Journal of Chromatography B, 872 (2008) 162–166

Contents lists available at ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

Weighted least squares in calibration: The problem with using "quality coefficients" to select weighting formulas

Joel Tellinghuisen

Department of Chemistry, Vanderbilt University, Box 1668 B, Nashville, TN 37235, United States

ARTICLE INFO

Article history:
Received 5 April 2008
Accepted 29 July 2008
Available online 6 August 2008

Keywords:
Calibration
Weighted least squares
Quality coefficient
Monte Carlo
Heteroscedasticity

ABSTRACT

The quality coefficient (Q) has frequently been used to select weighting formulae in calibration, and especially so in bioanalytical work, where there has been increasing awareness of the importance of data heteroscedasticity in recent years. However, this quantity is statistically flawed and should not be used for this purpose. The quality coefficient is computed from the differences between the apparent and true concentrations of the calibration samples as obtained from the least-squares calibration fit. Q is defined as the sum of either the squares or the absolute values of these differences, taken directly or as percentage (relative) deviations. It is calculated for several different trial weighting formulae, and the lowest Q value is then deemed to identify the best weighting choice. However, these Qs are predisposed to favor data consistent with their definitions—homoscedastic data for tests employing absolute differences, and data having proportional error (constant coefficient of variance) for tests using relative differences—because the Q in each case closely resembles the quantity actually minimized by the least-squares fit of the calibration data. The problem is examined and illustrated through Monte Carlo computations on data having either constant or proportional uncertainty and subjected to both tests. A modified Q based on the results of both the absolute and relative tests is much more reliable than either test alone but is still not recommended as a solution to the weighting problem, as other, statistically sound approaches are available and readily used.

1. Introduction

In classical univariate calibration, n calibration points (x_i, y_i) define a calibration curve ($y = f(x)$), and the amount of the unknown (x_0) is determined by solving the equation $y_0 = f(x_0)$, where y_0 is the response for the unknown. In its simplest and most widely used implementation, the response function is assumed to be linear ($y = a + bx$), the values of the independent variable x for the calibration data are taken to be error-free, and the response variable y is assumed to possess normally distributed random error of constant standard deviation σ. Unweighted linear regression (ordinary least squares, or OLS) is used to obtain estimates of the calibration parameters a and b, from which $x_0 = (y_0 - a)/b$ [1,2].

The limitations of this calibration procedure in bioanalytical work have increasingly been recognized in recent years, especially the need for weighted regression (weighted least squares, or WLS) to accommodate the often strong heteroscedasticity in the data for chromatographic techniques [3–11]. The interest in better calibration is driven in part by government guidelines and regulations, existing or proposed [10,12–15]. For the usual straight-line calibration, heteroscedastic data lead to altered equations for a and b, through the incorporation of weights w_i in the sums that occur in the LS equations (see below). Statistical theory shows that the weights should be taken as $w_i \propto 1/\sigma_i^2$, where σ_i is the standard deviation in y for the ith calibration value [16]. This ensures that the estimates of the parameters will be minimum variance; any other choice must yield less precise estimates of the calibration parameters and hence of the unknowns to be determined with the calibration curve. Further, incorrect weights lead to wrong estimates of the actual uncertainties [17].

In many bioanalytical techniques, including chromatographic methods, replicate measurements typically display greater spread at large x (and y) than at small [4–6,8,11]. This means that the data variance σ^2 is increasing with x, and the effect of taking this into account through WLS is to place more emphasis on the small-x data that largely determine the intercept for the calibration line. This in turn lowers the percent error of calibration at small x, which can significantly improve unknown determinations in this important region. To handle the data heteroscedasticity, many workers have employed weights proportional to $1/x$, $1/y$, $1/x^2$, or $1/y^2$, often in trial-and-error fashion, with the results judged by a "quality coefficient" (Q), the minimum value of which is said to identify the best

E-mail address: joel.tellinghuisen@vanderbilt.edu.

1570-0232/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2008.07.043

weighting function. The primary purpose of the present contribution is to demonstrate that this assumption about the significance of Q is fundamentally flawed, because this parameter is predisposed to prefer data consistent with its definition.

