Pergamon

J. Aerosol Sci., Vol. 26, No. 1, pp. 109–126, 1995
Elsevier Science Ltd
Printed in Great Britain

0021-8502(94)E0071-5

THE PERFORMANCE EVALUATION OF AEROSOL SAMPLERS TESTED WITH MONODISPERSE AEROSOLS

L. C. Kenny* and D. L. Bartley†

*Health and Safety Laboratory, Health and Safety Executive, Broad Lane, Sheffield S3 7HQ, U.K.
and † National Institute for Occupational Safety and Health, 4676 Columbia Parkway, Cincinnati, OH 45226, U.S.A.

(*First received* 22 *November* 1993; *and in final form* 28 *March* 1994)

Abstract—Performance requirements are currently being established for instruments used for the particle size-selective sampling of potentially hazardous aerosols in workplaces. In this paper, two methods for the estimation of the sampler performance characteristics bias, imprecision and accuracy are described, for aerosol sampler evaluation data obtained from experiments carried out with monodisperse test aerosols. Application of the methods is illustrated using experimental data for two sampler types, the MRE 113A static respirable sampler and the PM-10 ambient aerosol sampler. Where both methods can be applied to the same data, the estimates obtained using the two methods are in broad agreement. The small number of data points obtained in typical monodisperse-aerosol experiments leads to large uncertainties in the estimated performance characteristics.

1. INTRODUCTION

Target sampling conventions have been defined for the health-related sampling of aerosols in workplaces and in the general environment (CEN, 1992; ISO, 1992; ACGIH, 1993; USEPA, 1987). The performance of a size-selective aerosol sampler may be described in terms of the bias and accuracy in aerosol mass concentrations measured with the sampler, relative to the concentrations which would have been obtained by using an ideal sampler perfectly following the relevant sampling convention (CEN, 1993). Since these ideal samplers do not actually exist, the performance of real samplers cannot be determined by direct comparison. Instead, the performance characteristics are computed numerically, using experimental data from measurements of the sampling efficiency as a function of particle aerodynamic diameter. The sampling efficiency is defined, for each particle aerodynamic diameter, as the ratio of aerosol mass concentration collected for measurement by the sampling process, to the corresponding ambient (i.e. undisturbed outside the sampler) concentration.

Evaluations of sampler bias and accuracy reported to date have been derived from experimental determinations of sampler efficiency which use polydisperse aerosols (Lidén and Kenny, 1992; Bartley *et al.*, 1994a). Current techniques for accurate and precise measurement of particle aerodynamic diameter restrict the size range for polydisperse sampler efficiency measurements to approximately 0.5–20 μm, whereas sampler efficiency measurements with monodisperse aerosols can be carried out for particle aerodynamic diameters up to ~ 100 μm (Mark and Vincent, 1986). In addition, it is extremely difficult to generate a uniform polydisperse aerosol, containing large particles, in a wind tunnel. In practice this means that whilst respirable samplers may be tested using polydisperse aerosols, at present thoracic, PM-10 and inhalable sampler inlets can only be tested with monodisperse aerosols. The laboratory evaluation of PM-10 sampler performance, using monodisperse aerosols, is currently limited to estimating the sampler bias (USEPA, 1987).

In general terms, the accuracy of a sampler describes the range around the true aerosol concentration in which the measurements obtained are likely to lie. More precise definitions

© British Crown Copyright.

are given by CEN (1993) and Gundersen and Anderson (1980). In order to calculate the accuracy, it is necessary to estimate four quantities: the bias of the sampler, the imprecision of repeat measurements with the sampler, and the uncertainties in the determination of bias and imprecision. The estimation of these four quantities from sampler efficiency measurements obtained with polydisperse aerosols is relatively straightforward, and has been described in the literature by, for example, Lidén and Kenny (1992), Bartley *et al.* (1994a) and others. For experiments using monodisperse aerosols, the estimation of sampler bias is also straightforward, but the estimation of the remaining three quantities is more complex than in the polydisperse case, principally because the number of data points is much lower.

This paper describes two methods for the estimation of sampler bias, imprecision and their uncertainties, from monodisperse sampler evaluation data. The two methods are suited to different experimental designs, although there are some designs for which either method can be used. Application of the methods is illustrated using published data from experimental performance tests of two types of sampler. The first example uses data for the MRE 113A sampler, a static sampler used to monitor respirable dust concentrations in British coal mines. The second example uses data for the PM-10 sampler, a static sampler used to monitor ambient aerosol concentrations in the United States of America. The two calculation methods give the same results, provided that the same model for underlying variation in the experimental data is applied in each case. The effect which the choice of this model can have on the performance evaluation is illustrated, and implications for the experimental design of sampler performance evaluations using monodisperse aerosols are considered.

2. DEFINITIONS OF SAMPLER PERFORMANCE CHARACTERISTICS

Each individual specimen s of a given type of sampler has a true, but unknown, efficiency curve $E_s(D)$, where D is the particle aerodynamic diameter. The true mean efficiency curve for the sampler type (averaged over many specimens) is $\bar{E}(D)$. If a single sampler specimen s is used to sample an arbitrary aerosol with normalised size distribution $A(D)$ many times, then the mean measured concentration (expressed as a fraction of the total aerosol mass concentration) will be

$$\bar{C}_s = \int_0^\infty E_s(D)\, A(D)\, \mathrm{d}D. \tag{1}$$

If a large number of specimens of the sampler type are used to sample the same aerosol under identical conditions, then the mean measured concentration will be

$$\bar{C} = \int_0^\infty \bar{E}(D)\, A(D)\, \mathrm{d}D. \tag{2}$$

Thus $\bar{C}$ (estimated as $\hat{C}$) is the mean for the sampler *type*, and $\bar{C}_s$ (estimated as $\hat{C}_s$) is the mean for a single *specimen* of the sampler type. An ideal sampling instrument, perfectly following the desired sampling convention $F(D)$, would give the ideal concentration value C_{std} when sampling aerosol $A(D)$:

$$C_{\mathrm{std}} = \int_0^\infty F(D)\, A(D)\, \mathrm{d}D. \tag{3}$$

The bias of the sampler type is defined as

$$\Delta = \frac{\bar{C} - C_{\mathrm{std}}}{C_{\mathrm{std}}}. \tag{4}$$

The physical differences between different specimens of the sampler type are embodied in the inter-specimen variance (assumed normally distributed with zero mean):

$$\sigma_s^2 \equiv \lim_{S \to \infty} \left[\frac{1}{S} \sum (\bar{C}_s - \bar{C})^2 \right]. \tag{5}$$

The sampler imprecision is usually expressed as a relative standard deviation (RSD), with respect to the ideal mean C_{std}. Under normal conditions of use the sampler imprecision will be increased by sources of intra-specimen variation such as analytical and flow rate errors, as will be discussed later.

In any experiment to measure the sampler efficiency curves $E_s(D)$ there will be experimental error. Thus the estimate obtained for $\bar{C}_s$ will be $\hat{C}_s = \bar{C}_s + \xi$, where ξ is the evaluation error. Over repeated observations, the evaluation error is assumed to be normally distributed with zero mean, and variance σ^2_{eval}. The variance of the estimated mean concentration $\hat{C}$, estimated using S specimens of the sampler type (and constant sample size per specimen, see Appendices for generalisation), will then be

$$\mathrm{Var}(\hat{C}) = \frac{1}{S}(\sigma^2_{eval} + \sigma^2_s) \tag{6}$$

and the estimated bias $\hat{\Delta}$ will have variance

$$\sigma^2_\Delta \equiv \mathrm{Var}(\hat{\Delta}) = \frac{\mathrm{Var}(\hat{C})}{C^2_{std}}. \tag{7}$$

The estimated sampler imprecision will also have an associated uncertainty, determined by the number of degrees of freedom with which the inter-specimen variance σ^2_s is estimated.

The challenge is to find means to estimate the quantities $\bar{C}$, σ^2_s and σ^2_{eval}, which arise from effects present over a range of particle diameters, from a series of experiments conducted at single diameters using monodisperse aerosols. The methods used in this paper rely on several important assumptions as to the nature of the effects to be quantified. Nevertheless they do provide approximate solutions, which may be used to compare the performances of different sampler types tested using similar experimental designs.

