FYI: [illegible]

Return to RC

WISCONSIN STATE LABORATORY OF HYGIENE

465 Henry Mall
Madison, WI 53706-1578
Phone: (608) 262-1293
FAX: (608) 262-3257

University of Wisconsin
Center for Health Sciences

[DNR-128]

Water Resources Center
University of Wisconsin - MSN
1975 Willow Drive
Madison, WI 53706

MEMORANDUM

RECEIVED - DNR
AUG 19 1997
BUREAU OF DRINKING WATER & GROUNDWATER

DATE: August 15, 1997

TO: Randell Clark

FROM: William C. Sonzogni

RE: Final Report- Ground Water Research Program on Improved Detection Limits for Ground Water Monitoring

CC: Ron Arneson, Mike Zorn

Please find enclosed a final report on our research project. We believe the project has been very successful. We still have some more research we expect to finish up in the next few months. It will be included in Michael Zorn's Ph.D. thesis, and I will be sure to make that information available to the Groundwater Coordinating Council and the Department.

On behalf of Mike and I, I want to thank the Groundwater Coordinating Council and the Department for sponsoring our research. We hope we have made some contributions that will result in an appreciable return on the investment.

051204

Improved Detection Limits for Ground Water Monitoring

Final Report

This study was designed to improve the detection of trace contaminants in ground water. The objectives were to (1) develop a new approach to concentrating sample analytes (while minimizing interferences) for the purpose of reducing detection limits, and (2) apply a more rigorous, but practicable, computational method for the analytical limit of detection.

We have accomplished our first objective by developing an on-line supercritical fluid extraction/gas chromatographic (SFE/GC) technique for determining trace quantities of pesticides and PCB's in water. This procedure is faster, cleaner, and cheaper than the conventional Soxhlet extraction/gas chromatographic technique. Also, the on-line SFE/GC method should realize gains in sensitivity of three orders of magnitude, resulting in greatly improved detection limits, much smaller sample size requirements, or both.

One of the developments in this work is a procedure for physically handling and transferring the XAD-2 resin at a very small scale. Rather than passing 80 L-160 L of water through 125 g of resin, the developed procedure uses 50 to 200 mL of water passed through approximately 350 mg of resin. The water sample is passed through a stainless steel extraction column where the analytes are retained by the XAD-2. The water remaining in the extraction column is removed prior to the SFE/GC analysis by passing high purity nitrogen through the sample at 50°C. The time required for the nitrogen drying process to quantitatively extract the PCB's from the XAD-2 resin still needs to be optimized.

Experiments designed to recover a PCB standard spiked onto an inert matrix yielded near quantitative recovery (between 85 and 115 percent). This suggests that the analytes are being quantitatively trapped for subsequent gas chromatographic analysis. Experiments involving a PCB standard spiked onto a clean sample of XAD-2 adsorbent resin have also yielded near quantitative recovery. Extraction of PCB's spiked onto a sample of XAD-2 may not completely mimic the process of extracting PCB's from an actual water sample, so we have some more work to do before completely verifying

the usefulness of this technique. However, we expect to test a water sample containing measurable levels of pesticides and PCB's and finalize the technique in the next few months. We are very excited about this technique as it could greatly improve our ability to do trace analyses of groundwater by chromatographic techniques. It also has tremendous potential for improving our ability to measure trace organics in surface waters, such as Lake Michigan.

We have also accomplished our second objective, completing work on a statistically rigorous method for calculating the limit of detection. We believe this procedure provides a significant improvement to the current method of calculating the detection limit based on the standard deviation of replicate measurements at a single concentration. This work was recently published in *Analytical Chemistry*. A copy of the paper is included with this final report; we will be happy to provide additional reprints if needed. This material was also presented at the American Water Resources Association (AWRA) meeting March 6 and 7, 1997 and at the Society of Environmental Toxicology and Chemistry (SETAC) meeting April 2-4, 1997. Also, we made a presentation on the topic to the Integrated Science

Services staff of the Department of Natural Resources. We are currently developing simplifications to the technique so that it can be applied routinely in analytical chemistry laboratories.

Finally, I want to mention that this grant helped support Michael Zorn in his pursuit of a Ph.D. Mike has performed extremely well, receiving straight A grades in his graduate courses. He has also made great progress on his dissertation, and he expects to be finished by the end of the year. The Wisconsin Coordinating Council deserves credit for helping to provide this opportunity to Mike (a Wisconsin native, by the way).

Anal. Chem. 1997, 69, 3069–3075

Weighted Least-Squares Approach To Calculating Limits of Detection and Quantification by Modeling Variability as a Function of Concentration

Michael E. Zorn,† Robert D. Gibbons,‡ and William C. Sonzogni*,†

Water Chemistry Program, University of Wisconsin—Madison, Madison, Wisconsin 53706, and Department of Biostatistics, University of Illinois—Chicago, Chicago, Illinois 60612

The limit of detection and limit of quantification are current critical issues in environmental testing. In most laboratories, limits are currently calculated on the basis of the standard deviation of replicate analyses at a single concentration. However, since the standard deviation depends on concentration, these single-concentration techniques result in limits that are directly dependent on spiking concentration. A more rigorous approach uses a weighted least-squares regression analysis of replicates spiked at a series of concentrations—a calibration design. In this work, the use of weighted tolerance intervals is introduced for estimating detection and quantification limits. In addition, models for estimating the weights used in calculating weighted prediction intervals and weighted tolerance intervals are presented. Using this method, detection and quantification limits were calculated for gas chromatographic analyses of 16 polychlorinated biphenyls. Results show that the approach developed provides improved estimates of analytical limits and that the single-concentration approaches currently in wide use are seriously flawed. Future work should reduce the data needed for the calibration design approach so that more rigorous detection and quantification limits can be routinely applied.