The quality coefficient has a long history [18]. Knecht and Stork [19] proposed the quantity

$$Q_{\rm rel} = \frac{1}{n-1}\sum\left(\frac{x_{ci}-x_i}{x_i}\right)^2, \tag{1}$$

where x_{ci} is the value of x computed for the ith y_i value from the calibration formula, $x_{ci} = (y_i - a)/b$, and n is the number of calibration data values. This definition was utilized by some later workers [20,21]; however, others [22,23] preferred a definition equivalent to

$$Q_{\rm abs} = \sum(x_{ci}-x_i)^2, \tag{2}$$

which Vankeerberghen and Smeyers-Verbeke noted was appropriate for data having constant uncertainty rather than constant relative uncertainty [24]. Recently, Almeida *et al.* used a quantity similar to $Q_{\rm rel}$, except with the summand being the absolute value of the ratio instead of the square [8]. This choice has been heavily used subsequently in bioanalytical applications of chromatography [25–33]. I will refer to this as $Q'_{\rm rel}$ below. For completeness, I define $Q'_{\rm abs}$ as the analog of $Q_{\rm abs}$ involving the sum of the absolute values.

In view of this heavy reliance on Qs to judge weighting formulas for calibration, it is important to understand their limitations. These Q tests are self-fulfilling: Use of $Q_{\rm abs}$ and $Q'_{\rm abs}$ lead overwhelmingly to the conclusion that the data are homoscedastic, while $Q_{\rm rel}$ and $Q'_{\rm rel}$ pick heteroscedasticity with proportional error (constant coefficient of variance), *regardless of the true error structure of the data*. The reason is that these tests are closely related to the quantities actually minimized in the respective LS fits. In the present work, these limitations are demonstrated through Monte Carlo (MC) computations in which synthetic data having the approximate structure of those used by Almeida *et al.* [8]—six calibration x_i values from 0.1 to 15, with replicates—are generated and then fitted using OLS and WLS with $1/x^2$ weighting, followed by computation of all four Qs defined above. A single MC computation includes 10^5 equivalent data sets, and the computations are performed for data that are homoscedastic or have proportional error, $\sigma_i \propto x_i$, since these are the limiting error structures typically encountered in data that are fitted directly (*i.e.*, without transformation). The tests include dependence on the numbers of replicates (1–8) at each calibration value, and the partitioning of these into calibration and test data. In each case, the statistics of wrong decisions are accumulated for these tests, and for a combined test that shows a much better overall performance. The tests are confined to data that follow a linear response function, and to just these two error structures and their corresponding LS fits. They show that use of just absolute or relative differences alone can be wrong more than 90% of the time; but use of a combined test yields the wrong decision <2% of the time for heteroscedastic data ($\sigma_i \propto x_i$) and <7% for homoscedastic, with squared differences being slightly better than absolute values, and with only two replicates sufficing to achieve this reliability for heteroscedastic data.

Even though the new combined Q test displays much better performance than the currently used versions, it is still largely limited to deciding between constant and proportional data error and is not recommended as a general solution to the weighting choice problem. Methods like variance function estimation from replicates or by generalized LS directly address the problem of determining the data variance and are thus much to be preferred [6,17,34–37].

2. Least squares background

For the response function, $y = a + bx$, the LS equations are obtained by minimizing the sum of weighted, squared residuals,

$$S = \sum w_i\delta_i^2 = \sum w_i(y_i - a - bx_i)^2 \tag{3}$$

with respect to a and b, where y_i is the measured value of y at $x = x_i$. The resulting equations are [1,2]

$$\begin{aligned} aS_w + bS_x &= S_y \\ aS_x + bS_{xx} &= S_{xy} \end{aligned} \tag{4}$$

where $S_w = \sum w_i$, $S_x = \sum w_i x_i$, $S_{xx} = \sum w_i x_i^2$, $S_y = \sum w_i y_i$, and $S_{xy} = \sum w_i x_i y_i$. The solutions to these equations are

$$\begin{aligned} a &= D^{-1}(S_{xx}S_y - S_x S_{xy}) \\ b &= D^{-1}(S_w S_{xy} - S_x S_y) \end{aligned} \tag{5}$$

where $D = S_w\, S_{xx} - S_x^2$. If the data y_i are unbiased estimates of y at x_i, the estimates of the parameters a and b will also be unbiased. If further, the data are normally distributed (*i.e.*, have Gaussian error) about their true y values, the parameter estimates will also be normally distributed about their true values; and if the weights are taken as $w_i = 1/\sigma_i^2$, the parameter estimates will be minimum-variance estimates, with variances given by

$$\begin{aligned} \sigma_a^2 &= S_{xx}/D \\ \sigma_b^2 &= S_y/D \end{aligned} \tag{6}$$

