Evaluation of Approximate Methods for Calculating the Limit of Detection and Limit of Quantification

MICHAEL E. ZORN,* ROBERT D. SIBBONS,† 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

In a previous paper, a computational method was presented for determining statistically rigorous limits of detection and quantification. The main purpose of this study is to evaluate similar but less computationally complex methods. These “approximate” methods use data at multiple spiking concentrations, are iterative, can be derived from either prediction intervals or statistical tolerance intervals, and require at a minimum ordinary least-squares regression for calculating the intercept and slope. Approximate detection and quantification limits calculated for various PCB congeners were similar to those calculated using the computationally exact method. Although the exact methods should be employed whenever possible (they include uncertainty in the calibration function and provide greater weight to less variable data), approximate methods can provide detection and quantification limits that are sufficiently accurate for most applications. Practical application of multiconcentration-based methods may involve the use of routinely generated quality control data in the statistical calculations.

Introduction

Limits of detection and quantification are used routinely by analytical chemists to evaluate data quality. The limit of detection is used to decide whether an analyte is present, while the limit of quantification is used to decide whether the concentration of an analyte can be reliably determined. Both limits are most commonly calculated based on variability in analyte response at a single, and often arbitrary, spiked concentration [e.g., U.S. EPA’s Method Detection Limit (MDL) (1) and Method Quantitation Limit (ML) (2)]. Although these “single concentration designs” are computationally simple, they can yield quite variable results depending on the choice of spiking concentration (3, 4). More rigorous methods use replicates spiked at a series of concentrations. These “calibration designs” provide more accurate estimates of the limits of detection and quantification because they account for variability in analyte response at multiple concentrations. However, these methods are computationally and experimentally more complex than single concentration designs.

In a previous paper (5), calibration-based detection and quantification limits were calculated for 16 polychlorinated biphenyl (PCB) congeners using weighted least-squares prediction intervals and statistical tolerance intervals (subsequently referred to simply as tolerance intervals); weights at the limits were estimated by modeling response variability as a function of concentration. Although this full model provides very accurate estimates of the detection and quantification limits, the statistical calculations are quite complex. The main purpose of this study is to evaluate less computationally complex (approximate) methods for calculating detection and quantification limits that reduce the statistical computations required. These approximate methods are also calibration-based; however, they do not incorporate uncertainty in the calibration curve (typically a small component of the overall variability). Limits calculated using the approximate methods are compared to results obtained using the full model provided in ref 5. Practical application of calibration-based methods is also briefly discussed.

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

\[ Y = b_0 + b_1X + \epsilon \]  

where \( b_0 \) is the intercept, \( b_1 \) is the slope, and \( \epsilon \) 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.

Often, the errors do not exhibit constant variance; nonconstant variance has been previously documented for various chemical analyses and analytes (5–17). Several authors (5, 11, 13, 14) have suggested using weighted least-squares (WLS) regression for calculating limits of detection and quantification in situations of nonconstant variance. WLS regression is a modification of OLS that gives greater emphasis (or weight) to lower variability, more reliable data. The WLS model is

\[ Y = b_{0w} + b_{1w}X + \epsilon \]  

where \( b_{0w} \) is the weighted intercept and \( b_{1w} \) is the weighted slope [see ref 5 or Draper and Smith (18) for a more detailed discussion of OLS and WLS regression].

Calibration design detection and quantification limit estimators are based on one-sided statistical intervals. Prediction intervals provide \((1 - \alpha)100\%\) confidence of including the next single instrument response at the true concentration (\(X_t\)), 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 \(P = 0.95\) would provide 99% confidence of including 95% of future instrument responses at \(X_t\). As such, tolerance intervals are better suited to the typical routine production laboratory in which detection limits are estimated once and applied to a large and potentially unknown number of future detection decisions. Weighted prediction and tolerance interval equations are provided in ref 5.

Approximate Prediction and Tolerance Intervals. One-sided prediction intervals around a predicted response (\(Y_p\))
at concentration $X_i$ can be approximated as

$$Y_i = \hat{Y}_i + \hat{b}_1 - \hat{a}_1 \frac{1}{n} \sqrt{1 + \frac{1}{n}}$$

(3)

where $\hat{b}_1$ is the upper $(1 - \alpha)100$ percentage point of Student's t-distribution on $n - p - 2$ df (where $p$ is the number of parameters used to model the standard deviation, see below) and $\hat{a}_1$ is the standard deviation at $X_i$. The approximation ignores uncertainty in the calibration function that relates instrument response (or measured concentration) to true concentration—this component of variability is typically small in practice.