Chemists are concerned with two types of limits when evaluating data quality. The first is a limit of detection, used to decide whether or not an analyte is present; the second is a limit of quantification, used to decide whether or not the concentration of an analyte can be reliably determined. Current calculation techniques derive detection and quantification limits from variability in analyte response at a single, arbitrary spiked concentration (U.S. EPA's MDL[1] and ML[2]). These single-concentration designs are inconsistent when response variability is not constant with concentration, resulting in calculated limits directly dependent on spiking concentration. More rigorous techniques that utilize a series of standards spiked over a range of known concentrations (calibration designs) can account for nonconstant variance as well as calibration error. In this study, we extend current calibration design techniques for calculating detection and quantification limits by applying a model of response variability as a function of concentration to linear least-squares regression.

Unweighted Least-Squares Regression Analysis. Calibration designs require measurement of replicate spikes at a series of analyte concentrations spanning the estimated detection and quantification limits. A least-squares regression analysis is performed of response (Y) on analyte concentration (X), expressed as a linear, first-order model of the form

$$Y = b_0 + b_1 X + \epsilon \tag{1}$$

where b_0 is the intercept, b_1 is the slope, and ϵ represents error in the response measurement or deviation from the fitted regression line. Errors are assumed to be independent and normally distributed with mean zero and constant variance.

Calibration design detection and quantification limit estimators are based on prediction intervals or tolerance intervals. Prediction intervals provide $(1 - \alpha)100\%$ confidence of including the next single instrument response (or measured concentration) at the true concentration (X), whereas tolerance intervals provide $(1 - \alpha)100\%$ confidence of including $(P)100\%$ of the entire population of instrument responses at the true concentration. For example, a tolerance interval with $\alpha = 0.01$ and $P = 0.99$ would provide 99% confidence of including 99% of future instrument responses at X.

Assuming constant variance, the prediction interval around a predicted response ($\hat{Y}_j$) at concentration X_j is defined by

$$Y_j = \hat{Y}_j \pm t_{(1-\alpha/2,n-2)} s \left[1 + \frac{1}{n} + \frac{(X_j - \bar{X})^2}{Sxx}\right]^{1/2} \tag{2}$$

where $t_{(1-\alpha/2,n-2)}$ is the $(1 - \alpha/2)100$ percentage point of Student's t distribution on $n - 2$ degrees of freedom, s is the residual standard deviation, n is the number of measurements, $\bar{X}$ is the mean concentration, and $Sxx = \Sigma(X_i - \bar{X})^2$. Prediction intervals are appropriate for determining detection and quantification limits calculated routinely for a small number of future analyses—for rigorous application, only one future sample is covered. Hubaux and Vos[3] developed detection limit theory using prediction intervals of this form.

Tolerance intervals are wider and will provide larger estimates of detection and quantification limits than corresponding prediction intervals. However, inference to a large and potentially unknown number of future detection decisions is possible with a high degree of confidence, making tolerance intervals attractive for routine application in commercial laboratories. Lieberman and

† University of Wisconsin—Madison.
‡ University of Illinois—Chicago.

(1) U.S. EPA. *Fed. Regist.* 1984, *49* (No. 209), 43430–43431.
(2) U.S. EPA. *Guidance on Evaluation, Resolution, and Documentation of Analytical Problems Associated with Compliance Monitoring*; EPA/821-B-95-001; U.S. Environmental Protection Agency: Washington, DC, 1993.
(3) Hubaux, A.; Vos, G. *Anal. Chem.* 1970, *42*, 849–855.

S0003-2700(97)00082-6 CCC: $14.00 © 1997 American Chemical Society

Miller[4] developed tolerance intervals for linear least-squares regression that simultaneously bracket the expected values ($\hat{Y}$) for all values of X. These "simultaneous" intervals provide an error band on the entire calibration line. Tolerance intervals at a given concentration (nonsimultaneous) may be more appropriate for calculating detection and quantification limits. A nonsimultaneous solution to the Lieberman and Miller equation can be obtained by substituting Student's t for the factor $(2F)^{1/2}$, as described by Miller.[5] Nonsimultaneous tolerance intervals are given by

$$Y_j = \hat{Y}_j \pm s\left\{t_{(1-\alpha/2,n-2)}\left[\frac{1}{n} + \frac{(X_j - \bar{X})^2}{Sxx}\right]^{1/2} + N(P)\left(\frac{n-2}{{}_{\alpha/2}\chi^2_{n-2}}\right)^{1/2}\right\} \quad (3)$$

where $N(P)$ is the two-sided P percentile point of the unit normal distribution, and ${}_{\alpha/2}\chi^2_{n-2}$ is the $(\alpha/2)100$ percentage point of the χ^2 distribution on $n-2$ degrees of freedom. The t and χ^2 test statistics provide the confidence level, while the normal test statistic provides the coverage—both can be set independently. Gibbons[6] proposed using similar (simultaneous) tolerance intervals to calculate detection limits.

The above discussion and equations are valid, provided the error assumptions are not violated. Error analysis can be performed by plotting the residuals (or deviations from the fitted response) versus the fitted response values. A horizontal band of residuals indicates constant variance, and unweighted least-squares regression is appropriate. A funnel shape opening toward larger values signifies increasing variability with concentration, resulting in incorrect estimates of the intercept, slope, and residual standard deviation using unweighted least-squares regression. Nonconstant variance has been previously documented for various chemical analyses and analytes.[6-14] For instance, this phenomenon has long been associated with nuclear analytical measurements as a result of Poisson counting statistics (see Currie[7]). Also, Morrison[9] noted nonconstant variance in analyzing for numerous elements in moon rocks from the Lunar Analysis Program of the U.S. National Aeronautics and Space Administration. Finally, Kurtz et al.[13] found similar behavior in analyzing several pesticides by gas chromatography. In addition, the International Union of Pure and Applied Chemistry (IUPAC) has incorporated the issue of "heteroscedasticity" or nonconstant variance into their recommendations for calculating limits of detection and quantification.[15]

There are two main solutions to the problem of nonconstant variance: transform the data or perform a weighted least-squares regression analysis. The objective of a transformation is to rescale the data so that variability becomes constant, allowing the original unweighted least-squares regression theory to be used. However, a variance-stabilizing transformation is often difficult to determine. Also, if the response values are transformed, linearity of the calibration curve will likely be compromised, and an appropriate transformation for the X axis will be required in order to retain linearity. This can be problematic, and, as several previous authors have pointed out,[12,16,17] a better solution is to use weighted least-squares regression.