The first method estimates the overall mean concentration $\bar{C}$ by polygonal approximation to the integral in equation (2). The variance of the estimated mean concentration can be expressed in terms of the inter-specimen and residual variances of the measured efficiency values themselves. These in turn are estimated by Analysis of Variance of the efficiency values. The assumptions which underpin this method are that both inter-specimen and residual variances in the efficiency values are the same at all the measured particle aerodynamic diameters. In principle this can be achieved by transformation of the measured variables, although in practice suitable transformations may not be found. Details of the polygonal approximation method are given in Appendix A.

The second method estimates a mean concentration $\bar{C}_s$ for each tested specimen, by fitting a curve to the efficiency values for that specimen using linear or non-linear least-squares regression. The evaluation error is estimated from the uncertainty in the fitted curve's parameters, assuming that the fitted model has no lack-of-fit. The mean concentration estimate and its variance are calculated directly, and the inter-specimen variance estimated from them. The assumptions which underpin this method are that the correct efficiency curves $E_s(D)$ for each specimen can be identified from the limited data available, and the errors arising from lack-of-fit are insignificant in comparison to the quantities being estimated. Details of the curve-fitting method are given in Appendix B.

Experimental design and data transformation

The experiments carried out measure the sampling efficiency of the instrument as a function of particle aerodynamic diameter. Tests are repeated for different values of other important variables such as external wind speed or sampler flow rate, however the data relating to each different experimental condition are usually processed separately. Typical layouts for experiments (taken from CEN, 1993), in which a number of sampler specimens are tested at a number of particle aerodynamic diameters (under the same external conditions), are shown in Table 1. These examples are balanced designs, with either six or 18 specimens tested at nine diameters. There are the same number of repeated tests or replicates

Table 1. Examples of balanced experimental designs

	Replicates with 6 specimens (A–F)						Replicates with 18 specimens (A–R)					
Diameter	1	2	3	4	5	6	1	2	3	4	5	6
1	A	B	C	D	E	F	M	C	K	O	A	E
2	A	B	C	D	E	F	I	P	E	L	O	Q
3	A	B	C	D	E	F	L	K	A	P	I	H
4	A	B	C	D	E	F	R	G	K	M	H	J
5	A	B	C	D	E	F	C	B	J	D	N	O
6	A	B	C	D	E	F	H	D	N	P	R	F
7	A	B	C	D	E	F	F	R	M	G	Q	E
8	A	B	C	D	E	F	I	B	Q	G	C	J
9	A	B	C	D	E	F	F	A	L	N	D	B

(six) at each diameter. The Analysis of Variance (ANOVA) of these experiments is straightforward, and the inter-specimen and residual variances can be obtained using standard algorithms, as discussed in Appendix A.

The six-specimen design in Table 1 can be analysed by either the polygonal approximation or curve-fitting approach, since each sampler specimen is tested at all nine diameters. The 18 specimen design can only be analysed by the polygonal approximation method, since each specimen is tested at only three diameters, which is too few to fit a model by the least-squares method. The diameter values are assumed to cover the range of interest to the sampler under consideration, and will not usually be equally spaced. The order of testing particle diameters would be randomised.

The maximum possible number of degrees of freedom (υ_Δ) for the estimated bias variance is equal to the number of specimens tested minus one (5 or 17). In the limit of no experimental error, the number of degrees of freedom υ_s of the inter-specimen variance estimate will be equal to υ_Δ, however if the residual variance is large in relation to the inter-specimen variance, then υ_s will be considerably reduced. A small value for the inter-specimen degrees of freedom leads to a large uncertainty in the estimated sampler imprecision.

The estimation of sampler bias depends on the choice of a model to describe how the sampling efficiency changes with particle aerodynamic diameter. Likewise, the estimation of sampler precision depends on the choice of a model to describe how the inter-specimen and residual variances change with particle aerodynamic diameter. With monodisperse experiments it is difficult to find the best variance model by directly examining the data, because the number of data points at each diameter is so small. For very minimal experimental designs (like those in Table 1) there are simply not enough data for the diameter-dependence of the inter-specimen and residual variance components to be examined individually. Variations between different sampler specimens may be expressed as shifts in the efficiency curve along the aerodynamic diameter axis, in which case the inter-specimen variance component will vary with aerodynamic diameter. A sampler with an S-shaped efficiency curve (for example a cumulative log-normal curve) will have maximum inter-specimen variance around 50% efficiency, whereas a sampler with a parabolic curve will have maximum inter-specimen variance at low efficiency values. The residual variance, which arises from experimental errors, may show a quite different structure to the inter-specimen variance. The combined variance at each diameter may not therefore reflect the underlying diameter-dependence of either component.

Within the polygonal approximation method, a variance model is incorporated by transforming the efficiency values, before carrying out the ANOVA and estimating the required variances. The transformed variances should be the same at all particle diameters (i.e. homogeneous). It is unlikely that a single transformation would be able to homogenise both inter-specimen and residual variances, so if one of the two variances is clearly dominant, then the transform should homogenise the dominant variance. Since the inter-specimen variance contributes to three of the four quantities which define sampler performance, the choice of transform is of particular importance when the inter-specimen

variance dominates. Examples of commonly used transforms are given in Draper and Smith (1981). The polygonal approximation method is able to accommodate different transforms for the inter-specimen and residual variances, however the quantity of data will generally be too small (for typical experimental designs) to allow individual variance models to be found for both components.

Within the curve-fitting method, the inter-specimen variance model is incorporated implicitly and arises from the choice of the family of efficiency curves that are fitted to the data, as will be illustrated for the PM-10 example. Inhomogeneity of the residual variances can be dealt with separately by transformation of the data prior to curve-fitting, or by using weighted least-squares regression.

Calculation of sampler imprecision and accuracy

Under the conditions of sampler use, the total imprecision of the measured aerosol concentrations will have components arising from diverse sources. The variation which may arise as a result of incorrect setting of the sampler flow rate can be estimated from knowledge of the flow dependence of the sampling efficiency, which can be measured experimentally. Likewise, variation arising from analytical errors can be estimated for typical applications of the sampling method. Examples of how these components may be quantified and combined are given in CEN (1993) and by Bartley *et al.* (1994a).

The inter-specimen variance, which is due to physical differences between sampler specimens, is estimated from the evaluation experiment as $\hat{\sigma}_s^2$. Dividing by the standard C_{std} gives the inter-specimen component of the relative standard deviation:

$$\hat{\mathrm{RSD}}_s = \frac{\hat{\sigma}_s}{C_{std}}. \tag{8}$$

This component is combined with the flow-related and analytical components to estimate the total imprecision RSD_t (for details see the above references).

Several approaches to computing the sampler accuracy, to some required confidence, have been discussed in the literature (Bartley and Fischbach, 1994; Bartley *et al.*, 1994b). All methods require the estimation of the quantities bias, imprecision, and their associated uncertainties. The simplest approach to estimating the accuracy, the Bonferroni method, computes individual upper limits at specified confidence for the sampler bias and imprecision (Fischbach *et al.*, 1993). The single-sided $(1-\alpha_1)\times 100\%$ confidence limit estimate of the absolute bias is given by

$$\Delta_{upper} = |\hat{\Delta}| + t\hat{\sigma}_\Delta, \tag{9}$$

where t is the $(1-2\alpha_1)\times 100$ percentile of Student's-t distribution with υ_Δ degrees of freedom. It is assumed that the uncertainty in the estimated total imprecision $\hat{\mathrm{RSD}}_t$ is due entirely to imperfect knowledge of the inter-specimen component; the upper $(\alpha_2)\times 100\%$ confidence limit estimate for RSD_t is therefore given by

$$\mathrm{RSD}_{upper} = \left[\hat{\mathrm{RSD}}_t^2 + \hat{\mathrm{RSD}}_s^2\left(\frac{\upsilon_s}{\chi^2_{\alpha_2}} - 1\right)\right]^{1/2}, \tag{10}$$

where υ_s is the number of degrees of freedom in the estimate of the inter-specimen variance, and $\chi^2_{\alpha_2}$ is the $(1-\alpha_2)\times 100$ percentile of the chi-square distribution with υ_s degrees of freedom.

3. EXAMPLES OF THE CALCULATION OF PERFORMANCE CHARACTERISTICS

The MRE 113A static respirable aerosol sampler

The MRE 113A static respirable aerosol sampler (Dunmore *et al.*, 1964) incorporates a horizontal elutriator to size-fractionate the aspirated aerosol, and is widely used for

monitoring respirable dust concentrations in British coal mines. The experimental work described here was carried out on behalf of the Health and Safety Executive at the Warren Spring Laboratory (Stevenage, U.K.), and the results of the study have been reported by Mark *et al.* (1993). Note that the results are used in this paper simply to show that the calculation methods work; the performance of the sampler itself will not be discussed.