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. A one-sided tolerance limit for a predicted response ($\hat{Y}_i$) at concentration $X_i$ can be approximated by

$$Y_i = \hat{Y}_i + K_{1-p,1-\alpha, n} s_i$$

(4)

where $K_{1-p,1-\alpha, n}$ is a single sample tolerance limit factor for confidence ($\alpha$) and coverage ($P$) based on $n$ available measurements—see Gibbons (15), Guttmann (19), or Hahn and Meeker (20) for tabulated values, or see Link (21) for an alternative method that does not require the use of tables. The method developed by Link is not recommended for calculating $K_{1-p,1-\alpha, n}$ with small values of $n$. For example, using $K_{0.90,0.99, 40}$ with fewer than 10 measurements results in an error of greater than 10%; however, at $n = 40$, the error is decreased to about 1%.

In situations of nonconstant variance, estimation of the standard deviation $s_i$ at $X_i$ is required to calculate prediction and tolerance limits using eqs 3 and 4. This can be achieved by modeling the standard deviation as a function of concentration. The following models of standard deviation $s_i$ have been previously proposed: a proportional model (13, 15), $s_i = a_i X_i$; a linear model (13, 15), $s_i = a_0 + a_1 X_i$; a quadratic model (5), $s_i = a_0 + a_1 X_i + a_2 X_i^2$; an exponential model (3, 5), $s_i = a_0 e^{a_1 X_i}$; and a two-component model (16), approximated by $s_i = (a_0 + a_1 X_i^2)^{1/2}$.

Limit of Detection. As described by Currie (6), the critical level (LC) provides a specifically defined false-positive (type I) error rate and is concerned with the signal or measured concentration that is significantly greater than background instrumental noise. However, the detection limit (LD) provides specific false-positive and false-negative (type II) error rates and represents the true concentration that is significantly greater than zero—see refs 6, 15, or 22 for a more detailed discussion. Previous authors have calculated the critical and limit of detection using WLS prediction (3, 5, 11, 13, 15, 22) and WLS tolerance intervals (5, 15, 22).

The prediction interval-based critical level in response units and detection limit in concentration units can be approximated as

$$Y_C = b_{\alpha^*} + \hat{b}_1 \sqrt{1 + \frac{1}{n}}$$

(5)

and

$$L_D = L_C + \frac{\hat{b}_1}{b_{\alpha^*}} \sqrt{1 + \frac{1}{n}}$$

(6)

respectively, where the intercept ($b_{\alpha^*}$) and slope ($b_{\alpha^*}$) are calculated using either OLS or WLS regression. Values of $s_{\alpha^*}$ (the standard deviation at zero and at the detection limit, respectively) can be estimated by modeling the standard deviation as a function of concentration.

Similarly, the tolerance interval-based critical level in response units and detection limit in concentration units can be approximated by

$$Y_C = b_{\alpha^*} + K_{1-p,1-\alpha, n} s_{\alpha^*}$$

(7)

and

$$L_D = L_C + \frac{K_{1-p,1-\alpha, n} s_{\alpha^*}}{b_{\alpha^*}}$$

(8)

respectively. The critical level in concentration units is calculated as $L_C = (Y_C - b_{\alpha^*})/b_{\alpha^*}$ using either prediction or tolerance intervals. Equations 7 and 8 are statistically equivalent to those used by the American Society for Testing and Materials to calculate the interlaboratory detection estimate, IDE (23).

Limit of Quantification. The limit of quantification is used to decide whether the concentration of an analyte can be reliably determined. Gibbons et al. (3) have suggested an alternative minimum level, or AML, procedure for calculating the limit of quantification. This approach defines $Y_Q$ (the determination limit in response units) as 10 times the standard deviation at the lowest detectable signal (LC) plus the weighted intercept (to accommodate bias). The AML is the concentration that provides an upper bound for the operationally defined level $L_Q$. In ref 5, WLS prediction and tolerance intervals were used to calculate $L_C$, $s_{\alpha^*}$, $Y_Q$, and AML.

In this study, $Y_Q$ is calculated as

$$Y_Q = 10 s_{\alpha^*} + b_{\alpha^*}$$

(9)

where $s_{\alpha^*}$ is the standard deviation at the critical level. The determination limit in concentration units is calculated as $L_Q = (Y_Q - b_{\alpha^*})/b_{\alpha^*}$. The prediction interval-based AML can be approximated as

$$AML = L_Q + \frac{\hat{b}_1}{b_{\alpha^*}} \sqrt{1 + \frac{1}{n}}$$

(10)

using eq 5 to obtain estimates for $L_C$, $s_{\alpha^*}$ and $Y_Q$ and the tolerance interval-based AML can be approximated by

$$AML = L_Q + \frac{K_{1-p,1-\alpha, n} s_{\alpha^*}}{b_{\alpha^*}}$$

(11)

using eq 7 to obtain estimates for $L_C$, $s_{\alpha^*}$, and $Y_Q$. [Note that $L_Q$ is based on variability at $L_C$ (see eq 9), whereas the AML is based on variability at $L_Q$ (see eqs 10 and 11).]