Weighted Least-Squares Regression Analysis. Weighted least-squares regression is a modification of ordinary least-squares that accommodates nonconstant variance. More reliable data (smaller variability) are given greater emphasis, or weight. Often, the inverse standard deviation ($1/s_i$) or the inverse variance ($1/s_i^2$) at a given concentration will work well as a weight. A major advantage of this technique is that one does not have to refit response versus concentration, since the original data remain unchanged. However, as was mentioned, it does require modification to the ordinary least-squares theory. The weighted least-squares model is

$$Y = b_{0w} + b_{1w}X + \epsilon \quad (4)$$

The weighted slope is calculated as

$$b_{1w} = \frac{Sxy_w}{Sxx_w} \quad (5)$$

where

$$Sxy_w = \sum w_i(X_i - \bar{X}_w)Y_i = \sum w_iX_iY_i - \frac{(\sum w_iX_i)(\sum w_iY_i)}{\sum w_i} \quad (6)$$

$$Sxx_w = \sum w_i(X_i - \bar{X}_w)^2 = \sum w_iX_i^2 - \frac{(\sum w_iX_i)^2}{\sum w_i} \quad (7)$$

$$\bar{X}_w = \frac{\sum w_iX_i}{\sum w_i} \quad (8)$$

and

$$\bar{Y}_w = \frac{\sum w_iY_i}{\sum w_i} \quad (9)$$

The weighted intercept and the predicted response (fitted regres-

(4) Lieberman, G. J.; Miller, R. G. *Biometrika* 1963, *50*, 155–168.
(5) Miller, R. G. *Simultaneous Statistical Inference*; McGraw-Hill: New York, 1966.
(6) Gibbons, R. D. *Statistical Methods For Groundwater Monitoring*; Wiley: New York, 1994.
(7) Currie, L. A. *Anal. Chem.* 1968, *40*, 586–593.
(8) Püschel, R. *Mikrochim. Acta* 1968, *4*, 783–801.
(9) Morrison, G. H. *Anal. Chem.* 1971, *43*, 22A–31A.
(10) Horwitz, W.; Kamps, L. R.; Boyer, K. W. *J. Assoc. Off. Anal. Chem.* 1980, *63*, 1344–1354.
(11) Horwitz, W. *Anal. Chem.* 1982, *54*, 67A–76A.
(12) Oppenheimer, L.; Capizzi, T. P.; Weppelman, R. M.; Mehta, H. *Anal. Chem.* 1983, *55*, 638–643.
(13) Kurtz, D. A.; Rosenberger, J. L.; Tamayo, G. J. In *Trace Residue Analysis, Chemometric Estimations of Sampling, Amount, and Error*; Kurtz, D. A., Ed.; ACS Symposium Series 284; American Chemical Society: Washington, DC, 1985; Chapter 9.
(14) Rocke, D. M.; Lorenzato, S. *Technometrics* 1995, *37*, 176–184.
(15) Currie, L. A. *Pure Appl. Chem.* 1995, *67*, 1699–1723.
(16) Currie, L. A. In *Trace Residue Analysis, Chemometric Estimations of Sampling, Amount, and Error*; Kurtz, D. A., Ed.; ACS Symposium Series 284; American Chemical Society: Washington, DC, 1985; Chapter 5.
(17) Owens, K. G.; Bauer, C. F.; Grant, C. L. In *Detection in Analytical Chemistry, Importance, Theory, and Practice*; Currie, L. A., Ed.; ACS Symposium Series 361; American Chemical Society: Washington, DC, 1988; Chapter 10.

sion line) are given by

$$b_{0w} = \bar{Y}_w - b_{1w}\bar{X}_w \tag{10}$$

and

$$\hat{Y}_{wj} = b_{0w} + b_{1w}X_j \tag{11}$$

and the weighted residual standard deviation is

$$s_w = \sqrt{\frac{\sum w_i(Y_i - \hat{Y}_{wi})^2}{n-2}} \tag{12}$$

Weighted prediction intervals around a predicted response ($\hat{Y}_{wj}$) at concentration X_j are calculated as

$$Y_j = \hat{Y}_{wj} \pm t_{(1-\alpha/2,n-p-2)}s_w\left[\frac{1}{w_j} + \frac{1}{\sum w_i} + \frac{(X_j - \bar{X}_w)^2}{Sxx_w}\right]^{1/2} \tag{13}$$

The weighted parameters have replaced the unweighted parameters (s, $\bar{X}$, and Sxx), the sum of the weights has replaced n, and $t_{(1-\alpha/2,n-p-2)}$ is the $(1 - \alpha/2)100$ percentage point of Student's t distribution on $n - p - 2$ degrees of freedom (where p is the number of parameters used to model the weights, see below). In addition, the inverse weight ($1/w_j$) at X_j has replaced 1 in the unweighted equation. Since variability is not constant with concentration, the appropriate multiplier for error in future measurements is $s_w/w_j^{1/2}$, rather than simply s as in the unweighted equation.[18]

To provide coverage for a given percentage of future measurements, rather than just the next single measurement, tolerance intervals must be applied. In the weighted case, the tolerance interval becomes

$$Y_j = \hat{Y}_{wj} \pm s_w\left\{t_{(1-\alpha/2,n-2)}\left[\frac{1}{\sum w_i} + \frac{(X_j - \bar{X}_w)^2}{Sxx_w}\right]^{1/2} + \left(\frac{1}{w_j}\right)^{1/2} N(P)\left(\frac{n-p-2}{\chi^2_{\alpha/2,n-p-2}}\right)^{1/2}\right\} \tag{14}$$

Similarly, the weighted parameters replace the unweighted parameters, the sum of the weights replaces n, and $s_w/w_j^{1/2}$ replaces s as the multiplier for error in future measurements, the right-most term.