From these equations, it can be seen that OLS is just a special case of WLS. Furthermore, if the σ_i are known absolutely—as, *e.g.*, in Monte Carlo calculations, where they are used to set the scale of the error on the synthetic data—Eq. (6) are exact. In the context of MC computations on such models, this means that one expects histograms of a and b to follow the Gaussian distribution with zero bias and with σ_a and σ_b values given by Eq. (6), within the statistical reliability of the MC computation. For $N = 10^5$ replicate data sets, the MC average of a should fall within $\sigma_a/N^{1/2}$ of its true value 68% of the time, and similarly for b [38]. Sampling estimates of standard deviations have relative uncertainty $(2\nu)^{-1/2}$, where ν = the number of degrees of statistical freedom; thus the MC sampling estimates s_a^2 and s_b^2 should be within 0.22% of the values given by Eq. (6) 68% of the time for $N = 10^5$.

Since the data error is presumed to be known *a priori* in Eq. (6), these expressions constitute *a priori* values of the parameter variances. Usually in OLS one presumes no advanced knowledge of the scale of the error (just that it is constant) and takes $w_i = 1$. Then the data error is estimated from the fit itself, as $s_y^2 = S/\nu = S/(n-2)$. Accordingly one must scale the right-hand sides of Eq. (6) by s_y^2 to obtain the corresponding *a posteriori* estimates s_a^2 and s_b^2 (which vary from data set to data set in MC computations, since s_y^2 also varies). The same scaling procedure is required for heteroscedastic data when the w_i are known in only a relative sense, in which case S/ν is known as the "estimated variance for data of unit weight." From Eqs. (3)–(5) it can be seen that the values for a, b, s_a^2, and s_b^2 are independent of a constant scale factor in the w_i. However, if the w_i are not correctly taken as $\propto\sigma_i^{-2}$, there are two immediate consequences: (1) the parameter variances will be larger than the minimum-variance values; (2) Eq. (6) will not correctly predict the actual results. The magnitudes of these flaws depend upon the structure of the data set and the variability of σ_i and w_i over the range of x_i [17].

With $w_i = 1$, the quantity minimized in OLS is similar to the test quantity in $Q_{\rm abs}$; and the first of Eq. (4) ensures that $\sum(x_{ci} - x_i) = 0$. Similarly the use of $w_i \propto 1/x_i^2$ is nearly equivalent to minimizing the relative error in y; and the second of Eq. (4) can be rewritten

as $\sum \left[(x_{ci} - x_i)/x_i\right] = 0$. It is for these reasons that Q_{abs} and Q_{rel} are predisposed to favor data fitted with $w_i = 1$ and $1/x_i^2$, respectively, regardless of the actual error structure for those data. In many calibration situations the intercept a is small enough to ensure that there is little statistical difference between weighting as $1/x^2$ and $1/y^2$, except that with the latter and noisy data, it may be necessary to iterate the weights, since it is better to evaluate these using $y_{calc,i} = a + bx_i$ than the measured y_i [37]. There is also an obvious problem with including a blank value, $x_i = 0$, when weighting as $1/x^2$.

