

THE QUASIMEME LABORATORY PERFORMANCE STUDY OF PAHS IN SEDIMENTS, 1996-1999.

by Robin Law, Carole Kelly, and Mike Nicholson*



Burnham Laboratory

Statistical Analysis

Analysis of z-scores

The statistical analysis focuses on the z-scores calculated from the concentrations reported by each laboratory for each determinand measured on each material.

Figure 1 shows a typical distribution of z-scores for e.g. Chrysenes, revealing more variability than would be expected given the assigned value and assigned total error. There are more scores outside the range ± 2 corresponding to satisfactory performance than would be expected, and a small peak of z-scores greater than 6, corresponding to extremely unsatisfactory performance.

A more formal statistical analysis of the percentages of results that were satisfactory showed statistically significant differences amongst determinands and materials, and between analytical methods.

Figure 2 shows the percentages of satisfactory results by analytical method for each determinand. HPLC (overall 75% satisfactory) performs better than GC (overall 63% satisfactory), although less clearly so for determinands 11 to 16.

Figure 3 and Figure 4 show that the picture is more complicated. Figure 3 shows the percentage of satisfactory results by analytical method for each material, but only for determinands 1 to 10. HPLC performs better than GC for the earlier releases of marine sediments (QPH0010MS to QPH0017MS). For the later releases of marine sediments (QPH0018MS to QPH0021MS) and the releases of biological tissue (QPH0008BT to QPH0011BT) the position is more ambiguous.

In Figure 4, the data are restricted to determinands 11 to 16. These were only analysed for the later releases of marine sediments and the releases of biological tissue. Now, the performances for each analytical method are more similar to each other, with considerable variation in performance for different materials.

Variation in Between-Laboratory Coefficient of Variance with Concentration

Robust means and robust between-laboratory standard deviations were calculated for each determinand measured in each material. Figure 5 shows the corresponding percentage coefficients of variance plotted against the robust mean on a \log_{10} scale. The decreasing trend with concentration is summarised by the solid line, corresponding to

$$CV\% = 2 \times 2C^{-0.1505}$$

where

$$CV\% = 2C^{-0.1505}$$

is the Horwitz curve, shown by the dotted line. This curve is proposed by Horwitz as a target for achievable between-laboratory variability independent of analyte, method, matrix or time. This suggests a potential for a 50% improvement in the variability between laboratories with further method development.

Introduction

The QUASIMEME LPS for PAHs in sediments comprised six exercises during 1996-1999, approximately two per year. In each exercise two sediment samples were circulated. Initially the PAHs determined comprised the core group of 10 parent compounds included in the earlier stages of the programme, but in the later rounds this was increased to 17 parent PAH. Approximately 20 laboratories submitted data for each exercise.

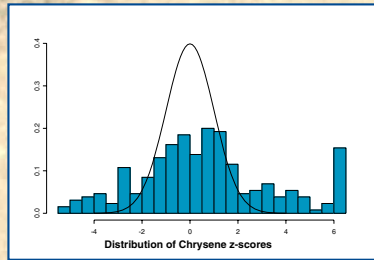


Figure 1 Distribution of z-scores, defined as $z = \frac{\text{measured concentration} - \text{assigned value}}{\text{total error}}$

for Chrysenes, combined over all laboratories and all materials. Overlaid is the density function corresponding to a Normal distribution with a mean of zero and a standard deviation of unity.

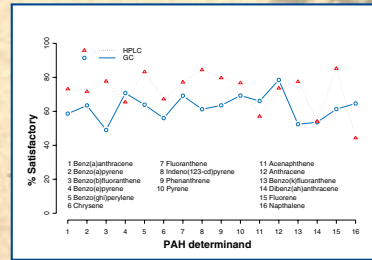


Figure 2 Percentages of satisfactory results taken over all laboratories and materials for each determinand and analytical method. Satisfactory is defined in terms of the z-scores i.e. $|z| < 2$.

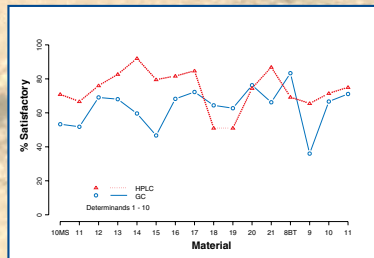


Figure 3 Percentages of satisfactory results plotted by analytical method against each material (QPH0010MS to QPH0021MS and QPS0008BT to QPS0011BT). The percentages are calculated over all laboratories, but for determinands 1-10 (as defined in Figure 2).

