and not being transported to the field).

Ambient blank - a blank solution put in the same type of bottle used for an environmental sample, kept with the set of sample bottles before sample collection, and opened at the site and exposed to the ambient conditions.

Source solution blank - a blank solution sent to a laboratory to confirm that it is free of the analyte of interest.

Lab blank - a blank solution prepared in the laboratory and analyzed the same as an environmental sample.

Blind sample - a sample submitted for analysis whose composition is known to the submitter but unknown to the analyst. A double blind sample is one of known composition that is submitted to the analyst in such a manner that neither its composition nor its identification as a check sample is known to the latter. A blind sample is one way to test the proficiency of a measurement process. Blind samples can be used to monitor the performance of an analytical system, check the analytical results of more than one laboratory, more than one analytical method, or the consistency of the same laboratory and method. Every blind sample analyzed should have an associated reference to the source and the possible dilution. Blind samples may be prepared from a reference material, as defined below.

Reference material - A material or substance, one or more properties of which are sufficiently well established, to be used for the assessment of a measurement method or for assigning values to materials.

Replicate (Duplicate) samples - A group of samples, collected in a manner such that the samples are thought to be essentially identical in composition. Replicate is the general case for which duplicates are the special case consisting of two samples. There are many types of replicate samples possible, each of which may yield slightly different results in a dynamic hydrologic setting, such as a flowing stream. The types of replicate samples are:

- ✓ *Concurrent samples* - samples collected by two or more people collecting samples simultaneously, or by one person alternating sub-samples between two or more collection bottles.

- ✓ *Sequential samples* - a type of replicate sample in which the samples are collected one after the other, typically over a short time.

✓ *Split sample* - a type of replicate sample in which a sample is split into sub-samples contemporaneous in time and space.

Spike sample - A sample to which known concentrations of specific analytes have been added in such a manner as to minimize the change in the matrix of the original sample. Every spiked sample analyzed should have an associated reference to the spike solution and the volume added.

Spike solution - A solution with one or more well established analyte concentrations that are added in known quantities to an environmental sample to form a spike sample.

2.11.2 Identification of Quality-Control Samples in NWIS

A system for identifying quality-assurance samples and maintaining the relationship with corresponding environmental samples has been established. The MEDIUM CODE, DATE, TIME and SAMPLE TYPE are used in the following manner to clearly identify all QA data:

Blank

- Use actual STATION NUMBER, DATE and TIME for blanks associated with a specific site and environmental sample. Use artificial STATION NUMBER, DATE and TIME for blanks that are associated with many sites and/or environmental samples. If more than one blank is analyzed, increment the TIME by 1 minute for each blank.

- Use a MEDIUM CODE of Q.

- Use a SAMPLE TYPE of 2.

- Use 991xx parameter codes listed below to designate which type of blank solution, the source of the blank solution, and the type of blank sample.

- If more than one blank is collected, use parameters 82073 and 82074 to enter the same STARTING TIME and ENDING TIME for each blank. Parameters 99109 and 99110 can be used to enter the SAMPLE SET START DATE and SAMPLE SET END DATE for a blank associated with environmental samples collected on more than 1 day.

Blind

- Use actual STATION NUMBER, DATE and TIME for blinds associated with a specific site and environmental sample. Use artificial STATION NUMBER, DATE and TIME for blinds associated with many sites and/or environmental samples. If more than one blind is analyzed, increment the TIME by 1 minute for each blind.

- Use a MEDIUM CODE from Q-Z.

- Use a SAMPLE TYPE of 4.

- Use 991xx parameter codes listed below to designate the source of the reference material and the source code number.

- If more than one blind is collected or if a blind is related to more than one environmental sample, use parameters 82073 and 82074 to enter the same STARTING TIME and ENDING TIME for each blind. Parameters 99109 and 99110 can be used to enter the SAMPLE SET START DATE and SAMPLE SET END DATE for blind samples that are associated with environmental samples collected on more than 1 day.

Reference Material

- Assign an artificial STATION NUMBER.

- Use the DATE and TIME that material was created or received.

- Select a MEDIUM CODE from Q-Z.

- Use a SAMPLE TYPE of 6 (new code). An existing SAMPLE TYPE of 3 (reference) has been used in the past, but there were no standards for its use. The new SAMPLE TYPE of 5 (reference material) was added and should be used for all samples that fit the definition of reference material included in this documentation.

- Use 991xx parameter codes listed below to designate the source of the reference material and the source code number.

Replicate

- Use the existing STATION NUMBER.

- Use actual DATE and TIME for the 1st sample and increment TIME by 1 minute (or actual times if greater than 1 minute) for each additional sample.

- Select a MEDIUM CODE from Q-Z.

- Use a SAMPLE TYPE of 7 for every sample, including the first one. A SAMPLE TYPE of 5 (duplicate) has been used in the past for the special case of replicates consisting of only two samples. The SAMPLE TYPE of 7 should be used for all samples that fit the definition of reference material included in this documentation.

- Use the 991xx parameter code listed below for replicates to designate which of the methods was used to create the replicates.

- Use the same STARTING TIME and ENDING TIME (parameters 82073 and 82074) for all samples.

- If the replicate samples will be stored in the regular database, use normal MEDIUM CODES (i.e., 6 or 9) and a sample type of 6.

Spike

- Use actual STATION NUMBER, DATE and TIME for spikes associated with a specific site and environmental sample. Use artificial STATION NUMBER, DATE and TIME for spikes that are associated with many sites and/or environmental samples. If more than one spike is analyzed, increment the TIME by 1 minute for each spike.

