

Proposed Approach to Analyzing Censored Data Sets for Trends Analysis

Compiled by Laura Bexfield, with input and/or review by Wayne Lapham, Michael Rosen, Mike Rupert, Steve Grady, Dennis Helsel, Jeff Martin, and Mark Sandstrom, and others.

When performing analysis of data sets that include censored data, the NAWQA Ground-Water Trends Team (GWTT) advocates following the general guidelines spelled out in books by Helsel and Hirsch (2002) and Helsel (2005), as well as other papers and memoranda authored principally by Helsel, Hirsch, Cohn, and Gilliom with the USGS. In a nutshell, these references state that direct substitution methods should never be used—that is, censored data should never be replaced by some arbitrary number (whether it's 0, the detection limit, or some fraction of the detection limit) and then treated as “real” values for use in parametric methods of statistical analysis. Instead, the investigator should use either **maximum likelihood estimation** (if certain requirements are met) or **nonparametric methods** of data analysis (which may require replacement of censored values in a software package with some other selected value in order for ranks of the data to be calculated properly, depending upon the software and the laboratory reporting methods, as described below):

Maximum Likelihood Estimation

As described by Helsel (2005), maximum likelihood estimation (MLE) computes parameters of a data set based on the fitting of a distribution (generally lognormal for environmental data) to observed values above each detection limit and to the percentage of data below each detection limit. MLE works well for relatively large datasets where the data can be reliably shown to fit the particular distribution that will be applied. However, MLE performs poorly for data sets with fewer than 25 to 50 observations, or even for larger data sets where the data do not fit any particular distribution well. Therefore, many trends data sets might not meet the requirements for use of MLE. For those data sets where MLE seems like a good option, the ESTREND package of S-PLUS and macros available for Minitab 14 at www.practicalstats.com/nada will perform MLE computations. Helsel (2005) details the use of MLE methods in computing summary statistics and interval estimates, comparing groups, determining correlations, and performing regression and trends analysis.

Nonparametric Methods

Nonparametric methods of statistical analysis are likely to be the most broadly applicable methods for analysis of the trends data sets that are currently available. Nonparametric methods use the relative positions (ranks) of data in analysis, rather than using the actual data values or assuming a certain data distribution. In many cases, for statistical software to properly rank data in performing nonparametric methods, censored values must be replaced by some other selected value. Appropriate treatment of data in statistical software for analysis by nonparametric methods depends upon the laboratory method of data reporting for each constituent of interest:

Quantitative and Censored Results Only

When a constituent is reported only as either a quantitative result (equal to or greater than the laboratory reporting level (LRL)) or a censored result (less than the LRL), with no reporting of estimated values at a magnitude lower than the LRL, treatment of data is easiest. Helsel (2005) discusses some methods/software (in particular, macros for Minitab available at the above website) that can handle censored data sets without having to actually replace censored values with some other selected value. This software also can typically handle multiple LRLs. However, for investigators without access to this software, the GWTT advocates replacing all censored values with $\frac{1}{2}$ the highest LRL for all censored results and all quantitative results less than the highest LRL. This method will properly tie all results that are less than the highest LRL at a single rank below the quantitative results.

Quantitative, Censored, and Estimated Results

Selection of an appropriate method for treating censored data is more difficult when laboratories report quantitative, estimated, and censored results, as the NWQL does (fig. 1). Laboratories that use this type of reporting generally report results above a “quantitation limit” (equivalent to the NWQL’s LRL) as quantitative, results between the “quantitation limit” and the “detection limit” (equivalent to the NWQL’s long-term method detection level, or LT-MDL) as estimated, and results below the detection limit as censored. In an effort to minimize the probability of reporting false negatives, the NWQL in particular reports censored results (non-detections) as “<LRL” rather than “<LT-MDL” (Childress and others, 1999). Helsel (2005) describes 3 primary approaches to dealing with this type of laboratory reporting without introducing bias to a data set (table 1): **1.** set all censored and estimated results below the LRL to a single value less than the LRL, which results in some loss of information provided by the estimated values; **2.** set all censored results to a single value less than the LT-MDL, while retaining all estimated results, which effectively treats all estimated (semi-quantitative) values as quantitative values; or **3.** set all censored results to a single value less than the LT-MDL and all estimated results to a single value between the LT-MDL and the LRL. Helsel (2005) emphasizes that whenever estimated results between the LRL and the LT-MDL are retained in a data set, either quantitatively (as in **2.**) or semi-quantitatively (as in **3.**), the methods used by the NWQL require that all values reported as <LRL be treated as though they were <LT-MDL in order not to positively bias the data set. Childress and others (1999) likewise state that when using nonparametric methods to interpret data from the NWQL, a censored result must be assumed to fall somewhere in the range of 0 to the LT-MDL, rather than in the range of 0 to the LRL.