3. Monte Carlo calculations

The MC computations were done using methods that have been fully described in other recent works [17,37,38]. As was already noted, the model was constructed to resemble that employed in Ref. [8], with x_i values at 0.1, 0.5, 1, 2.5, 5, 10, and 15. The intercept and slope were taken as −0.005 and 0.28, respectively. The synthetic data were obtained by adding random Gaussian error to the true y value at each x_i; this error had $\sigma = 0.1$ (homoscedastic model) or $\sigma = 0.02x_i$ (heteroscedastic). (While specific values are required for these computations, their magnitudes have no effect on the results of the Q tests.) Each data set could contain as many as 8 replicates at each x_i value, and these could be partitioned into calibration and test data. The calibration data were then fitted by OLS, and by WLS using $w_i = 1/x_i^2$; and in each case the Q coefficients were computed for the test data and were then used to pick a preferred weighting. For example, fitting homoscedastic data with OLS and WLS yields $Q_{abs,O}$, $Q_{abs,W}$, $Q_{rel,O}$, and $Q_{rel,W}$. If $Q_{abs,W}/Q_{abs,O} > 1$, the data were deemed to be homoscedastic (a correct decision); if this ratio was <1, they were designated heteroscedastic (incorrect). Similarly, if $Q_{rel,W}/Q_{rel,O} < 1$, the data were heteroscedastic (wrong), *etc.* For the purpose of these tests, the Q and Q' coefficients were treated separately. The wrong decisions resulting from the Q tests were accumulated to yield an overall error rate for each MC run.

Since the error rates are the result of a yes/no decision, they follow binomial statistics, according to which the expected mean for a particular outcome having probability p is pN, and the variance is $Np(1-p)$ [16,39]. When p is expressed as a percent P, the corresponding uncertainty (%) is $N^{-1/2}[P(100-P)]^{1/2}$. Thus, for example, with $N = 10^5$ and $p = 0.5$, the expected mean is 5×10^4 and $\sigma = 158$, yielding $P = 50.00 \pm 0.16\%$. The error drops as P increases or decreases, yielding, *e.g.*, $3.000 \pm 0.054\%$ when $p = 0.03$. These sampling-based uncertainties are too small to show in the graphical illustrations of the results given below.

The computations were programmed in FORTRAN and run on a Compaq laptop computer. MC runs of 10^5 data sets took typically 5–30 s, depending on the number of replicates (up to 8).

4. Results and discussion

Fig. 1 shows the results of the Q tests for the two data error structures, as a function of the number of replicate data sets, with all calibration data both fitted and used to compute the Q values. With a single set of 6 calibration points, each data model confirms its own error structure more than 99.9% of the time, when tested with its "own" Q. At the same time, it gives the wrong decision about as often when tested with the other Q. To be more specific, homoscedastic data almost always yield a smaller Q_{abs} when fitted with OLS than with WLS (as they should). But the same data almost as often yield a smaller Q_{rel} when fitted with WLS (hence the wrong decision). The same holds for the heteroscedastic data, now giving the right result almost always when tested with Q_{rel}

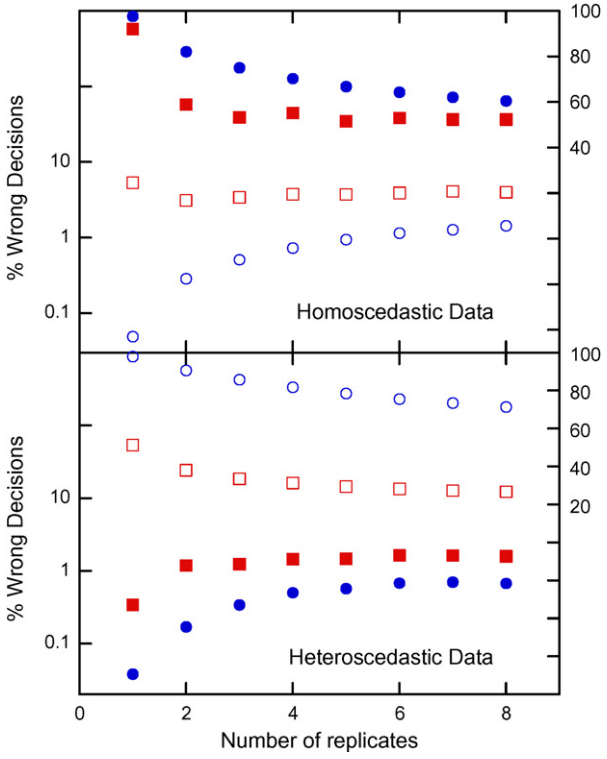

Fig. 1. Results of applying the Q tests to homoscedastic (top) and heteroscedastic data models, as functions of the number of replicate data sets taken. Round symbols refer to sums of squared test quantities (Q) and square to sums of absolute values (Q'); open points represent Q_{abs} (Q'_{abs}), while filled symbols represent Q_{rel} (Q'_{rel}). In each case the data are fitted with OLS ($w_i = 1$) and with WLS ($w_i = 1/x_i^2$). The exact standard errors in a and b are $0.05845/r^{1/2}$ and $0.007583/r^{1/2}$, respectively for the homoscedastic model, and $0.002278/r^{1/2}$ and $0.009496/r^{1/2}$, respectively for heteroscedastic, where r is the number of fitted replicates. Note logarithmic scales to left for the small error rates near bottom in each grid, and absolute scales (right) for the large error rates at top.

but wrong when tested with Q_{abs}. The summing of absolute values (yielding Q') instead of squares is less definitive in both directions.

