White Paper

Integrated Report Improvements

Use of Censored Data

Date: May 11, 2018

State of Oregon
Department of
Environmental
Quality

Water Quality Standards and Assessments
700 NE Multnomah St.
Suite 600
Portland, OR 97232
Phone: 503-229-5696
800-452-4011
Fax: 503-229-5850
Contact:
Becky Anthony

www.oregon.gov/DEQ

DEQ is a leader in restoring, maintaining and enhancing the quality of Oregon's air, land and water.

Introduction and background

Due to limitations in field and laboratory chemical analysis procedures, small concentrations of some substances cannot be precisely measured. Analytical test procedures typically have both a Method Detection Level (MDL) and a Minimum Reporting Level (MRL). The MDL is essentially the concentration at which a sample can be discerned from a sample blank, while the MRL is the lowest concentration where an analyte can be both detected and an accurate concentration quantified. Both values are laboratory and instrument dependent and can be significantly different for the same analyte. There is not a consistent way labs are required to report this data. For example, some labs will report to the MRL while others report to the MDL. For this analysis, DEQ will use the generic term Quantitation Limit (QL) to include MRL, MDL and any other reporting limits used by third parties.

Sample concentrations below the QL are referred to as censored values and, depending on the laboratory, are reported in several different ways. The most common method for reporting censored data is to use non-detect (ND) or < the value of the QL being used.

In past Integrated Report (IR) cycles, DEQ eliminated censored data from its data analysis. However, revisions to the assessment methodology now favor use of more robust data sets for making both listing and delisting decisions regarding beneficial use attainment. With the adoption of the exact binomial test statistical method for toxic substances and conventional pollutants for the protection of aquatic life, and the calculation of the geometric mean to apply to the human health criteria, DEQ must now define a method for the use of censored data in the IR. This paper describes different ways of treating censored data in the IR and identifies a preferred method for future assessments. The objective of the method is to reduce bias and develop an approach that is applicable and reproducible for all parameters..

Methods used by other states

DEQ reviewed several states IR methodologies for comparison on how censored data are treated in their respective assessments (Table 1). Methods varied by state and the treatment of censored data was often tied to whether numeric water quality criteria are above or below the QL. The most common methods for using censored data in the IR are summarized below:

1. Set values to zero
2. Eliminate non-detect results when water quality criteria is <QL
3. Set the value to the QL
4. Substitute to ½ the QL
5. Substitute to ½ the water quality criteria value
6. Substitute the lowest value of either ½ the QL or ½ the water quality criteria
7. Assign value to non-detect results, based on proportion of non-detect samples

Table 1. Comparison of states use of censored data in assessment methodologies.

State	Use of Censored Data
CO	Replace ND with zero. In cases of drastically different MDL or ML values possible unintentionally biased toward zero. May be appropriate to omit the dataset with a higher MDL or ML from the assessment of the data.
CA	When available data are less than or equal to the quantitation limit and the quantitation limit is less than or equal to the water quality standard, the value will be considered as meeting the water quality standard, objective, criterion, or evaluation guideline. When the sample value is less than the quantitation limit and the quantitation limit is greater than the water quality standard, objective, criterion, or evaluation guideline, the result shall not be used in the analysis.
FL	The Department shall treat any result less than the MDL of the method ... as being one half the MDL (if the criterion equals or exceeds the MDL) or one half the criterion (if the criterion is less than the MDL), for any pollutant. For criteria that are expressed as averages, which ever of the following measurements is smaller is used in calculating the average—half of the LOQ or half of the criterion. For values expressed as greater than the LOQ, the whole value is used.
ID	If censored measurements comprise less than 50% of the measurements of an analyte and the data set appears to be parametrically distributed (either normal, lognormal, or gamma), then the statistical parameters of the distribution are best inferred using distributional methods such as the maximum likelihood estimator (e.g., Helsel 1990; 2005; the utilities available in ProUCL 5.0 [EPA 2013a]) are recommended for such situations. If censored measurements comprise more than 50% of the data set, nonparametric analysis is generally preferred unless special circumstances apply (EPA 2009); in that case multiple methods for estimating the distribution's parameters should be evaluated, including a sensitivity analysis of the results, before deciding on the best outcome. In special cases, such as where the nondetect percentage is very high, DEQ may approve alternative methods for handling censored data on a case-by-case basis.
MN	Values below the level of detection, even if greater than the standard, will not be considered an exceedance of the standard. Values below the level of detection will be considered a data point for the purposes of meeting the minimum data requirement. For calculating geomeans: Value assigned to "less thans" = LOD [1 – (Number of values < LOD / Total number of values)] Where LOD = level of detection
MT	Nutrients - Convert non-detects in the dataset to 50% of reported detection limit; if > 15% of dataset is non-detect, consult WQPB Standards Section. Metals - Include non-detects in the dataset if the water quality standard (WQS) is higher than the laboratory detection limit for that metal parameter.
NE	Rather than eliminating the "non detects" from the assessment data, values measured below detection limits will be calculated as 50% of the method detection limit. This approach may not be appropriate during the analysis of water quality trends.
NM	Results from samples that are flagged by the laboratory as "below the minimum quantification or reporting limit" (generally referred to as minimum reporting limit or MRL) may only be used during the assessment process if the MRL is less than the applicable water quality criterion (WQC) or numeric threshold being assessed.

