

## Fitting a linear functional relationship to data with error on both variables

It is fairly well known that a basic assumption of regression is that the  $y$ -values (the dependent- or response variable) are random variables while the  $x$ -values (the independent- or predictor variable) should be error-free. This model often prevails (or is approximated to) in analytical chemistry applications, for example in many calibrations. If the condition is violated, however, the results of the regression are in principle incorrect and in practice can be sufficiently incorrect to be misleading. What is far less well known is that a general method, functional relationship estimation by maximum likelihood (FREML), is available for use when the regression assumption is incorrect. FREML provides estimates of the intercept ( $a$ ) and slope ( $b$ ) of the line, plus their standard errors, that do not suffer from the biases introduced by the inappropriate use of regression. The method is symmetric in that the  $x$ - and  $y$ -variables can be interchanged without affecting the outcome. It is capable of handling heteroscedastic data, that is, data points with different precisions.

### Regression and functional relationship models

Normal weighted regression is based on a model of the paired data ( $x_i, y_i, i = 1, \dots, n$ ) such that

$$y_i = a + bx_i + e_i$$

where  $a$  and  $b$  are the parameters describing the true line and  $e_i$  is a random normal error of variance  $\text{var}(y_i)$ . Estimates ( $a, b$ ) of  $a$  and  $b$  are obtained by minimising the function

$$\sum_{i=1}^n (y_i - a - bx_i)^2 / \text{var}(y_i)$$

with respect to  $a$  and  $b$ . The minimisation is done by a simple application of calculus to give the familiar regression equations.

In functional relationship estimation we use a different model, namely

$$\begin{aligned}x_i &= u_i + e_i \\y_i &= v_i + h_i \\v_i &= a + bu_i\end{aligned}$$

where  $e_i$  and  $h_i$  are independent, normally distributed errors of variances  $k_i$  and  $l_i$  respectively. The functional relationship is found by minimising

$$\sum_{i=1}^n (x_i - u_i)^2 / k_i + \sum_{i=1}^n (y_i - a - bu_i)^2 / l_i.$$

This formulation follows from the maximum likelihood principle, which is a statistical estimation procedure that is more general than the familiar least squares method. Unlike regression, however, the above minimisation cannot be solved algebraically, but requires an

iterative numerical method. Details of the procedure can be found elsewhere.<sup>1</sup> Fortunately it is simple to program: both a Minitab macro and a compiled Fortran program to carry out FREML are available in AMC Software.

### Analytical applications of FREML

Typical applications of FREML in analytical chemistry are (i) calibration with solid reference materials (where the uncertainty on the reference value may be considerable) and (ii) comparing the results of two methods (or analysts, or laboratories, etc.) over an appreciable concentration range and characterising the bias, if any, between them. Both of these activities may require tests of significance on  $a$  and  $b$  and for lack of fit. FREML can accommodate these requirements by providing standard errors  $se(a), se(b)$  and a lack of fit statistic.

For example, we could test the hypothesis  $H_0: a = 0$  (i.e., that the line passes through the origin of the graph) by calculating  $z_a = a/se(a)$ . Likewise we could test  $H_0: b = 1$  (that the slope of the line is unity) with  $z_b = (b-1)/se(b)$ . The  $z$ -values can be interpreted as a standard normal distribution, so for 95% confidence we should use a critical value of 1.96 for  $|z|$ .

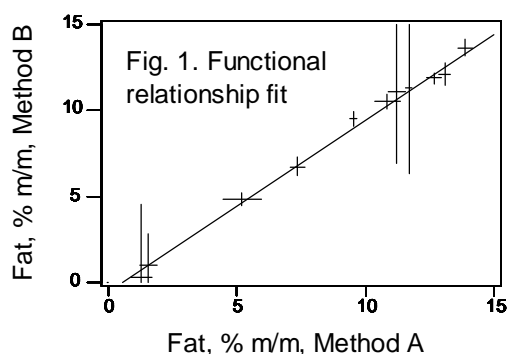
The lack of fit statistic is the sum of the squares of the scaled residuals, which we can treat approximately as a chi-squared variable with  $n-2$  degrees of freedom. A significantly high value of this statistic suggests either that the variance estimates were optimistically low (a common failing) or lack of fit. The latter could be brought about *inter alia* by uncorrected interference or by non-linearity in the true form of the relationship. An examination of a plot of scaled residuals would help to elucidate that difference.

### Example

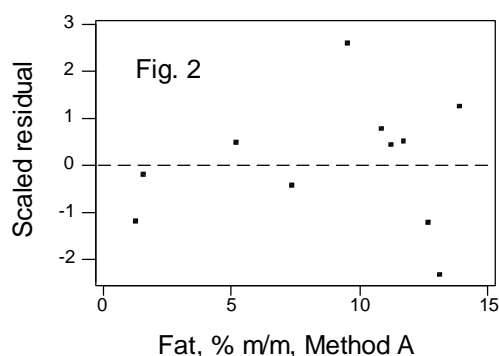
Consider results from the determination of fat in 11 different foodstuffs by two different methods. The experiment was carried out to determine whether there was bias between the results of the two methods over the whole of the usual concentration range. Each material was analysed by one of the two methods in a number of different laboratories, and Table 1 shows the mean result and the variance of the mean result for each material and both methods. The wide range of the variances is due mainly to the variation in the number of results that provide the mean. The general level of relative standard deviation is high, because the data reflect reproducibility (between laboratory) precision.