This behavior can be seen as a consequence of the importance of the end points. WLS ensures that the first residual will be small regardless of the true error structure; at the same time it allows the last residual to become large. Accordingly, the Q_{abs} test strongly prefers OLS for both error structures, while Q_{rel} just as strongly prefers WLS, confirming the self-fulfilling nature of these tests. These tendencies are reduced somewhat when replicates are taken but are still strong, in disagreement with my earlier speculations [17].

Partitioning the replicates into calibration and test data does not solve the problem, as shown in Fig. 2. The number of wrong decisions is greatly reduced but is still significant (~40%) in both cases; and the preference for the correct decision under proper weighting is somewhat reduced. The dependence on the manner in which the replicates are partitioned is surprisingly weak.

Closer examination of the statistics of the Q tests suggests a way of combining the absolute and relative tests in a way that greatly reduces the overall error rate. The preference for the correct weighting is on average much stronger than the (erroneous) preference for the wrong weighting. For example, with 5 replicates, all fitted and all used to compute the needed Qs for the points at $X = 5$ in Figs. 1 and 2, the average value of $Q_{rel,O}/Q_{rel,W}$ for heteroscedastic data is 19.8, while that for $Q_{abs,O}/Q_{abs,W}$ is 0.94 ± 0.09. The latter translates into an error rate of 79%, but if the two tests are com-

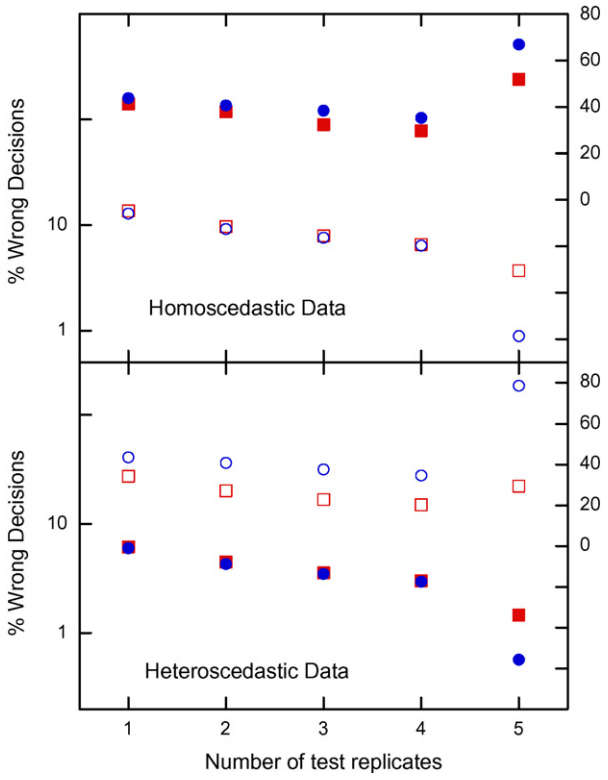

Fig. 2. Testing of data for 5 replicates, partitioned into the indicated number of test sets, with the remainder being fitted to obtain the calibration curve. The last points are reproduced from Fig. 1, for 5 replicate sets being used for both the calibration fit and the test calculations. Other quantities are as indicated in the caption to Fig. 1.

bined, using a product of the two ratios, the former wins out most of the time, yielding an overall error rate of only 1.4%. A similar behavior occurs for homoscedastic data, but the discrimination is less efficient, as shown in Fig. 3. With this combined test, use of all replicates for both fitting and testing gave better discrimination than partitioning into calibration and test sets, so results are shown just for the former case. Again, Q' rarely performs as well as Q. It is also interesting that there is little need for more than