The aim of the experiments was to measure the sampling efficiency of the MRE 113A as a function of particle size, at three windspeeds. Other more limited tests were carried out to examine the effects of sampler flow rate, pump pulsation and particle type (solid or liquid) on the sampling efficiency. The sampling efficiency was measured by sampling fluorescently tagged monodisperse aerosols with the MRE 113A samplers, and also with isokinetic probes to measure the reference concentration. A total of six specimens (A–F) were available for testing, although only two specimens (chosen at random) were used in each experimental run. This limitation was imposed by the size of the wind tunnel available, and the requirement to keep wind tunnel blockage to below 10%. Each experimental run was replicated three times with the same two specimens. The design matrix for a typical experiment (nine particle sizes, one windspeed) is shown in Table 2.

Sampling efficiency of the MRE 113A

Figure 1 shows the measured sampling efficiencies of the MRE 113A, for solid particles, at one of the three tested windspeeds, and at the sampler's design flow rate of 2.5 l min^{-1}. Reference to Table 2 shows each specimen was tested at only two to four diameters so there

Table 2. Design matrix for MRE 113A evaluation experiment

	Replicates					
Diameters (μm)	1	2	3	4	5	6
1 (1.39)	A	A	A	B	B	B
2 (2.33)	C	C	C	D	D	D
3 (2.81)	B	B	B	F	F	F
4 (3.83)	A	A	A	C	C	C
5 (4.54)	E	E	E	F	F	F
6 (4.99)	D	D	D	E	E	E
7 (5.39)	B	B	B	E	E	E
8 (6.02)	B	B	B	E	E	E
9 (7.63)	A	A	A	C	C	C

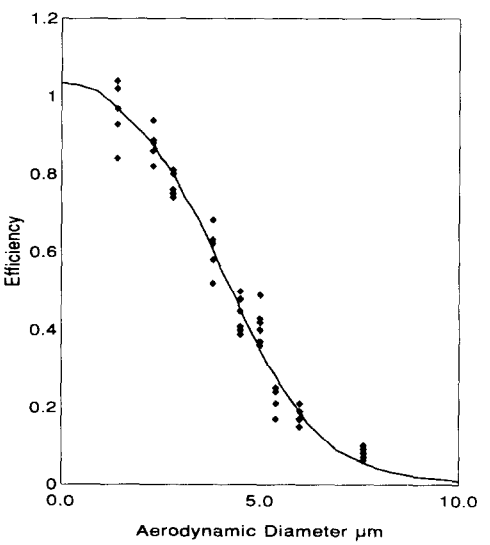


Fig. 1. Measured sampling efficiencies for the MRE 113A sampler; curve fitted to the efficiency values.

are insufficient data to fit a curve to the efficiency values for each specimen, although a curve can be fitted to the pooled data for all specimens. The experimental efficiency values can be modelled by a parabolic function compounded with Gaussian noise, which produces a model of the form

$$\bar{E}(D) = X\left[\Phi\left(\frac{\ln(Y/D)}{Z}\right) - \left(\frac{D}{Y}\right)^2 e^{(2Z^2)}\,\Phi\left(\frac{\ln(Y/D)}{Z} - 2Z\right)\right],$$

where X, Y and Z are empirical parameters and Φ is the normal distribution function. Non-linear least-squares regression was used to find optimal values for the three model parameters, and the resulting fitted curve is plotted in Fig. 1.

Calculation of sampler bias for the MRE 113A

The normalised mean aerosol concentrations that would be sampled from a range of unimodal log-normally distributed aerosols were estimated, both by using the polygonal approximation (equation (A2)) and by numerical integration of the fitted curve (equation (2)). The percentage bias in sampled aerosol concentration, relative to a sampler perfectly following the respirable sampling convention (CEN, 1992) was also estimated (equation (4)). The estimated bias values are plotted in Fig. 2 as a function of aerosol size distribution parameters, in this case for the mean sampling efficiency predicted by the fitted curve. For most of the aerosol size distributions considered, the bias values estimated by polygonal approximation differed by less than $\pm 1\%$ from those shown in the figure. Larger differences (up to $\pm 5\%$) occurred for the size distributions with very low geometric standard deviations and high mass median aerodynamic diameters, where the stepped nature of the polygonal model becomes evident.

Statistical analysis of the MRE 113A data set

A comprehensive statistical analysis of the data was carried out and reported by Chalmers (1992), and is summarised below. Previous measurements of this sampler's performance in calm air indicated that the inter-specimen variance is small (Lidén and Kenny, 1992), and in the results from this experiment, the residual variance was found to be the largest component of the variance of the efficiency values. The variance of the efficiency values increased at large diameters, and a square-root transform was found to improve the homogeneity. This

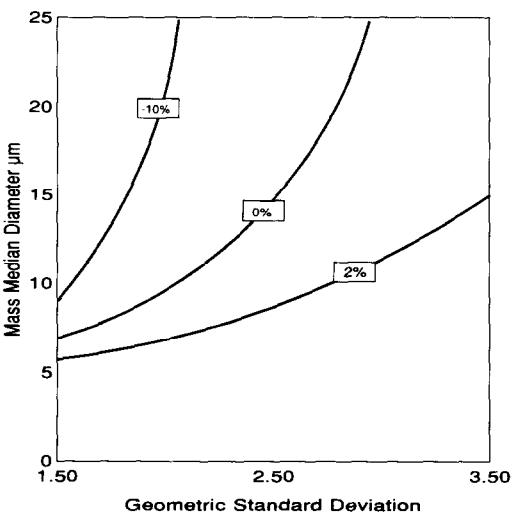


Fig. 2. Bias of the MRE 113A sampler, estimated from the fitted curve, for a range of unimodal log-normal size distributions.

variance model is not in accordance with the expected diameter-dependence of the inter-specimen variance, given the shape of the efficiency curve. Analysis of variance was carried out therefore on both the transformed and the untransformed data from experiments at different particle sizes, windspeeds and flow rates. A series of models was used to separate and estimate the main systematic effects; these were pooled together and the remaining variance attributed to sampler specimens. The inter-specimen variance was estimated from the ANOVA's expected mean squares with approximately 2 degrees of freedom.

Calculation of the uncertainty of the MRE 113A bias

Using equations (A10) and (A11) for transformed variances, and equations (A6) and (A7) for untransformed variances, the inter-specimen and evaluation error components of the variance of the mean concentration were estimated for a range of idealised log-normal aerosol size distributions. The evaluation error was found to be much greater than the inter-specimen variance. The variance of the bias, $\hat{\sigma}_\Delta^2$, was estimated in each case using equation (7); the RSD of the bias is plotted in Fig. 3(a) and (b) as a function of aerosol size distribution parameters. With the untransformed variances the bias is larger, particularly for size distributions having small geometric standard deviation. As $\upsilon_\Delta = 5$, the upper 95% confidence limit for the absolute bias is (from equation (9))

$$\Delta_{\text{upper}} = |\hat{\Delta}| + 2.015\,\hat{\sigma}_\Delta,$$

since the 90th percentile of Student's-t distribution for 5 degrees of freedom is $t = 2.015$. Referring to Figs 2 and 3, one can see that the estimated upper limit of the absolute bias is around 5% for aerosol size distributions in the middle of the range considered. The calculation method makes little difference to this result.