There is an additional factor that must be considered when working with NWQL data for “information-rich” analytical methods, which include the NAWQA pesticide schedules. For these analytical methods, the NWQL reports estimated results not only between the LRL and the LT-MDL, but also below the LT-MDL (fig. 1). If quality control criteria are met, the NWQL reports an estimated concentration for “all positively identified analytes” below the LT-MDL. There is an increased risk (greater than 1%) of false positives in this region, in addition to an even larger risk of false negatives. So, an important question is whether to treat estimated results in this region as quantitative, semi-quantitative, or non-quantitative (censored) results.

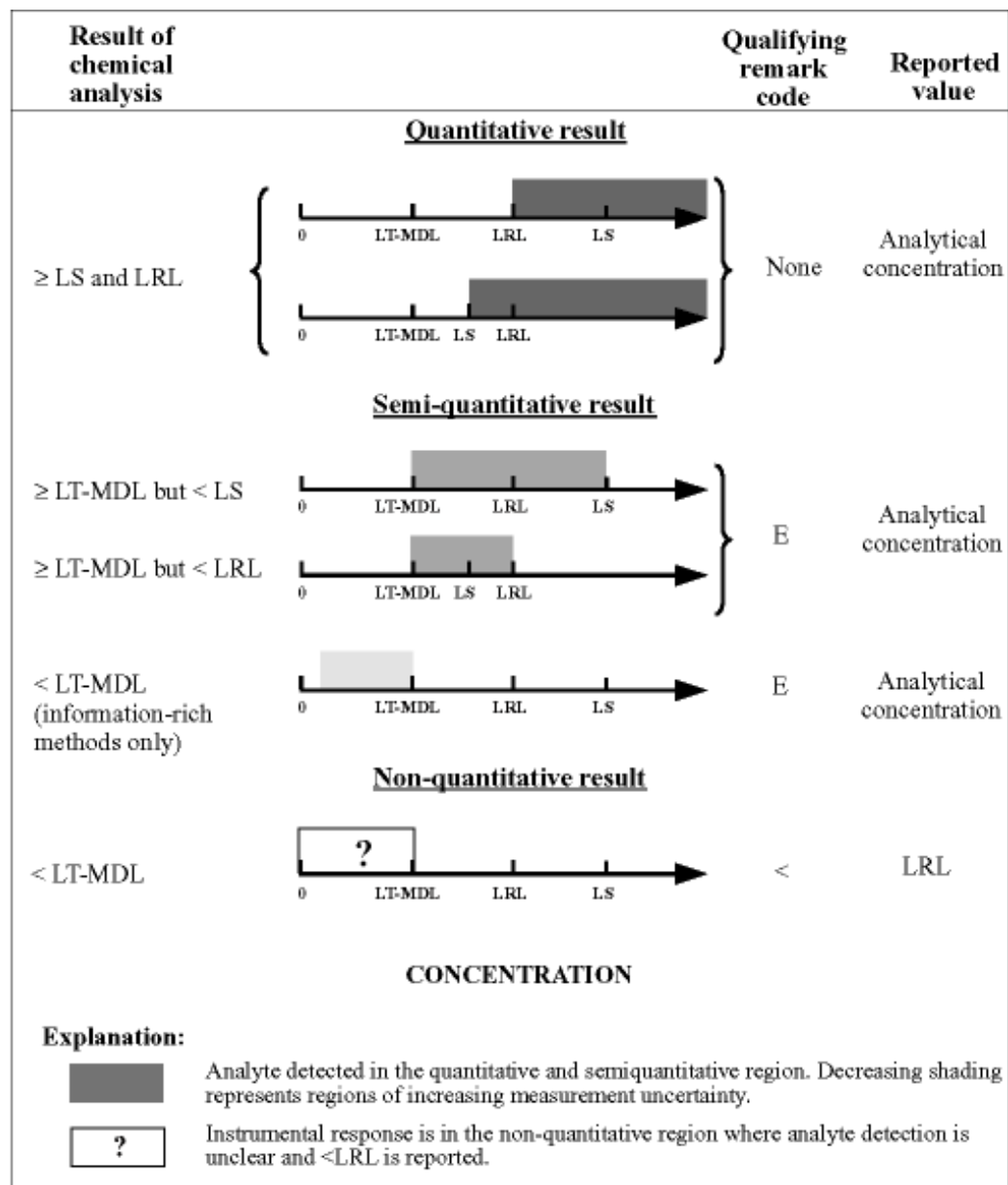


Figure 1. NWQL low-concentration reporting conventions showing the reported value and associated qualifying remark code in relation to the long-term method detection level (LT-MDL), the laboratory reporting level (LRL), and the lowest calibration standard (LS). [From Childress and others (1999); $\geq$, greater than or equal to; $<$, less than; E, estimated]

Table 1.—Summary of the three primary approaches to working with data from the NWQL without introducing bias, based on Helsel (2005).

Data approach	Quantitative result	Semi-quantitative (estimated) result*	Non-quantitative (censored) result
1.	No action	Replacement with a single value less than the LRL—the same value used for a censored result	Replacement with a single value less than the LRL—the same value used for an estimated result
2. (Recommended for trends analysis)	No action	Treatment as a quantitative result**	Replacement with a single value less than the lowest estimated value
3.	No action	Replacement with a single value between the LT-MDL and the LRL**	Replacement with a single value less than the LT-MDL

*This information does not apply to estimated results above the LRL

** If blank data indicate that results below the LT-MDL should not be treated as quantitative or semi-quantitative results and should instead be treated as censored results, another valid alternative would be to replace estimated results below the LT-MDL with a single value less the LT-MDL (the same value used for a censored result), while treating estimated results above the LT-MDL as indicated in the table

For trends analysis, we currently are advocating Helsel's (2005) data approach number **2.