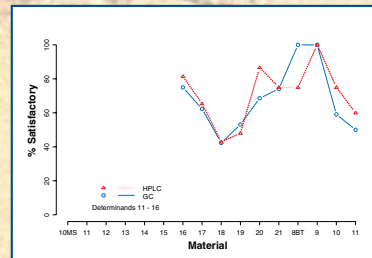


Figure 4 Percentages of satisfactory results plotted against each material by analytical method. The percentages are calculated over all laboratories, but for determinands 11-16 (as defined in Figure 2).

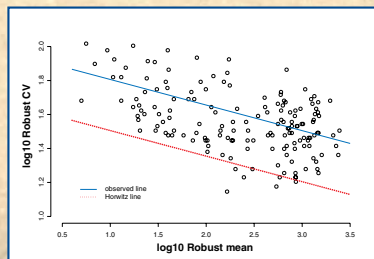


Figure 5 Plot of the robust coefficients of variation against robust mean concentration on a \log_{10} scale. The dotted line corresponds to the Horwitz curve. This defines a target for the between-laboratory variance as a function of concentration, found by Horwitz to apply independently of analyte, method, matrix and time. The solid line is a fit to the data, and corresponds to robust CVs at twice the level predicted by the Horwitz curve.



Lowestoft Laboratory

Laboratory Code	No. of Obs.	No. of Obs. by Method	% Satisfactory $ z < 2$	% Questionable $2 < z < 3$	% Unsatisfactory $3 < z < 6$	% Extremely Unsatisfactory $ z > 6$
Q311A	29	29	100			
Q138C	109	109	97	3		
Q107A	40	40	95	5		
Q357A	19	19	95	5		
Q115A	91	71	20	93	5	1
Q143A	78	78	91	4	5	
Q251A	31	31	90	6	3	
Q131D	40	40	90	5	5	
Q114A	140	140	88	7	4	1
Q152A	195	195	0	88	4	8
Q127A	45	45	87	7	4	2
Q111B	65	65	86	9	5	
Q134B	20	20	85	5	5	5
Q116A	91	91	80	19	1	
Q126C	149	149	80	14	6	
Q210A	117	117	79	12	7	2
Q104C	46	46	74	7	13	7
Q131B	161	161	71	11	16	2
Q275A	79	79	67	9	16	8
Q108A	60	48	12	65	17	18
Q116A	46	46	63	15	22	
Q130B	70	70	63	19	16	3
Q121A	46	46	61	20	11	9
Q153A	182	164	18	60	21	14
Q114D	94	94	57	15	18	10
Q110A	40	40	55	13	15	18
Q205A	44	44	55	16	23	7
Q234A	126	126	54	2	3	41
Q125A	102	102	51	22	10	18
Q224A	89	89	49	16	26	9
Q157A	37	37	46	22	30	3
Q113B	82	82	45	9	7	39
Q134D	51	51	45	27	22	6
Q152C	18	18	44	28	28	
Q240A	78	78	42	5	29	23
Q128A	59	59	41	10	12	37
Q134A	125	71	54	38	10	38
Q250A	10	10	30	10	50	20
Q204A	20	20	25	10	40	25
Q226A	79	79	8	16	24	52
Q242A	1	1	0	100		

Table 1 Summary of Laboratory Performance. The numbers of analyses made by each laboratory have also been split by analytical method. Laboratories have been ranked in descending order by the percentage of satisfactory results.

Summary of Laboratory Performance

Table 1 summarises the overall performance by each laboratory. They have been ranked in descending order of their percentage of satisfactory results. Also shown are the percentages of questionable, unsatisfactory and extremely unsatisfactory performance. This latter category is likely to reflect a gross error more than poor analytical performance: for instance data transcription and reporting errors. These rankings should be used with caution, as the numbers of observations vary between laboratories, and laboratories analysed different subsets of determinands and materials.

The Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Burnham Laboratory, Remembrance Avenue, Burnham on Crouch, Essex CM0 8HA, UK.

*The Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK <http://www.cefas.co.uk>