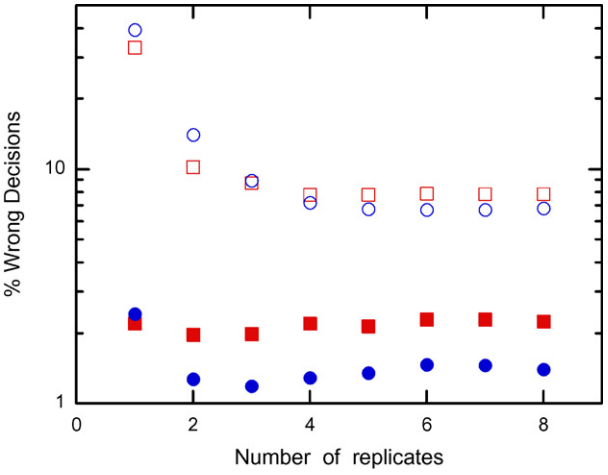

Fig. 3. Results of applying combined Q test to data that are homoscedastic (open points) or heteroscedastic (solid), using squared test quantities (Q, round) or absolute values (Q', square). For each number of replicates, all data are both fitted and used to compute the Q values; partitioning into fit and test subsets yielded poorer performance. Note logarithmic ordinate scale.

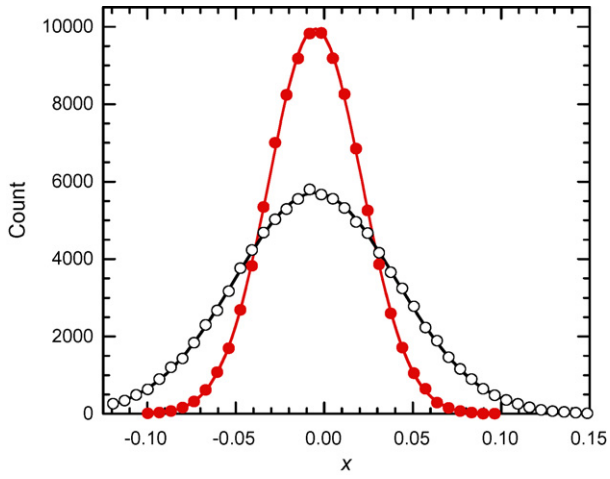

Fig. 4. Histograms of estimated a (intercept) values from 10^5 MC data sets having five replicates (30 points total), all fitted to the linear response function, with constant $\sigma_i = 0.1$. The solid points represent the results from OLS, the open from WLS using $w_i = 1/x_i^2$. Both histograms are well fitted by Gaussians and show no significant bias in the average of a (=−0.005), but σ_a for the weighted fit is 75% larger than the minimum-variance value (solid points and curve).

three replicates, with performance actually deteriorating slightly for more replicates in the case of heteroscedastic data.

The manner in which incorrect weighting affects the LS results is worth examining in more detail, especially in light of the realization that $1/x^2$ weighting must surely make the fitted line better match the data at the first calibration value ($x = 0.1$) in a given data set. As Fig. 4 shows, this power is illusory, because when different, statistically equivalent data sets are similarly fitted, the resulting a values move in accord with the actual data, and the result is a 75% increase in σ_a over the minimum-variance value achieved with proper weighting (OLS in this case). And yet, an analyst who uses Eq. (6) (scaled by s_y^2) to estimate σ_a concludes that it is much smaller than it actually is—43% of its true value. At the same time, the loss of precision in the slope (not shown) is even greater—a factor of 8 increase in σ_b over minimum variance; but now the analyst using Eq. (6) thinks it is even worse, by another factor of 3. These comparisons illustrate both of the consequences of incorrect weighting noted earlier.

Since the present tests deal with the extremes of homo- and heteroscedasticity for nontransformed data, we would expect the discriminating ability to be even worse for intermediate heteroscedasticity. Accordingly, when the data (5 replicates, all fitted and used to compute Q) were given uncertainty $\sigma_i \propto x^{1/2}$ ($w_i = 1/x$) and tested for weighting as 1, $1/x$, and $1/x^2$, they yielded the correct weighting at best 46.7% of the time (Q_{rel}) and worst 41.3% (Q_{abs}). The same two Q tests preferred $1/x^2$ weighting in 50.8% of the cases and OLS in 51.0%, respectively—again illustrating the proclivity of these tests to select their own natural data type.