To this point, all parameters of eqs 13 and 14 can be determined except for the weight w_j at X_j, which must be estimated. This can be achieved by modeling the weights as a function of concentration. In this study, the weights have been set equal to the inverse variance, $1/s_j^2$. Therefore, w_j can be estimated by modeling the standard deviation (or variance) as a function of concentration X. The following models of standard deviation s_x will be evaluated: a quadratic model (where $p = 3$),

$$s_x = a_0 + a_1X + a_2X^2 \tag{15}$$

(18) Caulcutt, R.; Boddy, R. *Statistics for Analytical Chemists*; Chapman and Hall: London, 1983.

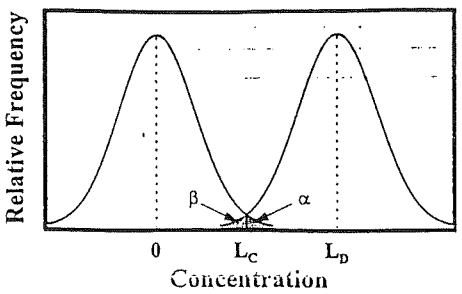

Figure 1. Relationship between a blank, the critical level (L_C), and the limit of detection (L_D). The distributions overlap at L_C with false positive (α) and false negative (β) error rates.

an exponential model (where $p = 2$),

$$s_x = a_0e^{a_1x} \tag{16}$$

and a two-component model (where $p = 2$) proposed by Rocke and Lorenzato[14] and approximated by

$$s_x = \sqrt{a_0 + a_1X^2} \tag{17}$$

Limit of Detection. A brief review of detection limit theory is provided here; for a more thorough review, see refs 3, 6, 7, and 19. Currie[7] defined the critical level (L_C) as "a decision limit at which one may decide whether or not the result of an analysis indicates detection". This level is concerned with the *signal* or *measured concentration* that corresponds to (unreliable) detection, and results from the hypothesis test H_0: $X = 0$. The critical level has a specifically defined false positive (type I) error rate (α)—of 1%, for example—but an undefined false negative (type II) error rate (β). According to Currie, the detection limit (L_D) is the *true concentration* "at which a given analytical procedure may be relied upon to lead to detection". A second hypothesis test (H_0: $X = L_D$) at the detection limit allows the false negative error rate to be set. A distribution at zero and a distribution at the detection limit overlap at the critical level with a given α and β (see Figure 1). Thus, when the true concentration is L_D, an instrument response indistinguishable from a blank (below L_C) is rare (e.g., 1%). It should be noted that setting the detection limit at the critical level incorrectly provides a false negative error rate of 50%. In other words, half of all measurements with concentration L_D (or L_C) will result in an instrument response indistinguishable from a blank.

The critical level and detection limit can be applied to calibration designs using prediction intervals or tolerance intervals as shown in Figure 2. Parameters include the critical level in response units (Y_C), equal to the upper interval at zero concentration; the critical level in concentration units (L_C), expressed as

$$L_C = \frac{Y_C - b_{0w}}{b_{1w}} \tag{18}$$

and the detection limit in concentration units (L_D), set to provide an appropriate false negative error rate.

For calculating detection and quantification limits, one-sided prediction and tolerance intervals are more appropriate than the

(19) Gibbons, R. D. *Environ. Ecol. Stat.* 1995, 2, 125–167.

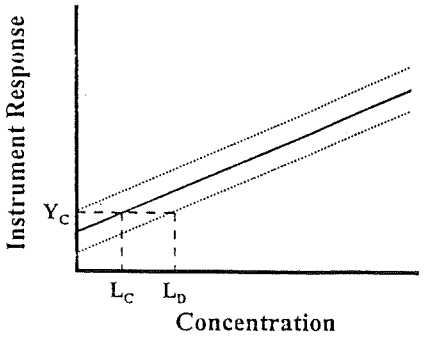

Figure 2. Calibration design critical level in response (Y_C) and concentration (L_C) units and limit of detection (L_D) in concentration units. Includes the calibration line (—) and prediction or tolerance intervals (- - -).

general two-sided intervals provided above (eqs 2, 3, 13, and 14). One-sided intervals are calculated using the $(1 - \alpha)100$ percentage point of Student's t distribution, the one-sided P percentile point of the unit normal distribution, and the $(\alpha)100$ percentage point of the χ^2 distribution.