Calculation of MRE 113A imprecision and its uncertainty

The inter-specimen component of the sampler imprecision, $\hat{\text{RSD}}_s$, estimated using equation (A11) for transformed variances and equation (A7) for untransformed variances, is plotted as a function of aerosol size distribution parameters in Fig. 4(a) and (b). The effects

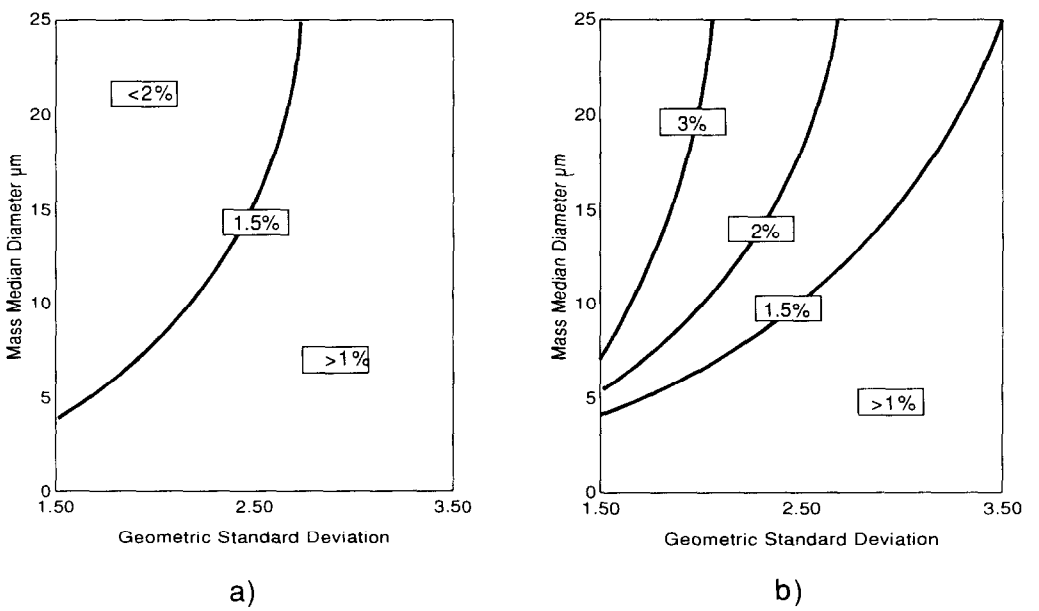


Fig. 3. Relative standard deviation of the MRE 113A bias, estimated using the polygonal approximation method: (a) with square-root transformed data; (b) with untransformed data (i.e. homogeneous variance model).

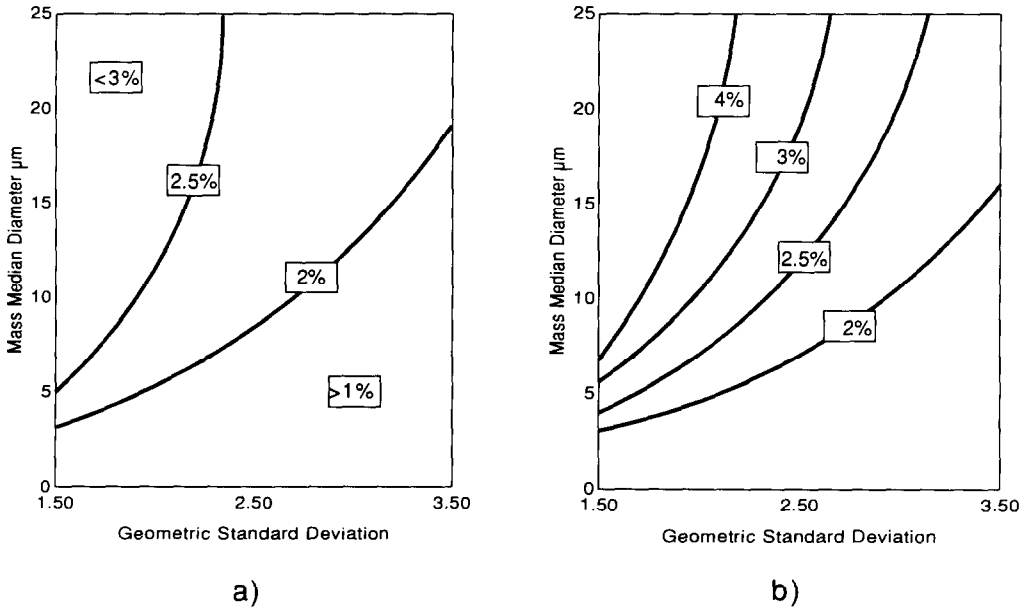


Fig. 4. Inter-specimen relative standard deviation of the MRE 113A sampler, estimated using the polygonal approximation method: (a) with square-root transformed data; (b) with untransformed data (i.e. homogeneous variance model).

arising from the choice of variance model are similar to those noted for the bias RSD. Although the inter-specimen RSD is generally small (around 2.5%), it has a large uncertainty due to the low number of degrees of freedom. The upper 95% confidence limit for the sampler imprecision is (from equation (10)):

$$RSD_{upper} = (R\hat{S}D_t^2 + 18.4\, R\hat{S}D_s^2)^{1/2},$$

since $\upsilon_s \approx 2$, and $\chi^2_{0.95}$ for 2 degrees of freedom is 0.103. The total imprecision, $R\hat{S}D_t$, for the MRE 113A sampler is estimated to be approximately 3%, and is dominated by the component arising from flow variations, as the analytical component for typical applications is small. The upper limit RSD_{upper} will therefore be around 10% in the middle of the range of aerosol size distributions considered, where it does not depend strongly on the variance model adopted. For aerosol size distributions with small geometric standard deviation, the upper confidence limit of the sampler imprecision is around 4% larger for untransformed data than it is for square-root transformed data.

The PM-10 ambient aerosol sampler

The data used in this example are from a number of evaluations of static PM-10 samplers used for monitoring ambient air quality in the United States of America. Experimental results for different specimens of different sampler types tested in different laboratories have been pooled together, in order to construct a data set that can be analysed by both the polygonal approximation and curve-fitting methods. The data are taken from McFarland and Ortiz (1984), Van Osdell and Chen (1989) and Hall *et al.* (1988), who also describe details of the instruments tested. The data are used here to illustrate the calculation methods; a realistic assessment of the performance of the PM-10 sampler cannot be made from this artificially constructed data set and will not be discussed.

Each of the five sampler specimens (A–E) was tested according to the methods laid down by the U.S. Environmental Protection Agency (USEPA, 1987). Tests were carried out at three windspeeds, with a number of monodisperse fluorescently tagged aerosols. Reference

concentrations were measured with isokinetic probes. The example given utilises data for one windspeed only. The design matrix is given in Table 3; data from very slightly different aerodynamic diameters have been pooled together where there was no observable difference in the measured sampling efficiencies. In this example there are different numbers of replicates at each tested diameter.

Sampling efficiency of the PM-10 samplers

Figure 5 shows the measured sampling efficiencies of the PM-10 specimens, at one windspeed only. From Table 3 it can be seen that there are sufficient data to fit a curve to the efficiency values for each sampler specimen. Since the PM-10 convention is a parabola, simple quadratic fits were used, forcing all curves to go through Efficiency = 1 at diameter = 0, and ignoring the data points at around 20 μm. The fitted functions were of the form

$$\hat{E}_s(D) = 1.0 + \hat{\theta}_s D^2.$$

Optimal values for the parameters $\hat{\theta}_s$ were found by weighted least-squares regression, and the resulting curves are plotted in Fig. 5.

Calculation of sampler bias for the PM-10 samplers

The normalised mean aerosol concentrations sampled from a range of ideal log-normally distributed aerosols were estimated, both by using the polygonal approximation (equation

Table 3. Design matrix for PM-10 evaluation experiment

Diameters (μm)	Replicates																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1 (2.8)	A	A	A	A	A	A	A	A	D	D	D	E	E	E			
2 (5.0)	A	A	A	A	B	B	B	C	C	C	C	D	D	D	E	E	E
3 (7.0)	C	C	C	D	D	D	E	E	E								
4 (7.6)	B	B	B														
5 (9.0)	C	C	C	C	D	D	D	E	E	E							
6 (10.1)	A	A	A	A	B	B	B	C	C	C	D	D	D	E	E	E	
7 (11.1)	C	C	C	D	D	D	E	E	E								
8 (11.8)	B	B	B														
9 (13.5)	A	A	A	A	D	D	D	E	E	E							
10 (14.0)	B	B	B	C	C	C	C										
11 (14.8)	D	D	D	E	E	E											
12 (20.0)	B	B	B	B	C	C	C	C									
13 (20.8)	D	D	D	E	E	E											

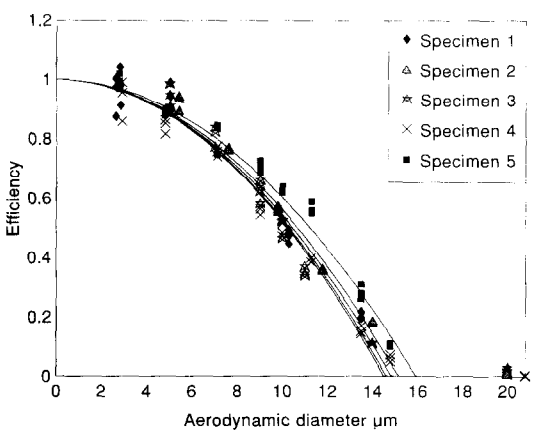


Fig. 5. Measured sampling efficiencies for the PM-10 samplers; curves fitted to the efficiency values for each specimen.