** (table 1), while treating all estimated results (even those below the LT-MDL) as quantitative. This approach retains the highest level of information provided by estimated concentrations, although these concentrations have greater uncertainty than quantitative concentrations above the LRL. For this approach, all censored values should be replaced with a single value less than the lowest estimated value but greater than zero (table 1). (Censored values could theoretically be replaced with zero, but zero values would cause problems for any transformations or graphs involving logarithms or reciprocals.) This approach should then allow statistical software to rank censored values lowest, followed by estimated values, and finally by quantified values above the LRL. Because for the NAWQA pesticide schedules the NWQL reports estimated results for every positively identified analyte below even the LT-MDL and censored values would be set below the lowest estimated result, changes in the LT-MDL are effectively invisible to this data approach (as are changes in the LRL and in the concentration of the lowest laboratory calibration standard relative to concentrations in environmental samples (fig. 1)). This consequence nevertheless appears to be consistent with the statement by Childress and others (1999) that for information-rich methods historical detections "are considered accurate even if the LRL is substantially higher than the historical MRL," where the MRL is the method reporting limit, which was generally equivalent to the LRL prior to adoption of the most recent NWQL reporting methods. However, by setting all censored values below all estimated values, the investigator is accepting a relatively high risk of false negatives.

Examples of the application of the recommended data approach and other approaches in table 1 to hypothetical NWQL data sets are given at the end of this document. Publications that include analysis/interpretation of such data sets should explicitly state the data approach that was used. Publications also should specify in the text, figures, and/or tables the detection levels that apply to each analyte, and should

state that there is less certainty associated with concentrations below these levels, along with less control over the likelihood of false negatives.