Weighted prediction intervals (one-sided) were applied to the calculation of a critical level and a detection limit by Oppenheimer et al.[12] The critical level in response units is described by

$$Y_C = b_{0w} + t_{(1-\alpha,n-p-2)}s_w\left[\frac{1}{w_0} + \frac{1}{\sum w_i} + \frac{\bar{X}_w^{\,2}}{Sxx_w}\right]^{1/2} \quad (19)$$

and the detection limit in concentration units by

$$L_D = L_C + \frac{t_{(1-\beta,n-p-2)}s_w}{b_{1w}}\left[\frac{1}{w_{L_D}} + \frac{1}{\sum w_i} + \frac{(L_D - \bar{X}_w)^2}{Sxx_w}\right]^{1/2} \quad (20)$$

Oppenheimer et al. made several conservative assumptions, resulting in larger estimates of L_D, to simplify these equations and the subsequent calculations. They set the weight at zero (w_0) and the weight at the detection limit (w_{L_D}) equal to 1 (the weight at the lowest spiking concentration) to circumvent modeling the weights as a function of concentration; they also performed calculations with $w_0 = w_{L_D} = 91.74$, an estimate of the weight at zero concentration. Also, $(L_D - \bar{X}_w)^2$ was replaced by $\bar{X}_w^2$ to avoid an iterative solution for L_D.

Currie[16] later combined weighted prediction intervals and a linear model of the standard deviation ($s_x = a_0 + a_1X$) to more accurately calculate the limit of detection. Since the lowest standard in this data set exceeded the calculated detection limit by more than an order of magnitude, the linear model was sufficient in describing the overall variability. In this study, data at and below the limit of detection will necessitate the incorporation of more complex models of the weights (quadratic, exponential, and two-component Rocke and Lorenzato[14]) and the iteration of the full equation in order to correctly calculate w_0 and w_{L_D} and to simultaneously solve for the detection limit.

One-sided weighted tolerance intervals can similarly be applied to the calculation of the critical level in response units by

$$Y_C = b_{0w} + s_w\left\{t_{(1-\alpha,n-2)}\left[\frac{1}{\sum w_i} + \frac{\bar{X}_w^{\,2}}{Sxx_w}\right]^{1/2} + \left(\frac{1}{w_0}\right)^{1/2} N(P)\left(\frac{n-p-2}{{}_{\alpha}\chi^2_{n-p-2}}\right)^{1/2}\right\} \quad (21)$$

and the detection limit in concentration units by

$$L_D = L_C + \frac{s_w}{b_{1w}}\left\{t_{(1-\beta,n-2)}\left[\frac{1}{\sum w_i} + \frac{(L_D - \bar{X}_w)^2}{Sxx_w}\right]^{1/2} + \left(\frac{1}{w_{L_D}}\right)^{1/2} N(P)\left(\frac{n-p-2}{{}_{\beta}\chi^2_{n-p-2}}\right)^{1/2}\right\} \quad (22)$$

A weighted tolerance interval analysis will also be performed by calculating w_0 and w_{L_D} and simultaneously solving for the detection limit.

Limit of Quantification. The limit of quantification is used to decide whether or not the concentration of an analyte can be reliably determined. It was suggested by Currie[7] that the "determination limit" (L_Q) be set at a concentration with sufficiently small standard deviation to allow for accurate quantification—a 10% relative standard deviation was suggested. Similarly, the U.S. Environmental Protection Agency has defined the minimum level (ML) as 3.18 times the method detection limit (MDL), or 10 times the standard deviation used to calculate the MDL.[2] This method is widely used to calculate the limit of quantification; however, it is based on variability in analyte response at a single concentration, making it inconsistent in situations of nonconstant variance.

Gibbons et al.[20] have suggested an alternative to the EPA's minimum level: an alternative minimum level, or AML. This approach defines Y_Q (the determination limit in response units) as 10 times the standard deviation at the lowest detectable signal (L_C) plus the weighted intercept, or

$$Y_Q = 10s_{L_C} + b_{0w} \quad (23)$$

where s_{L_C} is the standard deviation at the critical level. Addition of the weighted intercept converts from a response deviation to an actual response value. The corresponding concentration L_Q can be obtained by

$$L_Q = \frac{Y_Q - b_{0w}}{b_{1w}} \quad (24)$$

The AML is the concentration that provides an upper bound for the operationally defined level L_Q. The original procedure uses weighted least-squares to obtain the calibration parameters b_{0w} and b_{1w} and standard deviation modeling and a single sample tolerance interval to simultaneously solve for L_C and s_{L_C}, to subsequently obtain Y_Q. A weighted prediction interval is then used to calculate the AML. In this work, either the calibration-based prediction interval (eq 19) or tolerance interval (eq 21) approach developed earlier will be used to obtain more rigorous estimates of L_C, s_{L_C}, and Y_Q. The weighted prediction interval-

(20) Gibbons, R. D.; Coleman, D. E.; Maddalone, R. F. *Environ. Sci. Technol.* In press.

Table 1. PCB Concentration Ranges

PCB no.[a]	concn (ng/mL) lowest	concn (ng/mL) highest	PCB no.[a]	concn (ng/mL) lowest	concn (ng/mL) highest
1	0.0775	77.450	61	0.0054	5.380
3	0.2047	204.650	65	0.0053	5.270
4	0.0484	48.405	77	0.0089	8.860
14	0.0230	22.960	101	0.0065	6.460
18	0.0128	12.840	128	0.0044	4.395
30	0.0044	4.400	155	0.0051	5.070
31	0.0088	8.800	166	0.0046	4.560
54	0.0117	11.665	180	0.0026	2.580

[a] Numbered according to Ballschmiter and Zell.[21]

based AML can then be calculated as

$$\mathrm{AML} = L_Q + \frac{t_{(1-\beta,n-p-2)}S_w}{b_{1w}}\left[\frac{1}{w_{L_Q}} + \frac{1}{\sum w_i} + \frac{(L_Q - \bar{X}_w)^2}{Sxx_w}\right]^{1/2} \quad (25)$$

or the weighted tolerance interval-based AML as

$$\mathrm{AML} = L_Q + \frac{s_w}{b_{1w}}\left\{t_{(1-\beta,n-2)}\left[\frac{1}{\sum w_i} + \frac{(L_Q - \bar{X}_w)^2}{Sxx_w}\right]^{1/2} + \left(\frac{1}{w_{L_Q}}\right)^{1/2} N(P)\left(\frac{n-p-2}{{}^{\beta}\chi^2_{n-p-2}}\right)^{1/2}\right\} \quad (26)$$