Experimental Section

Gas chromatographic peak areas of 16 PCB congeners [numbered according to Ballschmiter and Zell (24)] were obtained from ref 5—see the original reference for information on PCB standards and gas chromatographic parameters. Various macro's were written to facilitate regression analyses and calculation of detection and quantification limits using the methods developed above. Due to significant lack of fit ($p < 0.05$) of the linear model for five of the PCB congeners (61, 77, 101, 128, and 180) reported in ref 5, these congeners are not included in this study. In practice, analysis of these congeners would typically
TABLE 1. Ordinary and Weighted Least-Squares Parameters Determined from Regression Analyses of Response (Y) as a Function of Concentration (X)

<table>
<thead>
<tr>
<th>PCB</th>
<th>n</th>
<th>OLS intercept</th>
<th>OLS slope</th>
<th>WLS intercept</th>
<th>WLS slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>56</td>
<td>12.2</td>
<td>3.4</td>
<td>18.4</td>
<td>19.6</td>
</tr>
<tr>
<td>3</td>
<td>56</td>
<td>-1.3</td>
<td>-0.2</td>
<td>7.2</td>
<td>7.0</td>
</tr>
<tr>
<td>4</td>
<td>56</td>
<td>25.2</td>
<td>2.0</td>
<td>35.0</td>
<td>40.5</td>
</tr>
<tr>
<td>14</td>
<td>56</td>
<td>34.7</td>
<td>-1.3</td>
<td>179.2</td>
<td>196.1</td>
</tr>
<tr>
<td>18</td>
<td>56</td>
<td>15.0</td>
<td>5.4</td>
<td>205.5</td>
<td>210.8</td>
</tr>
<tr>
<td>30</td>
<td>54</td>
<td>16.7</td>
<td>17.8</td>
<td>548.6</td>
<td>536.2</td>
</tr>
<tr>
<td>31</td>
<td>56</td>
<td>-14.1</td>
<td>14.9</td>
<td>815.2</td>
<td>748.3</td>
</tr>
<tr>
<td>54</td>
<td>56</td>
<td>20.6</td>
<td>7.6</td>
<td>226.9</td>
<td>233.9</td>
</tr>
<tr>
<td>65</td>
<td>56</td>
<td>-4.3</td>
<td>1.5</td>
<td>425.5</td>
<td>403.0</td>
</tr>
<tr>
<td>155</td>
<td>56</td>
<td>18.7</td>
<td>4.7</td>
<td>445.8</td>
<td>469.8</td>
</tr>
<tr>
<td>166</td>
<td>55</td>
<td>-4.5</td>
<td>1.5</td>
<td>597.4</td>
<td>586.2</td>
</tr>
</tbody>
</table>

* Numbers according to Ballschmiter and Zell (24).

Results and Discussion

OLS and WLS ($w_i = 1/s^2$) regression analyses were performed of response (Y) as a function of analyte concentration (X), and the intercepts and slopes are listed in Table 1. Because the data exhibit increasing variability with concentration (5), the intercept and slope are more accurately estimated using WLS; nevertheless, there is reasonable similarity in estimates of the intercept and slope for OLS and WLS regression models.

Table 2 lists the detection limit and alternative minimum level (AML) calculated using the rigorous methods presented in ref 5. Values were calculated using WLS regression analyses of response (Y) as a function of analyte concentration (X) with weighted prediction intervals, WPIs ($\alpha = \beta = 0.01$), and weighted tolerance intervals, WTLs ($\alpha = \beta = 0.01, P = 0.99$). A quadratic model of the standard deviation as a function of concentration was used to estimate the weights. In ref 5, a quadratic model (where $p = 3$) was shown to provide a better fit to these data than an exponential model or a two-component model (16). As shown in Figure 1, the width of the weighted prediction and tolerance intervals change with concentration, accurately representing the actual error.

