Data Reporting Rules

The following is a brief summary of data reporting rules used by the MassDEP, Division of Watershed Management, Watershed Planning Program (WPP) for water chemistry data. This information has been excerpted from WPP standard operating procedures for data management (CN 000.83, CN 056.7). For additional and related information, see the Data Dictionary.

On-line Data File Format:
The WPP data files posted on the MassDEP web are in basic, "flat file"-type format, with columnar/header fields filtered to allow data views and searches to be defined by the user (e.g., by project, by watershed, by waterbody, by parameter, etc.).

QC Status Levels:

- QC1: Raw data. Generally not suitable for use, dissemination or publication.

- QC2: Preliminary data that has been entered into the appropriate WPP database and for which data entry proofing has taken place. Generally not suitable for use, dissemination or publication.

- QC3: Preliminary data for which technical QA/QC review and validation has taken place.   This may include evaluation of QC sample results, identification of outliers, comparison to project data quality objectives (DQOs), assessment of holding time exceedances, project-level review, etc. QC3 data are generally not suitable for use, dissemination or publication.

- QC4: Final Data. This QC level results from completion of the WPP data validation process, including technical and project-level review of accuracy, precision, representativeness, reasonableness, completeness and acceptability.   These data can be freely used and cited in documents without caution or caveat, adhering to the following guidelines:

    o Qualified data are reported with one or more qualifier symbols, usually under the "…SS" or "DWMQual" fields. Data qualifiers that are used to qualify data denote one or more specific, minor issues encountered during data validation that are considered by WPP to be insufficient cause for censoring, but that were otherwise noteworthy. These symbols are integral to the final results, and must be included when data are reported. Qualified data are generally usable, but data users must be cognizant of potential limitations depending on how data are used.

    o For censored data, numeric results are NOT reported, but the "##" symbol is applied to the result field, AND one or more qualifiers are used in the "…SS" or "DWMQual" fields to denote the major issues encountered during data validation that were considered by WPP to be sufficient justification for censoring.   Censored data are considered NOT usable for decision-making, but are included in data reports (using "##") for completeness.

- QC5: Final data that have been published (e.g., citable report, on-line, etc.). All information described under QC4 applies.

Data Qualifiers:
The following data qualifier symbols are used by WPP to denote qualified and censored data. These qualifiers are contained in the "…SS" and "DWMQual" fields in the data files, with the exception of the general qualifiers, which are presented in the results fields. For data that are qualified or censored by a laboratory, the lab-specific qualifiers are retained, but matched to the appropriate WPP qualifier for reporting purposes. Decisions regarding censoring vs. qualification for individual datum are made by WPP based on a thorough review of all pertinent information and meta data, including extent of exceedance of project DQOs. Dependent on the nature and extent of the quality control issue,

related sampling trip and/or laboratory batch sample results may also be qualified or censored. <u>As noted above, qualifier symbols are integral to the final results, and must be included when data are reported</u>. While qualified data are generally usable, censored data are considered NOT usable for decision-making.

General Qualifiers (applicable to all data types):

- " **--** " =   No data (i.e., data not collected/not required/not calibrated for/otherwise not available).

- "**##**" =   Censored data (i.e., data that have been discarded for some reason; check qualifier symbol for cause(s)).

- "******" =   Missing data (i.e., data that were intended to be collected but were not reported for any reason (e.g. site access issues, broken bottles, probe malfunction, etc.) other than no water).

- "**^^**" =   No water (i.e., a special case of missing data due to dry/no water conditions).

Probe-Specific Qualifiers:

- "**c**" =   unit not calibrated for a particular parameter and/or calibration standard(s) used for pre-calibration of the instrument was beyond the expected range.

- "**i**" =   inaccurate probe readings (based on field and/or lab QC data); may be due to significant pre-survey calibration problems, post-survey checks outside typical acceptance ranges for the low ionic and deionized water checks, lack of calibration of the depth sensor prior to use, or to checks against laboratory analyses.

- "**m**" =   method not followed; one or more WPP or manufacturer protocols not followed; may include operator error, instrument failure, lack of documentation, etc.

- "**r**" =   data may not be representative of the waterbody, due to circumstances and/or conditions at the time of sampling (e.g., out-of-water, side channel, backwater, etc.).

- "**s**" =   field sheet recorded data were used to accept data (i.e., data electronically recorded in a data logger were not available or hand-held probes were used where data logging was not possible).

- "**t**" =   tidal influence likely (not indicative of freshwater conditions).

- "**u**" =   unstable readings, due to lack of sufficient equilibration time prior to collecting final readings, non-representative conditions, highly-variable water quality conditions, etc.

Sample-Specific Qualifiers:

- "**a**" =   accuracy, as estimated using laboratory QC results, including matrix spikes, Proficiency Testing sample recoveries, internal check standards and lab-fortified blanks, did not meet project DQOs identified in a lab QA Plan or program/project Quality Assurance Project Plan (QAPP). Also used for Secchi depths where disk contacts lake bottom.

- "**b**" =   blank contamination in lab reagent blanks and/or field blank samples (indicating possible bias).

- "**d**" =   precision of field duplicates (as Relative Percent Difference (RPD)) did not meet project DQOs.

- "**e**" =   not theoretically possible. Specifically, used where E. coli bacteria counts (colonies per unit volume) exceed fecal coliform bacteria counts, where a numeric lake Secchi depth is greater than the reported station

depth, where a dissolved fraction is greater than the total (for the same elemental analysis and from the same sample bottle; e.g., orthophosphate vs. total phosphorus), and for other incongruous or conflicting results.

- "**f**" =   frequency of quality control samples collected did not meet expectations or requirements, as identified in a program/project QAPP.

- "**h**" =   holding time violation (indicating possible bias)

- "**j**" =   'estimated' value; used for lab-related issues where certain lab QC criteria are not met and re-testing is not possible. Lab-specific qualifiers are retained, but matched to the appropriate WPP qualifier for reporting purposes (e.g., a lab may use "B" for lab blank contamination, which would be changed to "b" by WPP for data reporting purposes).

- "**m**" =   field/lab method not followed with documented or potential effect(s) on data quality.

- "**p**" =   samples not preserved per required protocols or analytical method requirements.

- "**r**" =   data may not be representative due to circumstances and/or conditions at the time of sampling, including the possibility of "outlier" data.

- "**t**" = tidal influence likely (not indicative of freshwater conditions).

Numeric Field Values Involving Non-Detects:
For sample results (text format) less than the lower reporting limit, the numeric fields (e.g., nResult, nDO, etc.) show these values as negative the lower limit (e.g., <10 = -10). Significant figures are not maintained.

Numeric Field Values for Results that Exceed Upper Quantitation Limits (UQLs):
For sample results (text format) greater than the upper reporting limit or UQL, the numeric fields (e.g., nResult, nSecchi) show these values as the upper reporting limit (e.g., >2419 = 2419).

Rounding and Significant Figures:
Standard DWM-WPP rules for rounding and significant figures have been applied to the final data files as shown in the text-format fields. These rules are parameter-specific and depend on the magnitude of the numeric (lab, probe) results in relation to the lower reporting limits (e.g., Minimum Reporting Limit, MRL) and accuracy limits. (CN 056.7)

Depth Data:
In general, river and stream water sample data are collected from an approximate depth of 0.2 meters. All reported depth results (e.g., max depths, Secchi depths) are in units of meters.

Flow Condition, Lake Level and Ice-Cover Conditions
Qualitative flow and water level conditions listed under the "FLOWSTAT" field are based on visual field notes.