EXPERIMENTAL SECTION

Standards. A stock solution containing 16 polychlorinated biphenyls (PCBs), ranging from congener 1 (monochloro, numbered according to Ballschmiter and Zell[21]) to congener 180 (heptachloro), was prepared in isooctane (EM Science, Gibbstown, NJ). Concentrations were chosen to produce a similar instrument response for each congener. Eight standards were prepared by diluting the stock mixture by factors of 2 (highest concentration), 5, 10, 20, 50, 100, 200, and 2000 (lowest concentration). Table 1 lists the concentration range for each PCB congener. Congener 204 was subsequently added to each standard (at a concentration of 7.275 ng/mL) to be used as a retention time reference. Seven replicates were analyzed at each of the eight concentrations. Samples were run in seven blocks, each block containing one replicate of each standard. The eight standards within each block were randomized to eliminate the effect of any systematic error.

Gas Chromatography. The standards were analyzed using a Hewlett-Packard 5890A gas chromatograph (GC) equipped with a ^{63}Ni electron capture detector (ECD). The detector was maintained at a temperature of 330 °C, with the flow rate of makeup gas (Ar/CH_4) at 25–30 mL/min. The carrier gas (H_2) velocity was approximately 50 cm/s. A 2 μL injection was performed in the splitless mode (0.7 min purge delay), with an injection port temperature of 300 °C. The column was a J&W Scientific (Folsom, CA) DB-5 capillary column (30 m × 0.25 mm, 0.25 μm film thickness). The oven temperature program was as follows: initial temperature 90 °C, 5 °C/min to 110 °C, 2 °C/min to 220 °C, 10 °C/min to 300 °C, retained for 25 min at 300 °C.

(21) Ballschmiter, K.; Zell, M. *Fresenius Z. Anal. Chem.* **1980**, *302*, 20–31.

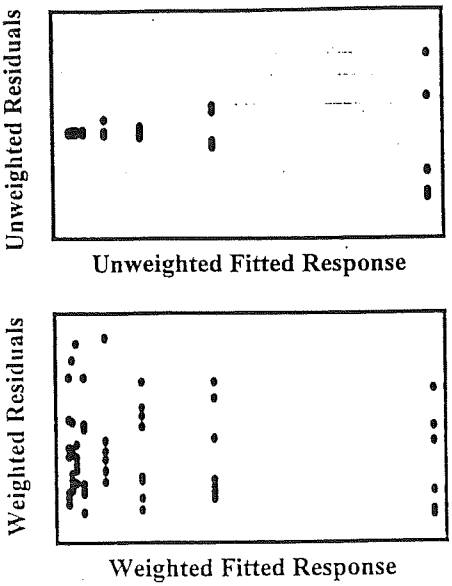

Figure 3. (Top) Unweighted least-squares residual plot for PCB congener 1; the funnel shape indicates nonconstant variance. (Bottom) Weighted least-squares residual plot for PCB congener 1; the horizontal band of residuals indicates constant variance.

Computations. Statistical analyses were performed using Minitab program software (Release 9.2). Various macros were written to facilitate regression analyses and calculation of detection and quantification limits using the methods developed above. Software packages are commercially available that will perform these computations directly.[22,23]

RESULTS AND DISCUSSION

Response variability increases with concentration for all 16 PCB congeners. Unweighted least-squares residual plots are similar to that for congener 1, shown in Figure 3 (top). The funnel shape opening toward higher concentration signifies increasing variance, making unweighted least-squares regression inappropriate. Weighted least-squares regression, where the weights are equal to the inverse observed variance ($w_i = 1/s_i^2$), improves the residual plots significantly (see Figure 3, bottom). The horizontal band of residuals, calculated as $w_i^{1/2}(Y_i - \hat{Y}_{wi})$, indicates constant variance and a valid analysis.

Table 2 lists the unweighted and weighted least-squares regression parameters. The slope is essentially unaffected by the weighting, while the intercept is moderately affected in several cases (PCB congeners 31, 61, 77, and 180). The residual standard deviation, however, is decreased considerably using weighted least-squares to near unity ($s_w \approx 1$) due to the inverse variance weighting scheme.[12] As a result, the multiplier for error in future measurements using weighted least-squares ($s_w/w_j^{1/2}$) becomes simply s_j, the standard deviation at concentration X_j. This allows the width of the weighted prediction and tolerance intervals to change with concentration, accurately reflecting the actual error (see Figure 4). This ultimately results in tighter prediction and tolerance intervals at low concentration; consequently, detection and quantification limits are significantly lower and more accurate using weighted least-squares regression.

(22) Gibbons, R. D. *DETECT: A computer program for computing detection and quantification limits*; Scientific Software International: Chicago, 1996.
(23) Gibbons, R. D. *AML: A computer program for computing the Alternative Minimum Level*; Scientific Software International: Chicago, 1996.