As an illustration of results for a single data set, the 30 data points used in the original calculations by Almeida *et al.* [8] were regenerated approximately from their paper and subjected to the Q_{rel}, Q_{abs}, and Q'_{abs} tests in addition to the Q'_{rel} test they used. Just the constant, $1/x^{1/2}$, $1/x$, and $1/x^2$ weightings were tested here. The Q'_{rel} test confirmed their results (their Table 2), and the Q_{rel} test preferred $1/x^2$ weighting even more strongly. The Q'_{abs} test also preferred $1/x^2$ weighting, but the Q_{abs} test picked OLS over $1/x^2$ weighting. These outcomes are all statistically consistent with results in Fig. 1 for heteroscedastic data and 5 replicates. However, in the Q'_{abs} test, $1/x^2$ weighting fared only very slightly better than $1/x$, and Q_{abs} picked $1/x^{1/2}$ as best of all.

5. Conclusion

Monte Carlo computations on a simple linear calibration model confirm that quality coefficients widely used to choose optimal weighting formulas fail miserably, because such coefficients are predisposed to find homoscedasticity when they are based on absolute residuals and proportional error when they are based on relative residuals. This conclusion is not much affected by partitioning the data into calibration and test subsets. However, by using both the absolute and the relative Q tests, it is possible to discriminate correctly in favor of truly heteroscedastic data ($\sigma_i \propto x_i$) with an error rate <2%, and homoscedastic data with error <7%. The use of squares of the test quantities yields slightly better performance than absolute values.

Even though the new Q test proves fairly reliable, I do not recommend it, because it attempts to solve the problem of determining the weighting formula the wrong way, and it tacitly assumes that the weighting function is simpler than it likely is. The goal of such efforts should be determination of the actual variance function for the data, from which it follows that $w_i = 1/\sigma_i^2$. Data variance functions almost always display constant σ in the low-signal limit and often show proportional error in the high-signal limit [40–44]. In recent work I have shown that as few as three replicates at each of 6 calibration values suffice to determine a two-parameter variance function with enough reliability to yield negligible loss of precision in the subsequent calibration fit [37]. Thus, in many studies where replicate data have been used to assess weighting formulas through Q tests, the authors could and should have used the same data to determine the data variance function, thereby obtaining fully reliable assessments of the weights for their calibration fits.

Those who have used Q tests based on relative residuals to conclude that their chromatographic data have proportional error should not now despair that this test has deceived them, because many studies of residuals have shown that chromatographic data in typical working ranges are indeed dominated by proportional error. However, some of the most demanding analytical problems push the limits of detection, and in such cases the error in this limit becomes all-important. As already noted, homoscedasticity usually rules in the low-signal limit. This behavior has been confirmed for HPLC data in a recent study of variance functions for several analytes over detection ranges spanning four orders of magnitude [45].