**Approximate Detection and Quantification Limits.** Table 3 lists detection limits and AMLs calculated using WLS regression analyses (parameters: $b_0w$ and $b_1w$) of response (Y) as a function of analyte concentration (X) with approximate prediction intervals (APIw: $\alpha = \beta = 0.01$) and approximate tolerance intervals (ATIw: $\alpha = \beta = 0.01, P = 0.99$) and a

![Figure 1. Weighted least-squares (WLS) regression analysis of response (Y) as a function of concentration (X) for PCB congener 14. Includes the calibration line (---), weighted prediction intervals (- - -), and weighted tolerance intervals (---).](image)

TABLE 2. Detection Limits and Alternative Minimum Levels (ng/mL) Calculated Using WLS Regression Analyses of Response (Y) as a Function of Concentration (X) with Weighted Prediction Intervals (WPI) and Weighted Tolerance Intervals (WTI)

<table>
<thead>
<tr>
<th>PCB</th>
<th>WPI</th>
<th>WTI</th>
<th>WPI</th>
<th>WTI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.832</td>
<td>1.339</td>
<td>2.326</td>
<td>2.757</td>
</tr>
<tr>
<td>3</td>
<td>1.810</td>
<td>2.589</td>
<td>5.138</td>
<td>5.867</td>
</tr>
<tr>
<td>4</td>
<td>0.384</td>
<td>0.590</td>
<td>1.026</td>
<td>1.202</td>
</tr>
<tr>
<td>14</td>
<td>0.200</td>
<td>0.289</td>
<td>0.567</td>
<td>0.649</td>
</tr>
<tr>
<td>18</td>
<td>0.067</td>
<td>0.107</td>
<td>0.190</td>
<td>0.225</td>
</tr>
<tr>
<td>30</td>
<td>0.139</td>
<td>0.230</td>
<td>0.388</td>
<td>0.470</td>
</tr>
<tr>
<td>31</td>
<td>0.061</td>
<td>0.096</td>
<td>0.169</td>
<td>0.200</td>
</tr>
<tr>
<td>54</td>
<td>0.159</td>
<td>0.255</td>
<td>0.439</td>
<td>0.522</td>
</tr>
<tr>
<td>65</td>
<td>0.047</td>
<td>0.078</td>
<td>0.124</td>
<td>0.151</td>
</tr>
<tr>
<td>155</td>
<td>0.037</td>
<td>0.062</td>
<td>0.102</td>
<td>0.123</td>
</tr>
<tr>
<td>166</td>
<td>0.066</td>
<td>0.096</td>
<td>0.177</td>
<td>0.203</td>
</tr>
</tbody>
</table>

* Numbers according to Ballschmiter and Zell (24). Numbers according to Ballschmiter and Zell (24). * Numbers according to Ballschmiter and Zell (24). * Numbers according to Ballschmiter and Zell (24). * Numbers according to Ballschmiter and Zell (24). * Numbers according to Ballschmiter and Zell (24). * Numbers according to Ballschmiter and Zell (24). * Numbers according to Ballschmiter and Zell (24).

be restricted to those concentrations for which the linear model could be applied.

Table 3 lists detection limits and AMLs calculated using WLS regression analyses of response (Y) as a function of analyte concentration (X) with approximate prediction intervals (APIw: $\alpha = \beta = 0.01$) and approximate tolerance intervals (ATIw: $\alpha = \beta = 0.01, P = 0.99$). A quadratic model for the standard deviation was used. Values of K were estimated as in Link (22), where $K_{0.01;0.99;54} = 3.113, K_{0.99;0.99;55} = 3.104$, and $K_{0.99;0.99;56} = 3.095$. Approximate limits are very similar to limits calculated using WLS prediction and tolerance intervals. All values in Table 3 are within 20% of the corresponding values in Table 2 except the tolerance interval-based detection limit for congener 65 (22%). As shown in Figure 2, approximate tolerance intervals (with WLS estimates of the regression parameters) are indistinguishable from WLS tolerance intervals at very low concentration.

The required calculations would be further simplified if ordinary estimates of the intercept and slope could be substituted for the weighted estimates, thereby eliminating the need to perform the more difficult WLS regression. As discussed above, the intercept and slope are not greatly affected by weighting. Calculations were repeated using OLS regression analyses (parameters: $b_0$ and $b_1$) of response (Y) as a function of analyte concentration (X) with approximate prediction intervals (APIo: $\alpha = \beta = 0.01$) and approximate tolerance intervals (ATIo: $\alpha = \beta = 0.01, P = 0.99$) and a
Tolerance Intervals (ATI) and Approximate Prediction Intervals (API) are used in the determination of detection limits. The use of an analytical method that does not require frequent recalculations of detection limits is preferred. Regulatory requirements are increasingly calling for frequent recalculations of detection limits for quality control purposes. Calculation of detection and quantification limits using a multiconcentration approach could add substantial time and expense to the analytical process relative to the use of single concentration methods. For example, many laboratories routinely analyze thousands of analytes using hundreds of analytical procedures. Also, regulatory requirements are increasingly calling for frequent recalculations of detection limits for quality control purposes. Calculation of detection and quantification limits using a multiconcentration approach could add substantial time and expense to the analytical process relative to the use of single concentration methods now widely in use.