Table 2. Least-Squares Regression Parameters

PCB no.[a]	n	unweighted intercept	unweighted slope	unweighted residual standard deviation	weighted intercept	weighted slope	weighted residual standard deviation
1	56	12.2	18.4	88.1	3.4	19.6	1.0
3	56	−1.3	7.2	101.1	−0.2	7.0	1.0
4	56	25.2	35.0	107.0	2.0	40.5	1.0
14	56	34.7	179.2	255.7	−1.3	196.1	1.0
18	56	15.0	205.5	165.7	5.4	210.8	1.0
30	54	16.7	548.6	159.2	17.8	536.2	1.0
31	56	−14.1	815.2	478.5	14.9	748.3	1.0
54	56	20.6	226.9	166.9	7.6	233.9	1.0
61	56	−4.8	580.2	235.8	10.3	522.2	1.2
65	56	−4.3	425.5	154.0	1.5	403.0	1.0
77	56	−6.6	150.9	97.7	9.0	120.5	1.5
101	56	13.5	360.2	159.0	15.5	330.0	1.4
128	49	−66.5	689.0	263.5	−24.8	566.5	1.2
155	56	18.7	445.8	150.3	4.7	469.8	1.0
166	55	−4.5	597.4	211.9	1.5	586.2	1.0
180	56	−2.2	601.2	118.5	6.1	542.8	1.2

[a] Numbered according to Ballschmiter and Zell.[21]

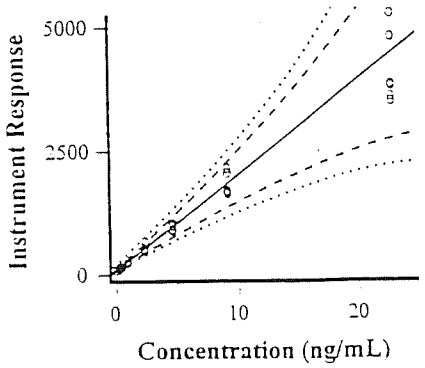

Figure 4. Weighted least-squares regression analysis for PCB congener 14. Includes the calibration line (—), weighted prediction intervals (- - -), and weighted tolerance intervals (···). The width of the weighted intervals changes with concentration, accurately representing the actual data.

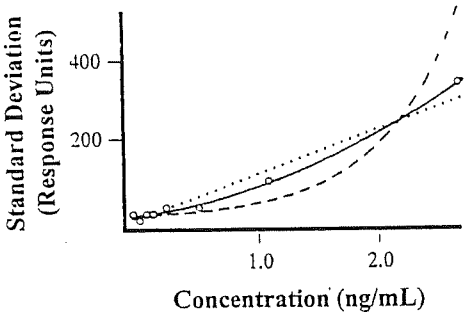

Figure 5. Standard deviation as a function of concentration for PCB congener 180. Fit using a quadratic model (—), an exponential model (- - -), and a two-component Rocke and Lorenzato model (···).

Table 3. Detection Limits and Alternative Minimum Levels (ng/mL) Using Weighted Prediction and Tolerance Intervals and a Quadratic Model for the Weights

PCB no.[a]	weighted prediction interval[b] L_D	weighted prediction interval[b] AML	weighted tolerance interval[c] L_D	weighted tolerance interval[c] AML
1	0.832	2.326	1.339	2.757
3	1.810	5.138	2.589	5.867
4	0.384	1.026	0.590	1.202
14	0.200	0.567	0.289	0.649
18	0.067	0.190	0.107	0.225
30	0.139	0.388	0.230	0.470
31	0.061	0.169	0.096	0.200
54	0.159	0.439	0.255	0.522
61	0.142	0.346	0.224	0.409
65	0.047	0.124	0.078	0.151
77	0.413	0.800	0.762	0.989
101	0.107	0.211	0.193	0.263
128	0.304	0.681	0.844	0.898
155	0.037	0.102	0.062	0.123
166	0.066	0.177	0.096	0.203
180	0.077	0.181	0.127	0.220

[a] Numbered according to Ballschmiter and Zell.[21] [b] 99% confidence (i.e., $\alpha = \beta = 0.01$). [c] 99% confidence and 99% coverage (i.e., $\alpha = \beta = 0.01$ and $P = 0.99$).

Detection and Quantification Limits. To calculate detection and quantification limits using weighted prediction and tolerance intervals, the weight at zero (w_0), the weight at the limit of detection (w_{L_D}), and the weight at the limit of determination (w_{L_Q}) must be estimated. A quadratic model, an exponential model, and a two-component Rocke and Lorenzato[14] model of the standard deviation as a function of concentration were examined (see Figure 5). In all cases, the quadratic model provides the best overall fit to the data. In general, the exponential model provides an overestimate at low and high concentrations and an underestimate at intermediate concentration, while the opposite is true for the two-component model.

Table 3 lists the detection limit and AML calculated for each of the 16 PCB congeners. Values were calculated using weighted prediction intervals ($\alpha = \beta = 0.01$) and weighted tolerance intervals ($\alpha = \beta = 0.01$, $P = 0.99$) with a quadratic model for the weights. As expected, tolerance interval-based estimates are larger than corresponding prediction interval-based estimates due to the increased coverage for a percentage of all future measurements, rather than just the next single measurement.

For several PCB congeners, the exponential and two-component Rocke and Lorenzato models provide fits comparable to that of the quadratic model. For example, using the exponential model for PCB congener 77 results in a prediction interval-based detection limit of 0.444 ng/mL, a prediction interval-based AML of 0.859 ng/mL, a tolerance interval-based detection limit of 0.667 ng/mL, and a tolerance interval-based AML of 0.998 ng/mL (compared to 0.413, 0.800, 0.762, and 0.989 ng/mL, respectively, using the quadratic model). Similarly, using the two-component model for PCB congener 1 results in a prediction interval-based detection limit of 0.903 ng/mL, a prediction interval-based AML of 2.490 ng/mL, a tolerance interval-based detection limit of 1.480 ng/mL, and a tolerance interval-based AML of 2.911 ng/mL (compared to 0.832, 2.326, 1.339, and 2.757 ng/mL, respectively, using the quadratic model). For other PCB congeners, the exponential and two-component models do not compare as well with the quadratic model.