References

[1] J.N. Miller, Analyst 116 (1991) 3.
[2] K. Danzer, L.A. Currie, Pure Appl. Chem. 70 (1998) 993.
[3] H.T. Karnes, C. March, J. Pharm. Biomed. Anal. 9 (1991) 911.
[4] N.V. Nagaraja, J.K. Paliwal, R.C. Gupta, J. Pharm. Biomed. Anal. 20 (1999) 433.
[5] S. Sadray, S. Rezaee, S. Rezakhah, J. Chromatogr. B 787 (2003) 293.
[6] K. Baumann, H. Wätzig, J. Chromatogr. A 700 (1995) 9.
[7] M.M. Castel-Branco, A.M. Almeida, A.C. Falcão, T.A. Macedo, M.M. Caramona, F.G. Lopez, J. Chromatogr. B 755 (2001) 119.
[8] A.M. Almeida, M.M. Castel-Branco, A.C. Falcão, J. Chromatogr. B 774 (2002) 215.
[9] M. Zhou, G. Wei, Y. Liu, Y. Sun, S. Xiao, L. Lu, C. Liu, D. Zhong, J. Chromatogr. B 798 (2003) 43.
[10] M.J. Burns, H. Valdivia, N. Harris, Anal. Bioanal. Chem. 378 (2004) 1616.
[11] M.K. Kiser, J.W. Dolan, LC–GC North Am. 22 (2004) 112.
[12] Guidance for Industry, Bioanalytical Methods Validation, United States Food and Drug Administration, Washington, DC, 2001, p. 6 (www.fda.gov/cder/guidance/).
[13] Eurachem Guide—The Fitness for Purpose of Analytical Methods: A Laboratory Guide to Method Validation and Related Topics, LGC, Teddington, 1998.
[14] Eurachem Guide—Quantifying Uncertainty in Analytical Measurement, Laboratory of the Government Chemist, London, 2000.
[15] EC Regulation 49/2000/EC.
[16] A.M. Mood, F.A. Graybill, Introduction to the Theory of Statistics, 2nd ed., McGraw-Hill, New York, 1963.
[17] J. Tellinghuisen, Analyst 132 (2007) 536.
[18] J.O. De Beer, T.R. De Beer, L. Goeyens, Anal. Chim. Acta 584 (2007) 57.
[19] J. Knecht, G. Stork, Fresenius J. Anal. Chem. 270 (1974) 97.
[20] L. De Galan, H.P.J. Van Dalen, G.R. Kornblum, Analyst 110 (1985) 323.
[21] Y. Hu, J. Smeyers-Verbeke, D.L. Massart, J. Anal. At. Spectrom. 4 (1989) 605.
[22] N.J. Miller-Ihli, T.C. O'Haver, J.M. Harnly, Spectrochim. Acta B 39 (1984) 1603.
[23] S.R. Bysouth, J.F. Tyson, J. Anal. At. Spectrom. 1 (1986) 85.
[24] P. Vankeerberghen, J. Smeyers-Verbeke, Chemom. Intell. Lab. Syst. 15 (1992) 195.
[25] A. Rakic, B. Miljkovic, M. Pokrajac, K. Vucicevic, J. Pharm. Biomed. Anal. 43 (2007) 1416.
[26] J. Wang, D. Wotherspoon, J. AOAC Int. 86 (2007) 550.
[27] E. Scheuch, J. Spieker, M. Venner, W. Siegmund, J. Chromatogr. B 850 (2007) 464.
[28] D. Durden, J. Chromatogr. B 850 (2007) 134.
[29] R.M. de Almeida, M. Yonamine, J. Chromatogr. B 853 (2007) 260.
[30] J. Wang, D. Leung, Rapid Commun. Mass Spectrom. 21 (2007) 3213.
[31] M. Zhou, X. Chen, D. Zhong, J. Chromatogr. B 854 (2007) 219.
[32] S.M.R. Wille, P. Van Hee, H.M. Neels, C.H. Van Peteghem, W.E. Lambert, J. Chromatog. A 1176 (2007) 236.
[33] M. Rosland, P. Szeto, R. Procyshyn, A.M. Barr, K.M. Wasan, Drug Dev. Ind. Pharm. 33 (2007) 1158.
[34] R.J. Carroll, D. Ruppert, Transformation and Weighting in Regression, Chapman and Hall, New York, 1988.
[35] M. Davidian, P.D. Haaland, Chemom. Intell. Lab. Syst. 9 (1990) 231.
[36] Ph. Hubert, J.-J. Nguyen-Huu, B. Boulanger, E. Chapuzet, N. Cohen, P.-A. Compagnon, W. Dewé, M. Feinberg, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, J. Pharm. Biomed. Anal. 45 (2007) 82.
[37] J. Tellinghuisen, Analyst 133 (2008) 161.
[38] J. Tellinghuisen, J. Chem. Educ. 82 (2005) 157.
[39] P.R. Bevington, Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill, New York, 1969.
[40] M. Thompson, Analyst 113 (1988) 1579.
[41] J.D. Ingle Jr., S.R. Crouch, Spectrochemical Analysis, Prentice-Hall, Englewood Cliffs, New Jersey, 1988.
[42] J. Tellinghuisen, Appl. Spectrosc. 54 (2000) 431.
[43] P. Modamio, C.F. Lastra, E.L. Marino, J. Pharm. Biomed. Anal. 14 (1996) 401.
[44] J. Tellinghuisen, Anal. Biochem. 343 (2005) 106.
[45] Q.C. Zeng, E. Zhang, H. Dong, J. Tellinghuisen, J. Chromatogr. A, in press.