(A2)) and by numerical integration of the fitted curves followed by weighted averaging (equations (1) and (B4)). The percentage bias in sampled aerosol concentration, relative to a sampler perfectly following the PM-10 sampling convention (USEPA, 1987) was also estimated (equation (4)). The estimated bias values are plotted in Fig. 6 as a function of aerosol size distribution parameters, in this case for the mean sampling efficiency predicted by the curve-fitting method. The bias values estimated by polygonal approximation differed by less than $\pm 1\%$ from those shown in the figure, except for the extreme size distributions, providing the data at around 20 μm were ignored in this case also.

Statistical analysis of the PM-10 dataset

The statistical analysis of the PM-10 data is complicated by the imbalance in the "experiment" and by the problems of variance homogenisation. Since there are replicate measurements for each specimen at each diameter, it is possible to estimate directly the inter-specimen and residual components of the variance, as a function of aerodynamic diameter. These estimates are plotted graphically in Fig. 7, which shows that the inter-specimen

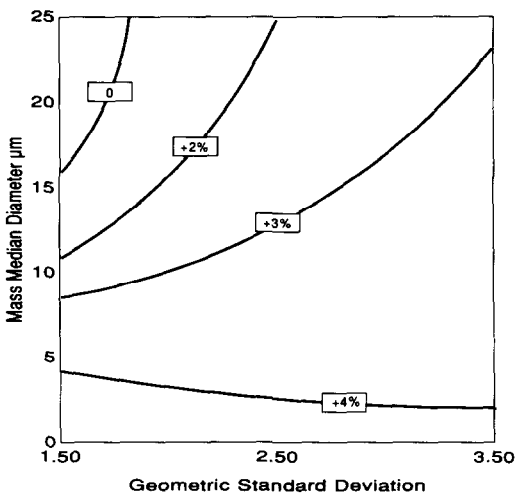


Fig. 6. Bias of the PM-10 sampler, estimated from the fitted curves, for a range of unimodal log-normal size distributions.

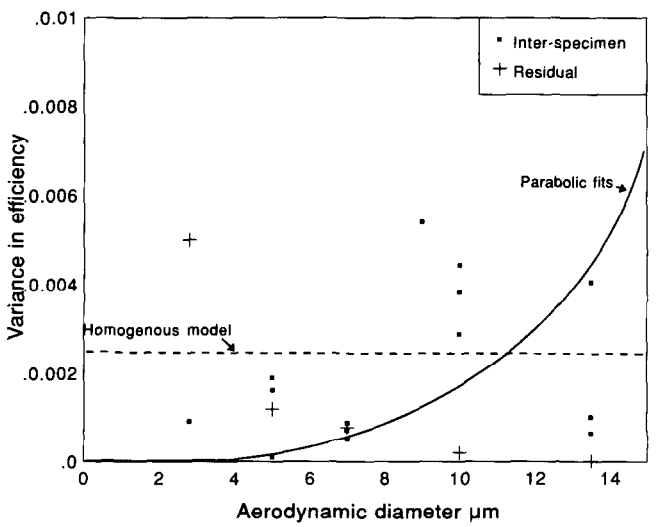


Fig. 7. Inter-specimen and residual components of the variance in PM-10 sampler efficiency, as a function of particle aerodynamic diameter. The two variance models used to calculate sampler imprecision are shown.

variance is generally larger than the residual variance, and both vary with diameter, though in the opposite sense. If the ANOVA is carried out with untransformed data, i.e. the variances are assumed to be the same at all diameters, the inter-specimen variance is estimated as 0.0024. This is shown as a straight line in Fig. 7. This homogeneous model appears to overestimate the inter-specimen variance at low aerodynamic diameters, and to underestimate it at high diameters.

The curve-fitting method incorporates an inter-specimen variance model implicitly, and this model should also be examined. The parabolic curve-fits assume that the sampler type is described in general by the model

$$\bar{E}(D) = 1.0 + \bar{\theta} D^2$$

and so the variance of the efficiency values at any given aerodynamic diameter D is given by

$$\sigma_e^2 = \sigma_\theta^2 D^4 = \sigma_\theta^2 \bar{\theta}^{-2} (1 - \bar{E})^2.$$

This expression is equivalent to the data transformation in equation (A9), i.e.

$$\sigma_\theta^2 \bar{\theta}^{-2} \equiv \sigma_{e_t}^2 \quad \text{and} \quad (1 - \bar{E}) \equiv \{g'(\mu)\}^{-1}.$$

The parameter values θ_s fitted to the data for each PM-10 specimen yield a value of ~ 0.006 for the variance of the transformed efficiency values. An ANOVA could also be used to obtain an estimate of this variance after applying a logarithmic transformation to the data, since

$$(1 - \bar{E}) = \{g'(\mu)\}^{-1} \Rightarrow g(\mu) = \log(1 - \bar{E}), \quad \text{and thus} \quad e^t = \log(1 - e).$$

Figure 6 shows this variance model alongside the variance estimates from the raw data. The model appears to underestimate the inter-specimen variance at low diameters and to overestimate it at high diameters. Since this contrasts with the variances estimated under the assumption of homogeneity, it is useful to compare the performance characteristics obtained using these two different variance models, and to examine the effect that the choice of model can have on the performance evaluation of the sampler.

Calculation of the uncertainty of the PM-10 sampler bias

The inter-specimen, evaluation error and bias variances were estimated using the polygonal approximation method with untransformed data (equations (A6), (A7) and (7)), for a range of log-normally distributed aerosols. Calculations were then repeated using the polygonal approximation method with the transformation derived from the parabolic curve-fits (equations (A10), (A11) and (7)). Finally, the same quantities were estimated using the curve-fitting method (equations (B5), (B6) and (7)).

Estimated bias RSDs, for the curve-fitting method and for the polygonal approximation method with untransformed data, are plotted in Fig. 8(a) and (b). When the polygonal approximation method was used with the data transformation derived from the parabolic curve fits it gave very slightly lower RSD values than those shown in Fig. 8(a). The results for untransformed data are generally larger than those for transformed data, and vary more slowly as a function of aerosol size distribution parameters. The upper 95% confidence limit for the bias is (from equation (9))

$$\Delta_{\text{upper}} = |\hat{\Delta}| + 2.132\, \hat{\sigma}_\Delta,$$

since the 90th percentile of Student's-t distribution for $\upsilon_\Delta \approx 4$ degrees of freedom is $t = 2.132$. Referring back to the estimated bias in Fig. 6, the upper limit of the bias is between 6 and 10% in the middle of the size distribution range considered, depending on which variance model is adopted.

Calculation of PM-10 sampler imprecision and its uncertainty

The inter-specimen component of the sampler imprecision was estimated using the polygonal approximation method with untransformed data (equation (A7)), for a range of

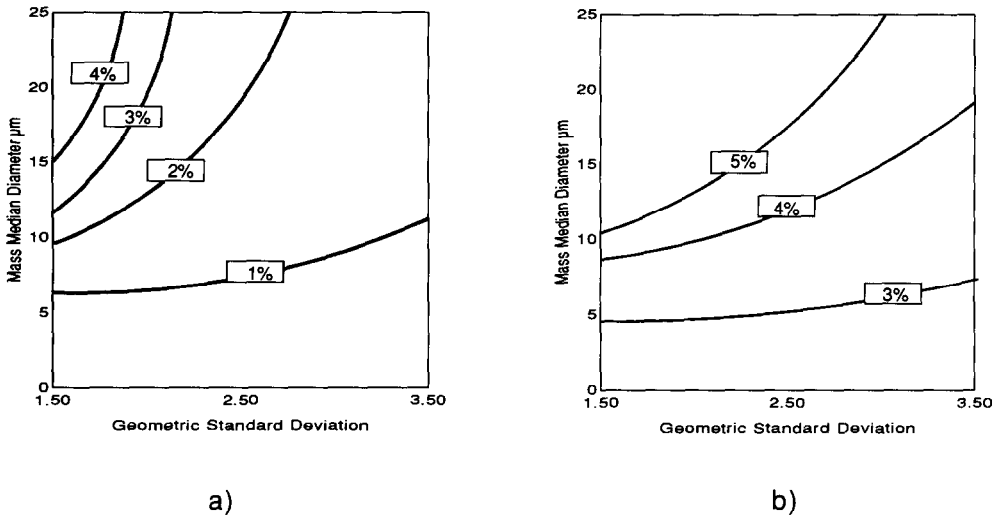


Fig. 8. Relative standard deviation of the PM-10 bias: (a) Estimated by curve-fitting using the parabolic variance model; (b) Estimated by polygonal approximation using the homogeneous variance model.