NH	When nondetect values were reported and an actual value was needed for making an assessment, 50 percent of the analytical detection limit was used as the value. For bacteria results reported as "0" counts, the zero values were replaced with 1 counts so that the geometric mean could be calculated. Care has been taken to ensure that waters were not listed based upon values below the detection limit where the detection limit was greater than the standard criteria. Results of some water quality samples are reported as above the analytical detection limit. In such cases, the actual value is not known. When detection limit exceedance values were reported and an actual value was needed for making an assessment, maximum detection limit of the analytical detection limit was used as the value. For example, bacteria results reported as ">2000" counts, were replaced with 2000 counts so that the geometric mean could be calculated.
NV	For development of the 2014 Integrated Report, samples with pollutant concentrations reported "as less than the detection limit" were assumed to comply with the water quality standards if the certified laboratory method is acceptable to NDEP, and no other information indicated that the substance in question existed in levels detrimental to the beneficial uses. For those water quality criteria requiring calculations, such as annual average or geometric mean, samples with values reported as below detection limit were included in the calculation at one-half of the detection limit.
TX	For criteria that are expressed as averages, including chronic toxicants (aquatic life use), bacterial indicators geometric mean (recreation use), human health criteria for water (fish consumption use), and primary organic substances (public water supply use), which ever of the following measurements is smaller is used in calculating the average—half of the LOQ or half of the criterion. For values expressed as greater than the LOQ, the whole value is used. When most of the reported values for a parameter are less than the LOQ, and the LOQ is significantly greater than the criterion (note that a margin of safety of about two for aquatic life and five to ten is incorporated into criteria), the samples are not used for calculation of averages or percent exceedances. A status of Not Assessed may be identified, rather than fully supporting or no concern. The assessor will use judgment when identifying parameters as fully supporting or delisting when the dataset includes nondetects.
UT	For sample results below detection, the reported result value or a value of 0.5 times the lowest reported detection limit is applied for purposes of the assessment. However, if one-half of the detection limit is above the water quality standard, the data will not be used in the assessment.
WA	Non-detect sample values will be considered in the assessment, but can only be used to show compliance with water quality criteria when the detection limit is less than the criteria. For calculating a geometric mean using non-detect samples, in which a zero cannot be used, a value will be chosen so as not to bias the geometric mean high or low.

Analysis of methods for using censored data in Oregon

DEQ conducted an analysis of five different methods of treating censored data for three different parameters; *E. coli*, lead and 4,4, DDT. These parameters differ in the relationship of the average QL to the numeric water quality criteria. Aquatic Life uses were assessed using the exact binomial test method and human health uses were assessed using the a geometric mean of the entire dataset following their respective methodologies. The geometric mean is defined mathematically as the n-th root of the product of n numbers. This analysis was done on a subset of available data at the site level and does not reflect actual expected conclusions for the 2018 IR.

Table 2. Comparison of methods for treating censored data in assessing water bodies for *E. coli* for human health uses. Water quality criteria (126 MPN/406 MPN) higher than QL (1 MPN).

	Number of Sites Assessed	Number of Sites in Cat. 3: Insufficient Data	Number of Sites in Cat. 3B: Insufficient Data – Exceedances	Number of Sites Listed (Cat. 5)
Remove from analysis	412	253	76	57
Use QL	699	537	76	57
Use 1/2 QL	699	537	76	57
Set to 0	699	NA	NA	NA
Use 1/2 WQ Criteria	699	537	76	57

"NA" indicates a geometric mean cannot be calculated based on a zero value.

Table 3. Comparison of methods for treating censored data in assessing water bodies for lead for aquatic life uses. Range of calculated water quality criteria (1.4 to 21.1 µg/L) near the range of QLs (0.02 to 1µg/L) .

	Number of Sites Assessed	Number of Sites in Cat. 3B: Insufficient Data – Exceedances	Number of Sites Listed (Cat. 5)
Remove from analysis	141	0	1
Use QL	184	1	1
Use 1/2 QL	184	0	1
Set to 0	184	0	1
Use 1/2 WQ Criteria	184	0	1

Table 4. Comparison of methods for treating censored data in assessing water bodies for 4,4, DDT for human health and aquatic life uses. Water quality criteria (Aquatic Life Chronic 0.001 µg/L and Human Health 0.000022 µg/L) below the average QL (0.0283 µg/L).