Caveats

Some common situations may prevent you from using data approach 2. described above (table 1). For example, if analysis of your blank data has led you to question the validity of estimated results below the LT-MDL as representative of the “real” presence of a compound in your environmental samples, you may want to treat these estimated results as equivalent to censored results by replacing them with the same value used to replace censored results (see the second footnote of table 1). For data sets spanning a time period that includes changing MRLs/LT-MDL’s, this approach would require determining an appropriate censoring level for each individual analyte by reviewing current and historical reporting level information provided by the NWQL on their internal website. Jeff Martin, working with NAWQA Pesticide National Synthesis, has used this approach in the past and found the highest historical LT-MDL to be an appropriate censoring level for most pesticides (written commun., 2004). Using this approach, all censored results and all estimated results below the selected censoring level would be assigned a single value below that censoring level. If your blank data lead you to question the validity of all estimated results (even those above the LT-MDL), at least in a quantitative sense, you may want to select between approaches 1. and 3. (table 1), which allow for the use of estimated results in a less quantitative way.

Other situations also would require use of a data approach besides 2. (table 1). One is a case where changes occurred in the NWQL reporting methods during the time range covered by your data. For example, if your earlier data are from older analytical and reporting methods (with only quantified and censored results), but your later data are from an “information-rich” analytical method using the new NWQL reporting method, you may need to set all censored results and all estimated and quantified results below the highest LRL to a single value less than the highest LRL before you analyze the data for trends. Also, be aware that all concentrations of some compounds have been reported by the NWQL with an “E” remark because of poor or variable recoveries. Interpretation of these data may need to be limited to a qualitative analysis.

Additional Information

To reiterate, if you need to replace censored values in your statistical software with some other value, these replacement values should never be used in any method of data analysis as though they were “actual” values. Instead, the replacement values should be used solely to allow statistical software to properly rank censored values for use in nonparametric methods of statistical analysis.

Helsel and Hirsch (2002) and particularly Helsel (2005) provide helpful discussions of the use of various nonparametric statistical methods for data sets with a wide range in the percentage of censored values. It is valuable to remember that highly censored data sets can be described and analyzed using methods that consider only categories of data (detected vs. non-detected, for example). Helsel (2005) also provides guidance on the proper presentation of boxplots and scatterplots for data sets with censored values, and he describes in detail a nonparametric method of estimating summary statistics (particularly means and percentiles) that he strongly recommends. This method, called the Kaplan-Meier method, is widely used in survival analysis and is

available in most statistical software, although use of the method on environmental data requires “flipping” the data set from left-censored to right-censored (see his book for instructions on this relatively simple procedure). Most other nonparametric methods of statistical analysis discussed in Helsel’s publications also are available in S-PLUS, Minitab (particularly with addition of macros from www.practicalstats.com/nada), and other commercially available software.

References

Childress, C.J.O, Foreman, W.T., Conner, B.F., and Maloney, T.J., 1999, New reporting procedures based on long-term method detection levels and some considerations for interpretations of water-quality data provided by the U.S. Geological Survey National Water Quality Laboratory: U.S. Geological Survey Open-File Report 99-193, 19 p. [Available at http://water.usgs.gov/owq/OFR_99-193/ofr99_193.pdf]

Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water sources: U.S. Geological Survey Techniques of Water Resources Investigations, Book 4, Chapter, A3, 512 p. [Available at <http://water.usgs.gov/pubs/twri/twri4a3/>]

Helsel, D.R., 2005, Nondetects and data analysis—Statistics for censored environmental data: New Jersey, Wiley-Interscience, 250 p. [This book is available to government employees at a 32% discount, with no shipping charges, directly from Wiley (877-762-2974). The total cost is \$61.17.]

Examples of applying various data approaches to hypothetical data sets

The following examples all indicate the values that would be used in software for performing statistical analysis for trends using **nonparametric** methods. The “replacement” values shown would never be used in any method of analysis as though they were actual values.