log-normally distributed aerosols. Calculations were then repeated using the transformation derived from the parabolic curve-fits (equation (A10)), and the curve-fitting method (equation (B5)). The number of degrees of freedom in each case was estimated to be approximately 3.5. Estimated inter-specimen RSDs, for the curve-fitting method and for the polygonal approximation method with untransformed data, are plotted in Fig. 9(a) and (b). The polygonal approximation method with transformed data gave inter-specimen RSDs that were lower by $\sim 0.5\%$ than the curve-fitting results in Fig. 9(a). The inter-specimen RSDs estimated using the untransformed data are larger than those obtained with the parabolic curve fits, and vary more slowly with aerosol size distribution parameters. The upper 95% confidence limit for the sampler imprecision was estimated to be (from equation (10)):

$$\mathrm{RSD}_{\mathrm{upper}} = (\hat{\mathrm{RSD}}_t^2 + 5.6\,\hat{\mathrm{RSD}}_s^2)^{1/2},$$

since $\upsilon_s \approx 3.5$, and $\chi^2_{0.95}$ for 3.5 degrees of freedom ≈ 0.53. The total imprecision of PM-10

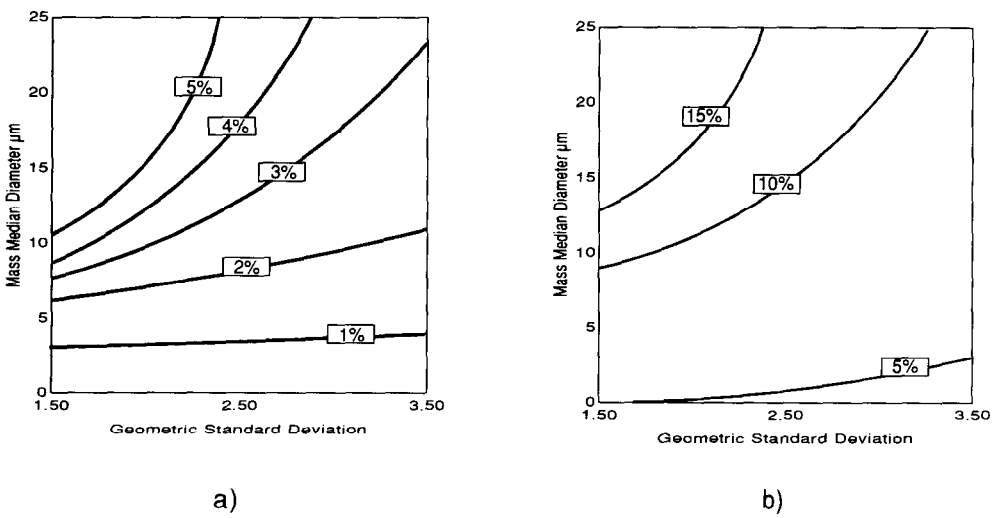


Fig. 9. Inter-specimen relative standard deviation of the PM-10 sampler: (a) Estimated by curve-fitting using the parabolic variance model; (b) Estimated by polygonal approximation using the homogeneous variance model.

samplers, measured in field trials, may not exceed 7% (USEPA, 1987). If this figure is combined with the estimates for the inter-specimen RSD obtained here, the upper limit of the sampler's imprecision is estimated to be between 10 and 24%, for aerosol size distributions in the middle of the range considered.

4. DISCUSSION AND CONCLUSIONS

Two methods have been described for estimating the performance characteristics of aerosol samplers. Each method has some advantages and some drawbacks, and both rely on the assumption of a suitable model to describe the important sources of variation in the experimental data.

The advantages of the polygonal approximation method are that it can be applied to unbalanced experiments, where there are insufficient data to fit a family of models to the experimental results for each specimen. One example is the experiment reported for the MRE 113A sampler, which could not have been analysed using the curve-fitting approach. Assuming that an analysis of variance of the results can be carried out satisfactorily, the polygonal approximation method is computationally simple. However, the weak point is the analysis of variance, especially if there are reasons to apply different transforms to the inter-specimen and residual components of the variation. In the case of the MRE sampler, the inter-specimen variance was small in relation to the residual variance (residual variance was approximately 15 times larger). Applying a transform to the data to homogenise the residual component made little difference to the estimated sampler imprecision.

The advantages of the curve-fitting method are that it is closer to the analysis method that has been successfully applied to polydisperse sampler evaluations, and in principle it avoids confusion between the inter-specimen and residual components of the variance by separating them at the outset. The main weakness of the method is that in order to correctly estimate the sampler precision one needs to identify, from a very restricted data set, the correct family of models for the sampler type. Any lack-of-fit in the curves fitted to each specimen will degrade the estimate of the sampler imprecision; different results could emerge from the same data set through the use of different models to fit the data.

The data from the PM-10 sampler enabled confirmation that the two methods described do in fact give the same results provided that the same model for variation in the data is used in each case. The bias can be estimated easily by either method, although the polygonal approximation method gave poorer estimates for aerosol size distributions with small geometric standard deviation. The PM-10 example showed that the choice of variance model is particularly important when the inter-specimen variation dominates (PM-10 inter-specimen variance was on average three times the residual variance). For this sampler it was possible to show directly that the diameter-dependence of the inter-specimen and residual variance components was quite different; the outcome of the performance evaluation depended on applying a suitable model to the dominant inter-specimen component. The two different models used in the calculations presented appeared to lead in one case to overestimated sampler imprecision, and in the other case to underestimated sampler imprecision.

The results of this work have implications for the design of monodisperse sampler performance evaluation experiments. First, in order to calculate the bias reliably, the test aerosols should include aerodynamic diameters close to the $D=0$ and $E=0$ axes, to minimise the region of extrapolation. The uncertainty in the bias depends mainly on the number of data points at each measured diameter, which should be as large as is practicable.

For the calculation of sampler imprecision, one needs to find a suitable model to describe how the inter-specimen variance, and less importantly the residual variance, depend on aerodynamic diameter. This is extremely difficult for experiments involving very few replicate measurements on a small number of sampler specimens; one really needs a sizeable group of specimens to be tested at each diameter. The residual variance can only be examined independently of the inter-specimen variance if there are replicate measurements, i.e. repeated tests on the same sampler specimen under the same conditions.

The conclusion to be drawn from this work is that very minimal experimental designs like those in Table 1 are too small for reliable estimation of the sampler imprecision; the number of data points should preferably be increased either by repeated measurements on the same group of specimens, or by enlargement of the test group. However, increasing the number of specimens tested is costly, and multiplies the practical difficulties. Current accepted practice for the evaluation of personal inhalable samplers is to mount the samplers on a life-size manikin and test them in a wind tunnel (CEN, 1993). In this case, the number of specimens tested together in one experimental run is limited to the number which can be mounted together, without mutual interference, in the manikin's breathing zone. At present the maximum practical number of specimens is thought to be around six.

The methods described in this paper will be applied, and their results compared, to a study of the performance of personal inhalable samplers at three windspeeds currently being carried out with the support of the European Commission (Kenny, 1993). The experimental design allows the results from the first tested windspeed to be analysed by either of the two methods discussed here. The initial analysis will examine the variance structure of the data, the potential for fitting curves to the response, and estimate the relative sizes of inter-specimen and evaluation errors. The experimental design for completing the tests at the remaining two windspeeds can be altered if necessary, to obtain the best possible result for an experiment of fixed overall size. Given the limitations of both analysis methods, performance comparisons of the tested sampler types will only be possible if the same data analysis method is used for all.