TABLE 4. Detection Limits and Alternative Minimum Levels (ng/mL) Calculated Using OLS Regression Analyses of Response (Y) as a Function of Concentration (X) with Approximate Prediction Intervals (API) and Approximate Tolerance Intervals (ATI)

<table>
<thead>
<tr>
<th>PCB*</th>
<th>API&lt;sup&gt;α&lt;/sup&gt;</th>
<th>ATI&lt;sup&gt;α&lt;/sup&gt;</th>
<th>API&lt;sup&gt;β&lt;/sup&gt;</th>
<th>ATI&lt;sup&gt;β&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.915</td>
<td>1.262</td>
<td>2.547</td>
<td>2.862</td>
</tr>
<tr>
<td>3</td>
<td>1.810</td>
<td>2.479</td>
<td>5.023</td>
<td>5.620</td>
</tr>
<tr>
<td>4</td>
<td>0.435</td>
<td>0.608</td>
<td>1.216</td>
<td>1.376</td>
</tr>
<tr>
<td>14</td>
<td>0.231</td>
<td>0.320</td>
<td>0.643</td>
<td>0.725</td>
</tr>
<tr>
<td>18</td>
<td>0.071</td>
<td>0.099</td>
<td>0.198</td>
<td>0.224</td>
</tr>
<tr>
<td>30</td>
<td>0.135</td>
<td>0.192</td>
<td>0.376</td>
<td>0.429</td>
</tr>
<tr>
<td>31</td>
<td>0.054</td>
<td>0.074</td>
<td>0.150</td>
<td>0.168</td>
</tr>
<tr>
<td>54</td>
<td>0.165</td>
<td>0.228</td>
<td>0.459</td>
<td>0.516</td>
</tr>
<tr>
<td>65</td>
<td>0.041</td>
<td>0.057</td>
<td>0.115</td>
<td>0.128</td>
</tr>
<tr>
<td>155</td>
<td>0.039</td>
<td>0.055</td>
<td>0.109</td>
<td>0.124</td>
</tr>
<tr>
<td>166</td>
<td>0.062</td>
<td>0.084</td>
<td>0.171</td>
<td>0.190</td>
</tr>
</tbody>
</table>

<sup>α</sup> Numbered according to Ballschmiter and Zell (24).<sup>β</sup> 99% confidence (i.e., α = β = 0.01).<sup>γ</sup> 99% confidence and 99% coverage (i.e., α = β = 0.01 and P = 0.99).

A potential option for reducing the labor involved in performing multiconcentration detection and quantification limit determinations is to utilize routinely generated quality control data in the statistical calculations. For example, results from laboratory-fortified matrixes (LFMs) (or matrix spikes, as defined in Standard Methods for the Examination of Water and Wastewater (25)) represent a large pool of data from which to apply a multiconcentration approach. It is recommended that, as a minimum, one LFM be included with each sample set (batch) or on a 5% basis, whichever is more frequent (25). By varying the fortification (i.e., spiking) level, a laboratory could accumulate matrix-specific data for multiconcentration calculations with little extra effort. Also, for laboratories not currently performing routine LFMs, it might be possible to use instrument calibration data in detection limit determinations. Perhaps a relationship could be established between calibration samples and real world (i.e., matrix specific) samples. An added advantage of using extant, routinely generated data (as opposed to using data generated on a single day) is that day-to-day variability (inevitable in routine measurement) is incorporated.

Even with reducing the effort to determine multiconcentration detection and quantification limits, single concentration designs are still required by many regulations. Consequently, single concentration techniques will likely continue in use for some time. Given this situation and the dependence of the calculated limits on the spiking concentration, the authors recommend that, when the single concentration technique must be used, the chosen spiking concentration always be reported along with the calculated limits. Without knowledge of the spiking concentration when using the single concentration designs, detection and quantification limits produced by different methods or different laboratories should not be compared.

Supporting Information Available
A computational example for PCB congener 65 is available as Supporting Information (4 pages with 2 tables). This
Material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited
(19) Guttman, I. Statistical Tolerance Regions; Classical and Bayesian; Hafner: Darien, CT, 1970.

Received for review November 4, 1998. Revised manuscript received March 10, 1999. Accepted April 16, 1999.