Methods for Handling Non-detect or Censored Data

Frequently, groundwater monitoring data include results reported as 'nondetect' or values only known to be somewhere between zero and the reporting limit.

Interpretation of data sets containing several nondetect results (or left-censored data) can be a challenge. Fortunately, statistical methods and software to interpret censored data have evolved over the past 10 years, and several good options are now available.

Methods for evaluating censored data include:
1) Substituting a value for non-detect results (i.e. ½ detection limit, lowest detection limit, etc.). This is the most common approach, but can lead to erroneous estimates of summary statistics (so called *invasive patterns*) in cases with a high proportion of nondetect results.
2) Cohen's method (1959)/Aitchison's method (1995) – not much better than substitution.
3) Maximum likelihood estimation (MLE) and regression on order statistics (ROS) are the methods most often cited in the literature for handling nondetects without introducing too much bias.

Dennis Helsel of the USGS has written extensively on handling non-detect values in environmental and water quality data sets. His book *Nondetects and Data Analysis: Statistics for Censored Environmental Data* (Helsel, 2005), describes the best approaches to handling nondetect data. Helsel's web site offers several resources for approaching this issue: http://www.practicalstats.com/nada/

Recently, the USEPA has sponsored the development of excellent free software for performing summary statistics (including MLE and ROS methods for handling nondetects). The software is available at: http://www.epa.gov/esd/tsc/software.htm The accompanying report to this software provides a good description of appropriate statistical methods for handling nondetects (Singh, 2006).

Some Additional References:

1.   Lee, L.A. and D. Helsel, *Baseline Models of Trace Elements in Major Aquifers of the United States.* Applied Geochemistry, 2005. **20**: p. 1560-1570.
2.   Helsel, D.R., *Fabricating Data: How Substituting Values for Nondetects Can Ruin Results and What Can Be Done About It.* Chemosphere, 2006. **65**: p. 2434-2439.
3.   Helsel, D., *Insider Censoring: Distortion of Data with Nondetects.* Human and Ecological Risk Assessment, 2005. **11**: p. 1127-1137.
4.   Helsel, D. *Nondetects and Data Analysis: Statistics for Censored Environmental Data.* 2005. Wiley.
5.   Helsel, D.R., *More Than Obvious: Better Methods for Interpreting Nondetect Data.* Environmental Science & Technology, 2005. **39**(20): p. 419A-423A.
6.   Shumway, R.H., R. Azari, and M. Kayhanian, *Statistical Approaches to Estimating Mean Water Quality Concentrations with Detection Limits.* Environmental Science and Technology, 2002. **36**: p. 3345-3353.
7.   Singh, A.S., *On the Computation of a 95% Upper Confidence Limit of the Unknown Population Mean Based Upon Data Sets with Below Detection Limit Observations.* 2006, U.S. Environmental Protection Agency, Office of Research and Development: Washington, D.C.