Acknowledgements—This work has been carried out under the Memorandum of Understanding between the U.K. Health and Safety Executive and the U.S. National Institute for Occupational Safety and Health. The authors wish to thank Colin Chalmers of Datastat Consultants for his contribution to the statistical work in this paper. Thanks also to David Mark of Warren Spring Laboratory, Göran Lidén of Arbetsmiljöinstitutet and Jean-François Fabriés of INRS.

Disclaimer—Mention of company names or products does not constitute endorsement by the Centers for Disease Control and Prevention or by the Health and Safety Executive.

REFERENCES

ACGIH (1993) Threshold limit values. *American Conference of Governmental Industrial Hygienists*, ISBN 1-882919-03-8.

Bartley, D. L., Chen, C.-C., Song, R. and Fischbach, T. J. (1994a) Respirable aerosol sampler performance testing. *AIHA J.* (in press).

Bartley, D. L., Fischbach, T. J., Shulman, S. A. and Song, R. (1994b) Generalised confidence limits: application to the estimation of accuracy. *Ann. Statist.* (submitted).

Bartley, D. L. and Fischbach, T. L. (1993) Alternative approaches for analyzing sampling and analytical methods. In *Proc. Int. Symp. on Air Sampling Instrument Performance*; *Appl. occup. environ. Hyg.* **8,** 381–385.

Chalmers, C. P. (1992) Statistical analysis of performance test data for respirable dust samplers from British coal mines. Final report for contract 2563/S42.12, available from the U.K. Health and Safety Executive.

Comité Européen de Normalisation (CEN) (1992) Workplace atmospheres: size fraction definitions for measurement of airborne particles in the workplace. CEN standard EN 481.

Comité Européen de Normalisation (CEN) (1993) Workplace atmospheres: assessment of performance of instruments for measurement of airborne particles. Document CEN/TC137/WG3/N125 or Revision Thereof.

Draper, N. and Smith, H. (1981) *Applied Regression Analysis*, ISBN 0-471-02995-5. Wiley, New York.

Dunmore, J. H., Hamilton, R. J. and Smith, D. S. G. (1964) An instrument for the sampling of respirable dust for subsequent gravimetric assessment. *J. Sci. Instrum.* **41,** 669–672.

Fischbach, T., Shulman, S. and Song, R. (1994) Some statistical procedures for sampling and analytical method accuracy tests and estimation. *Cincinnati: National Institute for Occupational Safety and Health* (in preparation).

Gunderson, E. C. and Anderson, C. C. (1980) Development and validation of methods for sampling and analysis of workplace toxic substances. *Cincinnati: National Institute for Occupational Safety and Health*, Pub no. 80–133.

Hall, D. J., Upton, S. L., Marshall, G. W., Waters, R. A. (1988) Wind tunnel measurements of the collection efficiency of two PM-10 samplers: the Sierra–Andersen model 321A hi-volume sampler and the EPA prototype dichotomous sampler. Warren Spring Laboratory Report no. LR 669 (PA)M, ISBN 0-85624-529-1.

International Standards Organisation (ISO) (1992) Air quality—particle size fraction definitions for health-related sampling. *ISO CD7708*, International Standards Organisation, Geneva.

Kenny, L. C. (1993) Report of progress towards the development of a performance standard for aerosol sampling instruments used in occupational hygiene. *J. Aerosol. Sci.* **23,** 773–779.

Lidén, G. and Kenny, L. C. (1992) The performance of respirable dust samplers: bias, precision and inaccuracy. *Ann. occup. Hyg.* **36,** 1–22.

Mark, D., Lyons, C. P. and Upton, S. L. (1993) Performance testing of the respirable dust sampler used in British coal mines. In *Proc. Int. Symp. on Air Sampling Instrument Performance. Appl. occup. environ. Hyg.* **8,** 370–380.
Mark, D. and Vincent, J. H. (1986) A new personal dust sampler for airborne total dust in workplaces. *Ann. occup. Hyg.* **30,** 89–102.
McFarland, A. R. and Ortiz, A. O. (1983) Characterisation of Sierra–Andersen Model 321 size-selective inlet for hi-vol samplers. Air Quality Laboratory Report no. 4716/04/02/84/ARM. Air Quality Laboratory, Texas A&M University, College Station, TX 77843.
Satterthwaite, F. E. (1946) An approximate distribution of estimates of variance components. *Biometrics* **2,** 110–114.
USEPA (1987) Ambient air monitoring reference and equivalent methods. United States Environmental Protection Agency, Federal Register 40CFR Part 53. 1st July 1987.
VanOsdell, D. W. and Chen, F. L. (1989) Wind tunnel test Report no. 27: PM-10 inlet testing at 8, 24 and 48 km/h in the modified wind tunnel. RTI project no. 432U-3999-93, Research Triangle Institute, PO Box 12194, Research Triangle Park, NC 27709-2194.

APPENDIX A: POLYGONAL APPROXIMATION METHOD

Estimation of mean sampled concentration

The raw data (efficiency values) are e_{ij}, where i denotes a particle aerodynamic diameter ($i = 1$ to I), j denotes a replicate result at a given diameter ($j = 1$ to J_i). The mean sampling efficiency at each measured diameter D_i is estimated as

$$\hat{E}_i = \frac{1}{J_i}\sum_j e_{ij}. \tag{A1}$$

This is a point estimate of the sampling efficiency curve $\bar{E}(D)$ of the sampler at particle aerodynamic diameter $D = D_i$. The polygonal approximation to estimate the mean normalised mass concentration sampled from an arbitrary aerosol $A(D)$ (see equation (2)) is given in this example by the trapezoidal rule, although a similar approach could be used with other integration schemes:

$$\hat{C} = \sum_{i=1}^{I} W_i \hat{E}_i, \tag{A2}$$

where

$$W_i = \frac{A_i + A_{(i+1)}}{2}, \quad i \neq 1; \quad W_1 = A_1 + \tfrac{1}{2}A_2, \qquad A_i = \int_{D_{(i-1)}}^{D_i} A(D)\,\mathrm{d}D.$$

The calculation of the first and last weighting factors, W_1 and W_I, require extrapolation from the data to give a value for E at $D = 0$, and a value for D at $E = 0$. In the example here it is assumed that E at $D = 0$ is equal to E at D_1; the value of D at which $E = 0$ is found by linear extension of the two highest diameter data points. Note that the upper extrapolation would not be necessary when analysing inhalable sampler data.

Statistical model

The efficiency values e_{ij} are analysed according to the model

$$e_{ij} = \mu_i + \varphi_{ij} + \varepsilon_{ij}, \tag{A3}$$

where μ_i is the true mean efficiency of the sampler type at particle diameter $D = D_i$, φ_{ij} is the systematic deviation at each diameter from the true mean efficiency which is associated with inter-specimen variability and ε_{ij} is the random experimental error.

An indicator function δ_{ijs} may be defined which describes the experimental design, i.e. $\delta_{ijs} = 1$ if observation ij was made with sampler specimen s, and is zero otherwise. Thus

$$\varphi_{ij} = \sum_{s=1}^{S} \delta_{ijs}\varphi_s. \tag{A4}$$

So each sampler specimen is assumed to have an individual efficiency curve that is displaced from the average efficiency curve by a fixed amount φ_s on the efficiency axis. In principal this assumption can be met by a suitable transformation of the efficiency values. Across the entire test group of S specimens, the systematic effects φ_s are assumed to be normally distributed, with zero mean and variance equal to σ_φ^2. The experimental errors are also assumed to be random, normally distributed with mean equal to zero and variance σ_ε^2. The polygonal approximation to the estimated mean concentration (equation (2)) may now be written as

$$\begin{aligned}\hat{C} &= \sum_{i=1}^{I}\sum_{j=1}^{J_i}\frac{W_i}{J_i}(\mu_i + \varphi_{ij} + \varepsilon_{ij}) \\ &= \sum_{i=1}^{I} W_i\mu_i + \sum_{i=1}^{I}\sum_{j=1}^{J_i}\frac{W_i}{J_i}\sum_{s=1}^{S}\delta_{ijs}\varphi_s + \sum_{i=1}^{I}\sum_{j=1}^{J_i}\frac{W_i}{J_i}\varepsilon_{ij}.\end{aligned} \tag{A5}$$

The variance of the mean concentration, $\mathrm{Var}(\hat{C})$, has two components, arising from the residual errors and from the