Method	Number of Sites Assessed		Percent Samples exceeding		Number of AUs Listed		Percent Sites Listed	
	Aquatic Life (Chronic)	Human Health	Aquatic Life (Chronic)	Human Health	Aquatic Life (Chronic)	Human Health	Aquatic Life (Chronic)	Human Health
Remove from analysis	6	6	100	100	6	6	100	100
Use QL	245	245	100	100	245	245	100	100
Use 1/2 QL	245	245	100	100	245	245	100	100
Set to 0	245	NA	0.49	NA	4	NA	1.63	NA
Use 1/2 WQ Criteria	245	245	0.49	0.49	4	4	1.63	1.63

This analysis shows that the method which generates the largest change in the categorical designation of sites is removing censored data from the analysis. Using Oregon's exact binomial test procedures, excluding censored results from the assessment would lower the effective sample size for the assessment unit. This would negate the benefits of the exact binomial test

and increase the probability of falsely concluding a waterbody exceeds the standard, resulting in erroneous listings. Using the 4,4, DDT example, this method considered results for only six locations, with all six locations assessed as a Category 5 (303(d) listing). The remaining 239 sampled locations would be excluded from assessment altogether. Setting results to zero preserves the advantages of the exact binomial test procedures for assessment of aquatic life uses by maintaining the actual sample size and listing procedures at accepted confidence levels, but a geometric mean (needed for assessment of human health uses) can not be calculated with a zero value. Using the QL or ½ the QL poses an issue when the water quality criterion is below common QLs. In the 4,4, DDT example, using either of these methods results in assigning all of the assessment units with data as Category 5 for both aquatic life and human health uses. This bias towards type II false-negative errors would be an inefficient use of TMDL resources. In addition, this method would likely reduce the incentive for collecting data for parameters that have numeric water quality criteria below or close to common QLs. Using ½ the criteria provides a non-zero value that can be used to calculate a geometric mean for human health uses and will reduce type II error (false-positive) when water quality criteria values are less than the QL.

New method for using censored data in the assessment

For water bodies with no quantifiable sample results:

- Water bodies will be assessed as *Category 2; Attaining* where samples have been collected but all values are reported below the lowest available QL and the QL is less than the numeric criteria.
- Water bodies will be assessed as *Category 3D; Not Technologically Feasible to Assess* where samples have been collected but all values are reported below the lowest available QL, and the QL is greater than the numeric criteria

For water bodies with a mix of quantifiable and censored data, DEQ will use the following methods for the application of the exact binomial test statistical method and the calculation of the geometric mean to apply to the human health criteria.

- When the QL is greater than the numeric criteria value, ½ of the value of the water quality criteria will be substituted for any sample reported as censored.
- When the QL is less than the numeric criteria, ½ of the value of the QL will be substituted for any sample reported as censored.
- Samples reported as greater than the Maximum QL use value.
    - For example, a bacteria sample reported as >2000 MPN, 2000 MPN will be used.

Through the IR assessment process, DEQ will be accepting data from multiple sources which may include different laboratory reporting limits. Sample concentrations measured between the MRL and the MDL are often reported as an estimated value, because the precision of the method is not enough to determine the exact concentration. For samples reported as estimated, DEQ will use the value and assign an assessment category based on these rules:

- When the QL is the less than the numeric criteria and an impairment determination is based on solely estimated or a combination of estimated and quantifiable results, water bodies will be assessed as Category 3B when quantifiable results alone do not indicate impairment.
    - In cases with drastically different QL values, it may be appropriate to omit the portion of the dataset with a higher QL from the assessment of the data.)
- When the QL is greater than the numeric criteria water bodies will be assessed using the estimated values.

Conclusion

DEQ is updating the method used to treat censored data in the assessment process. This is a change from the method used in 2010 and 2012 in which censored data was removed from the analysis. The new method minimizes type I false-positive and type II false-negative errors. False-positive errors are reduced because censored data will never result in an exceedance of a water quality criteria and can only be used to show attainment with numeric criteria when the QL is less than the criteria. Additionally, the method allows for the ability to calculate a geometric mean and maintains the actual sample size, preserving the statistical significance of the exact binomial test procedures. False-negative errors are reduced by the addition of a new assessment category, *3D: Not Technologically Feasible to Assess: Insufficient data to determine use support because numeric criteria are less than quantitation limits*. With this additional category, DEQ avoids classiyfing water bodies as *Category 2: Attaining* when current technology can not achieve quantifiable results at or near numeric water quality criteria values.

Alternative formats

Documents can be provided upon request in an alternate format for individuals with disabilities or in a language other than English for people with limited English skills. To request a document in another format or language, call DEQ in Portland at 503-229-5696, or toll-free in Oregon at 1-800-452-4011, ext. 5696; or email deqinfo@deq.state.or.us.