- Select a MEDIUM CODE from Q-Z.
- Use a SAMPLE TYPE of 1.
- Use the 991xx parameter codes listed below to designate the source code number of the spike solution, the spike type, and the volume of the spike. Use an existing code (32000) to designate the sample volume.
- If more than one spike is collected, use parameters 82073 and 82074 to enter the same STARTING TIME and ENDING TIME for each spike. Parameters 99109 and 99110 can be used to enter the SAMPLE SET START DATE and SAMPLE SET END DATE for blind samples that are associated with environmental samples collected on more than 1 day.

Spike Solution

- Assign an artificial STATION NUMBER.
- Use the DATE and TIME that source was created or received.
- Use a MEDIUM CODE of Q.
- Use a SAMPLE TYPE of 8.
- Use the 991xx parameter codes listed below to designate the source of the spike solution and the source code number of the spike solution.

A DISTRICT PROCESSING STATUS code of Z (local use only) should be assigned to each QA sample to prevent the data from being transferred to the national database and subsequently to the EPA's STORET database.

The following remark codes have been added to the system to allow further identification of some types of QA data:

A Average value

S Most probable value

Parameter code 99111, quality-assurance data indicator, can be stored with environmental samples to indicate that there are associated quality-assurance samples. The fixed values for 99110 indicate the type of quality-assurance samples associated with the environmental samples.

There are several existing parameter codes in addition to 82073 and 82074 mentioned in the identification of samples above that would clarify the nature of the sample (both environmental and QA samples). A list is provided below to encourage the use of these parameters. The definitions for the fixed-value parameters are in Appendix B.

00115 **Sample Treatment**
71999 **Sample Purpose - fixed value**
72005 **Sample Source - fixed value**
74200 **Sample Preservation - fixed value**
82075 **Amount of rinse, in liters**

      82398 Sampling Method - fixed value

      84164 Sampler Type - fixed value

 The parameter codes referenced in the sample identification are defined below:

        99100 Blank, Type of solution (fixed value)

        99101 Blank, Source of solution (fixed value)

        99102 Blank, Type of sample (fixed value)

        99103 Reference Material, Source (fixed value)

        99104 Reference Material or Spike Source, Code   Number

        99105 Replicate, Type (fixed value)

        99106 Spike, Type (fixed value)

        99107 Spike, Source (fixed value)

        99108 Spike Volume, in mL

        99109 Starting Date for a Set of Samples (YMDD)

        99110 Ending Date for a Set of Samples (YMDD)

        99111 Quality Assurance Data Type Associated with Sample (Fixed value code)

      1.00   No Associated QA Data

      10.00   Blank

      20.00   Blind Sample

      30.00

      40.00

      100.00 More than One Type of QA Sample

      110.00 Cross-section Information Stored

      200.00 Other

2.12 Data Entry Options

Two methods are available for entering data using QWDATA -- interactive data entry and batch data entry. Interactive data entry is entering data using queries and entry screens within the software. Batch data entry is using files as input to put data into the database. Both methods can be used to enter or modify data in the water-quality database if the user has the required security access. In general, interactive data entry is used to enter sample information (record creation), enter field data, and modify existing records for relatively small changes. Batch data entry is generally used for entering or modifying large amounts of data, for example the input of results from laboratories. Both methods can be used to remove data from the water-quality database.

2.12.1 Interactive Programs

Option 1 on the QWDATA main menu is used to interactively log samples into the database. This option can be used to enter sample information and field data. A *record number (see section 2.1.5)* is generated after the sample information has been entered. For specific information about using this program see *section 3.1 Option 1 – Login Samples*.

Option 2 on the QWDATA main menu contains options to enter field data, enter laboratory results, and edit samples or results. Option 2.1, Enter Field Data, allows a user to add field water-quality data to existing records in the QWDATA database. Option 2.2, Enter Laboratory Results, allows a user to add laboratory data to existing records in the QWDATA database. This option was primarily designed to enter data from non-USGS laboratories; field data can also be entered using this program. Option 2.3, Edit Samples or Results, can be used to edit and delete sample and result information for samples that exist in the QWDATA database. For more information on any of these options see *section 3.2 Option 2 – Modify Samples or Results.*

The entry of field or laboratory results, using any of the above interactive programs, is accomplished using field forms. These forms can be customized to provide parameters of interest for particular data entry needs. More information about field forms can be found in *section 2.9* and *Tip sheet 5.3.*

2.12.2 Batch Programs

Option 8 on the QWDATA main menu contains options to enter data using batch programs. The programs available from this menu work with files of water-quality data and are used for a variety of data needs. The most common application is entering data from laboratories, most typically the USGS NWQL. Other uses of the batch programs and files are output of entire records or groups of records by specifying the record information and making large scale changes to water-quality data in the QWDATA database.

Two batch file formats are accepted, a fixed format and a tab-delimited format. The fixed format, known as the '1 and *(or star)' card format, is a single file that was designed for use with NWIS versions prior to version 4.1. The tab-delimited format uses a pair of files, one for sample information and one for result information. This format was designed to accommodate new fields introduced in NWIS 4.1 and is expandable for fields introduced in future versions of the software. These two file formats are discussed in *section 3.8.* Both formats and example batch files are shown in *Appendix F*.

The programs available when using option 8 can retrieve NWQL results stored in a repository, process data into the database for existing and new records, produce batch output in either of the two formats, and review or edit a set of tab-delimited batch files. Detailed explanation of these programs is included in *section 3.8.*