An  $x$ - $y$  plot of the results and the fitted FREML line are shown in Figure 1. The error bars show the 95% confidence interval of each mean. The statistics derived from the FREML fit are shown in Table 2. (The same outcome is obtained if the  $x$ - and  $y$ -variables are interchanged.) Taking a critical value of  $|z| = 1.96$  for an approximately 95% confidence level, we see that the intercept is clearly significantly different from zero, indicating the presence of a rotational (constant) bias between the methods over the concentration range studied. In contrast, the slope of the line is not

significantly different from unity, showing that there is no detectable rotational (proportional) bias. In other words, the FREML line is, for all practical purposes, parallel to but distinct from the  $y = x$  line.



In addition to the above, we find that the 'lack-of-fit statistic' takes the value of 18.2, suggesting a marginally significant, but not worrying, lack of fit. The scaled residuals are shown in Figure 2. There is no obvious pattern in the results, and the lack of fit is accounted for by two rather (but not extremely) large residuals.



It is interesting to note that weighted regression applied to the above data (both  $y$  on  $x$ , and  $x$  on  $y$ ) tells us (incorrectly) that  $b$  is significantly different from unity, indicating a rotational bias.

#### Calibration with reference materials

Calibrating with reference materials is common in the analysis of solids. However, the relative uncertainty of the reference values may be comparable with, or even greater than, that of the measured responses. In such instances regression is inappropriate for estimating the calibration function, but FREML could be used.

The lack of fit statistic would be particularly useful here, because two different types of lack of fit would be plausible. The most obvious type would be caused by a non-linear true calibration function. That could be easily seen as a curved trend on a plot of the scaled residuals against the  $x$ -values or the  $y$ -values. In the other type of lack of fit, which is due to unique matrix effects in each calibrator (and is quite common), the residual plot would

show a random distribution. This latter style of lack of fit would be useful information to have, but its validity would depend on the variance estimates being realistic.

#### How to estimate weights

In the example above the variances were well estimated from (usually) moderate numbers of replicated results. In some instances where the use of FREML was indicated, we may have only one or a small number of results per point. How can we estimate the weights? If we first assume that the data are homoscedastic (that is, with uniform variances), we could use a common variance for each of the  $y$ -values and another for each of the  $x$ -values. Such variances would be relatively easy to estimate or simply guess.

If the data turn out to be heteroscedastic, a scaled residual plot may demonstrate the fact. Then different variances would have to be attached to each  $x$ -value and  $y$ -value. If the data are grossly heteroscedastic, even rough estimates of variance (for instance from duplicated data) will greatly improve the resulting statistics. Moreover, in calibration, we can often use a simple linear model of standard deviation as a function of concentration, which could be used to smooth crude point-estimates derived from small numbers of data.

#### Reference

1. B D Ripley and M Thompson, *Analyst*, 1987, **112**, 337-383.

Further information can be obtained also from:

P Sprent, *Contemporary Math.*, 1990, **112**, 3-15.

J Riu and F X Rius, *J. Chemomet.*, 1995, **9**, 343-362.

K I Mahon, *Internat. Geol. Review*, 1996, **38**, 293-303.

**Table 1 Determination of fat (% m/m) in 11 foodstuffs**

x (Method A)	y (Method B)	var(x)	var(y)
13.11	12.12	0.0100	0.0243
7.36	6.72	0.0173	0.0370
11.21	11.10	0.0294	0.9296
1.28	0.30	0.0422	0.1105
9.55	9.52	0.0049	0.0351
11.69	11.35	0.0045	0.1568
1.54	0.98	0.0268	0.0200
13.89	13.64	0.0223	0.0392
12.67	11.87	0.0211	0.0182
5.20	4.83	0.0990	0.0110
10.85	10.51	0.0448	0.0267

**Table 2. Functional relationship statistics from the fat data**

	Result	se	$H_0$	$z$
Intercept ( $a$ )	-0.516	0.080	$a = 0$	-6.45
Slope ( $b$ )	0.9966	0.0078	$b = 1$	-0.43

AMC Technical Briefs are informal but authoritative bulletins on technical matters of interest to the analytical community. Correspondence should be addressed to: The Secretary, The Analytical Methods Committee, The Royal Society of Chemistry, Burlington House, Piccadilly, London W1J 0BA. AMC Technical Briefs may be freely reproduced and distributed in exactly the same form as published here, in print or electronic media, without formal permission from the Royal Society of Chemistry. Copies must not be offered for sale and the copyright notice must not be removed or obscured in any way. Any other reuse of this document, in whole or in part, requires permission in advance from the Royal Society of Chemistry.

Other AMC Technical Briefs can be found on: [www.rsc.org/lap/rsccom/amc/amc\\_index.htm](http://www.rsc.org/lap/rsccom/amc/amc_index.htm)