Table A1. Analysis of variance of experimental designs in Table 1

Source of variance	Expected mean squares (EMS)	Degrees of freedom (υ)
Diameters	$k_1\sigma_\varphi^2 + 6\sigma_{\text{diam.}}^2 + \sigma_\varepsilon^2$	$I-1$
Specimens	$k_2\sigma_\varphi^2 + \sigma_\varepsilon^2$	$S-1$
Residual	σ_ε^2	$53-(I-1)-(S-1)$
Total		53

inter-specimen variability. The residual error component is used to estimate the evaluation error in the mean concentration:

$$\frac{\hat{\sigma}_{\text{eval}}^2}{S} \equiv \hat{\sigma}_\varepsilon^2 \sum_{i=1}^{I} \frac{W_i^2}{J_i}, \tag{A6}$$

where $\hat{\sigma}_\varepsilon^2$ is the estimated residual variance of the efficiency values. The inter-specimen component is used to estimate the inter-specimen variance in the mean concentration:

$$\frac{\hat{\sigma}_s^2}{S} \equiv \hat{\sigma}_\varphi^2 \sum_{s=1}^{S} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{Ji} \delta_{ijs} W_i / J_i \right\}^2, \tag{A7}$$

where $\hat{\sigma}_\varphi^2$ is the estimated inter-specimen variance of the efficiency values.

Analysis of variance of the sampling efficiency values

Estimates of the variances σ_φ^2 and σ_ε^2 are obtained by analysis of variance (ANOVA) of the efficiency values. An example of an ANOVA table, for the experimental designs given in Table 1, is shown in Table A1. In the case of the experiment with $S=6$ specimens, the constants in Table A1 take the values $k_1=0$ and $k_2=9$. In the case of the experiment with $S=18$ specimens, $k_1=3/4$ and $k_2=45/17$ (CEN, 1993). The variances σ_φ^2 and σ_ε^2 can be estimated from the expected mean squares and degrees of freedom for specimens and residuals:

$$\sigma_\varphi^2 = (\text{EMS[specimens]} - \text{EMS[residual]})/B; \quad \sigma_\varepsilon^2 = \text{EMS[residuals]} \text{ and } \upsilon_\varepsilon = 53-(I-1)-(S-1).$$

Since the inter-specimen variance is estimated as a linear function of two expected mean squares U = EMS [specimens] and V = EMS[residual], the approximate degrees of freedom for the inter-specimen variance (i.e. the sum $aU+bV$, where a and b are constants) can be estimated using Satterthwaite's approximation (Satterthwaite, 1946):

$$\upsilon = \frac{(aU+bV)^2}{\left(\dfrac{a^2U^2}{m} + \dfrac{b^2V^2}{n}\right)}. \tag{A8}$$

Here m is the degrees of freedom of U and n is the degrees of freedom of V.

If the efficiency values are transformed in order to homogenise the variances, then the ANOVA is carried out on the transformed values $e^t = g(e)$, where g is the transforming function. The estimated variances σ_φ^2 and σ_ε^2 will be in terms of e^t. To calculate these variances in terms of efficiency e, the variances are back-transformed, using the approximation

$$\sigma^2(e^t) = \{g'(\mu)\}^2 \sigma^2(e), \tag{A9}$$

where μ and $\sigma^2(e)$ are the mean and variance of e, $\sigma^2(e^t)$ is the variance of the transformed variable estimated by the ANOVA, and $g'(\mu)$ is the derivative of g with respect to e, evaluated at $e=\mu$. For transformed variables, equation (A6) for the variance in mean concentration arising from evaluation error becomes:

$$\frac{\hat{\sigma}_{\text{eval}}^2}{S} = \hat{\sigma}_{\varepsilon^t}^2 \sum_{i=1}^{I} \frac{W_i^2}{J_i \{g'(\hat{E}_i)\}^2}, \tag{A10}$$

where $\hat{\sigma}_{\varepsilon^t}^2$ is the estimated residual variance of the transformed efficiency values. Equation (A7) for the component arising from inter-specimen imprecision becomes

$$\frac{\hat{\sigma}_s^2}{S} = \hat{\sigma}_{\varphi^t}^2 \sum_{s=1}^{S} \left(\sum_i \sum_j \frac{W_i \delta_{ijs}}{\{g'(\hat{E}_i)\} J_i} \right)^2, \tag{A11}$$

where $\hat{\sigma}_{\varphi^t}^2$ is the estimated inter-specimen variance of the transformed efficiency values.

APPENDIX B: CURVE-FITTING METHOD

Suppose that a linear combination of a small number P of functions $E_p(D)$ can describe the sampling efficiency adequately. The combinations of functions $E_p(D)$ may be obtained, for example, by non-linear least-squares regression. For each sampler specimen s, regression of that specimen's efficiency values e_{ij} determines parameter

estimates $\hat{\theta}_{ps}$:

$$e_{ij} = \sum_{p=1}^{P} \hat{\theta}_{ps} E_p(D_i) + \varepsilon_{ij}. \tag{B1}$$

The random experimental errors ε_{ij} are assumed to be normally distributed with zero mean, and to be homogeneous in diameter. If they are not, the regression may be carried out by weighted least squares, or the efficiency values suitably transformed. The efficiency curve for sampler specimen s is then estimated as

$$\hat{E}_s(D) = \sum_{p=1}^{P} \hat{\theta}_{ps} E_p(D), \tag{B2}$$

which, by substitution in equation (1), is used to estimate the mean concentration for specimen s, $\bar{C}_s$. The variance at fixed s of the estimate $\hat{C}_s$ may be estimated from the variance–covariance matrix of the parameter estimates used to fit the curves $\hat{E}_s(D)$:

$$\mathrm{Var}(\hat{C}_s) = \sum_p \sum_q \frac{\partial \hat{C}_s}{\partial \hat{\theta}_p} \mathrm{cov}_{pqs} \frac{\partial \hat{C}_s}{\partial \hat{\theta}_q} \tag{B3}$$

where cov_{pqs} are the variance–covariance matrix elements for the estimated model parameters $\hat{\theta}$ for sampler specimen s. The estimated mean concentration for the sampler type is the average of the concentrations for the individual specimens, weighted according to the number of efficiency values for each specimen:

$$\hat{C} = \sum_s \frac{N_s}{N} \hat{C}_s \tag{B4}$$

where N_s is the total number of efficiency values available for specimen s, and $N = \Sigma N_s$. The variance in the estimated mean concentration can be estimated directly, again weighting for the number of efficiency measurements for each specimen:

$$\begin{aligned} \mathrm{Var}(\hat{C}) &= \frac{\hat{\sigma}^2_{\mathrm{eval}}}{S} + \hat{\sigma}^2_s \sum_s \left(\frac{N_s}{N}\right)^2 \\ &\equiv \frac{1}{S}(\hat{\sigma}^2_{\mathrm{eval}} + \hat{\sigma}^2_s) \quad \text{when } N_s = \frac{N}{S} \end{aligned} \tag{B5}$$

and

$$\left[1 - \sum_s \left(\frac{N_s}{N}\right)^2\right] \hat{\sigma}^2_s = \sum_s \frac{N_s}{N} (\hat{C}_s - \hat{C})^2 - \frac{(S-1)}{S} \hat{\sigma}^2_{\mathrm{eval}}. \tag{B6}$$

The evaluation error component of the variance in the mean concentration, σ^2_{eval}, is estimated from the pooled variances for the individual sampler specimens:

$$\hat{\sigma}^2_{\mathrm{eval}} \equiv \frac{S}{N - PS} \sum_{s=1}^{S} \frac{N_s}{N} (N_s - P) \mathrm{Var}(\hat{C}_s) \tag{B7}$$

where $\mathrm{Var}(\hat{C}_s)$ is given by equation (B3). The number of degrees of freedom in the estimate of σ^2_{eval} is approximately $(N - PS)$. Note that σ^2_{eval} refers to the experimental error in the determination of the mean response of each specimen, not the error in each replicate determination. The inter-specimen variance, σ^2_s, can now be estimated by subtracting the estimated evaluation error component (equation (B6)) from $\mathrm{Var}(\hat{C})$ (equation (B5)). The number of degrees of freedom on the estimate of the inter-specimen variance can be estimated using Satterthwaite's approximation (equation (A8)), given that the number of degrees of freedom for the sum $(\sigma^2_{\mathrm{eval}} + \sigma^2_s)$ is approximately $(S - 1)$.