The authors would like to point out, however, that the quadratic model does not provide the best fit in all situations. For example, using the quadratic model can lead to a negative estimate of the standard deviation at zero concentration, ultimately resulting in a negative estimate of the detection and quantification limit. This result is typically not a problem using exponential or two-component Rocke and Lorenzato fits. In our experience, none of the models examined works best in all situations. It is important

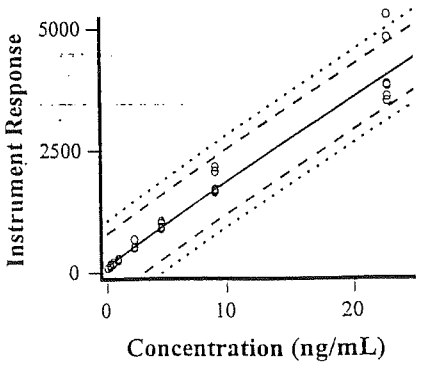

Figure 6. Unweighted least-squares regression analysis for PCB congener 14. Includes the calibration line (—), unweighted prediction intervals (- - -), and unweighted tolerance intervals (···). The unweighted intervals provide a poor fit to the actual data, especially at low concentration.

to be flexible and to have the ability to apply several models to ensure calculation of accurate detection and quantification limit estimates.

For comparison, the simplifications proposed by Oppenheimer et al.[12] were evaluated with respect to calculation of the prediction interval-based detection limit. The factor $(X_j - \bar{X}_w)^2$ was replaced by $\bar{X}_w^2$, and w_j was replaced by the weight at the lowest standard concentration. For four of the 16 PCB congeners (54, 61, 65, and 155), this method provides a detection limit that is larger than the detection limit calculated using the full equation with a quadratic model for the weights—a conservative estimate. However, for the 12 remaining PCB congeners, this method provides a detection limit that is smaller (generally by about 50%) than the detection limit calculated with the full equation—a nonconservative estimate. The most extreme example is PCB congener 3. The detection limit calculated using the simplified method is 0.184 ng/mL, which is about an order of magnitude smaller than 1.810 ng/mL, the detection limit calculated using the full equation with a quadratic model for the weights.

Depending on variability as a function of concentration, as well as the location of w_j with respect to the detection limit, different choices of w_j give detection limits that are closer to the estimates obtained using the full equation. However, the same relative choice of w_j does not work best in all cases. In addition, retaining the factor $(X_j - \bar{X}_w)^2$ in the equations changes the calculated detection limits by only a few percent. It is evident that using a single value of w_j to avoid an iterative solution can result in inaccurate detection limits that are not necessarily conservative estimates.

As a final comparison, unweighted prediction and tolerance interval-based detection limits were calculated. The unweighted intervals are wider than the weighted intervals at very low concentration and provide a poor fit to the actual data (see Figure 6). As a result, applying unweighted least-squares regression in situations of nonconstant variance generally yields overestimated detection limits.[6,12,17] Table 4 lists the unweighted detection limit for all 16 PCB congeners. Unweighted values range from 3 times the weighted value (tolerance interval-based detection limit for PCB congener 128) to nearly 60 times the weighted value (prediction interval-based detection limit for PCB congener 18). This illustrates the importance of performing routine diagnostic tests (e.g., residual plots) and using weighted least-squares regression in situations of nonconstant variance.

Table 4. Unweighted Detection Limits

PCB no.[a]	L_D (ng/mL) prediction interval[b]	L_D (ng/mL) tolerance interval[c]	PCB no.[a]	L_D (ng/mL) prediction inverval[b]	L_D (ng/mL) tolerance interval[c]
1	23.242 (28)[d]	32.285 (24)	61	1.972 (14)	2.761 (12)
3	68.042 (38)	94.858 (37)	65	1.756 (38)	2.448 (31)
4	14.800 (38)	20.570 (35)	77	3.142 (8)	4.393 (6)
14	6.920 (35)	9.614 (33)	101	2.141 (20)	2.985 (15)
18	3.909 (58)	5.433 (51)	128	1.869 (6)	2.690 (3)
30	1.410 (10)	1.972 (9)	155	1.635 (44)	2.277 (37)
31	2.847 (46)	3.964 (41)	166	1.723 (26)	2.422 (25)
54	3.567 (22)	4.958 (19)	180	0.957 (13)	1.341 (11)

[a] Numbered according to Ballschmiter and Zell.[21] [b] 99% confidence (i.e., $\alpha = \beta = 0.01$). [c] 99% confidence and 99% coverage (i.e., $\alpha = \beta = 0.01$ and $P = 0.99$). [d] Numbers in parentheses express unweighted detection limit divided by weighted detection limit.

CONCLUSIONS

Weighted least-squares regression can be used to calculate limits of detection and quantification by applying a model of response variability as a function of concentration to calculate the weight (w_i) at specific concentrations. Methods that do not incorporate a model of the weights are dependent on the choice of w_i, much like single-concentration designs are dependent on the spiking concentration. In this study, a weighted prediction interval approach has been used to calculate detection and quantification limits that can be applied to a limited number of future analyses. In addition, a weighted tolerance interval equation has been defined and likewise used to calculate detection and quantification limits. These values are larger than weighted prediction interval-based estimates but can be applied to a large and potentially unknown number of future analyses. The use of unweighted least-squares regression in situations of nonconstant variance results in significantly overestimated detection and quantification limits and is strongly discouraged.

ACKNOWLEDGMENT

The authors thank Dr. Robert J. Noll of the University of Wisconsin—Madison Water Chemistry Program for insightful discussions and critical review of the manuscript. We would also like to thank a manuscript reviewer for providing helpful suggestions and additional pertinent references.

SUPPORTING INFORMATION AVAILABLE

Listing of gas chromatographic peak areas at each spiking concentration for the 16 PCB congeners (8 pages). Ordering information is given on any current masthead page.

Received for review January 23, 1997. Accepted May 8, 1997.[a]

AC9700821

[a] Abstract published in *Advance ACS Abstracts*, July 1, 1997.

89072244973

b89072244973a