Example 1

Pesticide x—consistent reporting method for an “information-rich” analytical method, with LRL = 0.006 and LT-MDL = 0.003

Reported 1994 data: <0.006, <0.006, E0.002, E0.004, 0.006, 0.010

Reported 2002 data: <0.006, <0.006, <0.006, E0.001, E0.005, 0.015

Resulting values for analysis using nonparametric methods (approaches from table 1):

1. 1994 data: 0.003, 0.003, 0.003, 0.003, 0.006, 0.010
2002 data: 0.003, 0.003, 0.003, 0.003, 0.003, 0.015
2. 1994 data: 0.0005, 0.0005, 0.002, 0.004, 0.006, 0.010
2002 data: 0.0005, 0.0005, 0.0005, 0.001, 0.005, 0.015
3. 1994 data: 0.003, 0.003, 0.0045, 0.0045, 0.006, 0.010
2002 data: 0.003, 0.003, 0.003, 0.0045, 0.0045, 0.015

** Blank data indicate estimated values below the LT-MDL may not represent the “real” presence of the compound, but estimated values above the LT-MDL can be used quantitatively

1994 data: 0.003, 0.003, 0.003, 0.004, 0.006, 0.010

2002 data: 0.003, 0.003, 0.003, 0.003, 0.005, 0.015

Example 2

Pesticide y—consistent reporting method for an “information-rich” analytical method, but with 1994 MRL = 0.001, 2002 LRL = 0.007, and 2002 LT-MDL = 0.004

Reported 1994 data: <0.001, <0.001, 0.002, 0.004, 0.006, 0.010

Reported 2002 data: <0.007, <0.007, <0.007, E0.001, E0.005, 0.015

Resulting values for analysis using nonparametric methods (approaches from table 1):

1. 1994 data: 0.0035, 0.0035, 0.0035, 0.0035, 0.0035, 0.010
2002 data: 0.0035, 0.0035, 0.0035, 0.0035, 0.0035, 0.015
2. 1994 data: 0.0005, 0.0005, 0.002, 0.004, 0.006, 0.010
2002 data: 0.0005, 0.0005, 0.0005, 0.001, 0.005, 0.015
3. 1994 data: 0.002, 0.002, 0.0055, 0.0055, 0.0055, 0.010
2002 data: 0.002, 0.002, 0.002, 0.0055, 0.0055, 0.015

** Blank data indicate estimated values below the LT-MDL may not represent the “real” presence of the compound, but estimated values above the LT-MDL can be used quantitatively

1994 data: 0.002, 0.002, 0.002, 0.004, 0.006, 0.010

2002 data: 0.002, 0.002, 0.002, 0.002, 0.005, 0.015

Example 3

Pesticide z—consistent reporting method for an “information-rich” analytical method, but with 1994 MRL = 0.008, 2002 LRL = 0.005, and 2002 LT-MDL = 0.002

Reported 1994 data: <0.008, <0.008, E0.002, E0.004, E0.006, 0.010

Reported 2002 data: <0.005, <0.005, <0.005, E0.001, 0.005, 0.015

Resulting values for analysis using nonparametric methods (approaches from table 1):

1. 1994 data: 0.004, 0.004, 0.004, 0.004, 0.004, 0.010
2002 data: 0.004, 0.004, 0.004, 0.004, 0.004, 0.015
2. 1994 data: 0.0005, 0.0005, 0.002, 0.004, 0.006, 0.010
2002 data: 0.0005, 0.0005, 0.0005, 0.001, 0.005, 0.015
3. 1994 data: 0.001, 0.001, 0.0035, 0.0035, 0.0035, 0.010
2002 data: 0.001, 0.001, 0.001, 0.0035, 0.005, 0.015

** Blank data indicate estimated values below the LT-MDL may not represent the “real” presence of the compound, but estimated values above the LT-MDL can be used quantitatively

1994 data: 0.001, 0.001, 0.002, 0.004, 0.006, 0.010

2002 data: 0.001, 0.001, 0.001, 0.001, 0.005, 0.015

Example 4

Pesticide k—older data reported as either quantified or censored, newer data from an “information-rich” analytical method, with 1990 MRL = 0.010, 2002 LRL = 0.005, and 2002 LT-MDL = 0.002

Reported 1990 data: <0.010, <0.010, <0.010, <0.010, <0.010, 0.010

Reported 2002 data: <0.005, <0.005, <0.005, E0.001, 0.005, 0.015

Resulting values for analysis using nonparametric methods (approaches from table 1):

1. 1990 data: 0.005, 0.005, 0.005, 0.005, 0.005, 0.010
2002 data: 0.005, 0.005, 0.005, 0.005, 0.005, 0.015
2. not applicable
3. not applicable