United States
Environmental Protection
Agency

Industrial Environmental Research
Laboratory
Research Triangle Park NC 27711

EPA-600/7-78-132a
July 1978

EPA

A Data Reduction System for Cascade Impactors

Interagency Energy/Environment R&D Program Report

RESEARCH REPORTING SERIES

Research reports of the Office of Research and Development, U.S. Environmental Protection Agency, have been grouped into nine series. These nine broad categories were established to facilitate further development and application of environmental technology. Elimination of traditional grouping was consciously planned to foster technology transfer and a maximum interface in related fields. The nine series are:

1. Environmental Health Effects Research
2. Environmental Protection Technology
3. Ecological Research
4. Environmental Monitoring
5. Socioeconomic Environmental Studies
6. Scientific and Technical Assessment Reports (STAR)
7. Interagency Energy-Environment Research and Development
8. "Special" Reports
9. Miscellaneous Reports

This report has been assigned to the INTERAGENCY ENERGY-ENVIRONMENT RESEARCH AND DEVELOPMENT series. Reports in this series result from the effort funded under the 17-agency Federal Energy/Environment Research and Development Program. These studies relate to EPA's mission to protect the public health and welfare from adverse effects of pollutants associated with energy systems. The goal of the Program is to assure the rapid development of domestic energy supplies in an environmentally-compatible manner by providing the necessary environmental data and control technology. Investigations include analyses of the transport of energy-related pollutants and their health and ecological effects; assessments of, and development of, control technologies for energy systems; and integrated assessments of a wide range of energy-related environmental issues.

EPA REVIEW NOTICE

EPA-600/7-78-132a
July 1978

A Data Reduction System for Cascade Impactors

by

J.D. McCain, G.I. Clinard, L.G. Felix, and J.W. Johnson

Southern Research Institute
2000 Ninth Avenue, South
Birmingham, Alabama 35205

Contract No. 68-02-2131
T.D. 10101

Program Element No. EHE624

EPA Project Officer: D. Bruce Harris

Industrial Environmental Research Laboratory
Office of Energy, Minerals, and Industry
Research Triangle Park, NC 27711

Prepared for

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office of Research and Development
Washington, DC 20460

ABSTRACT

A computer based data reduction system for cascade impactors has been developed.[1] The system utilizes impactor specific calibration information together with operating conditions and other pertinent information such as stage weights, sampling duration, etc., to determine particle size distributions in several forms for individual runs. A spline technique is applied to fit a curve to the cumulative size distribution obtained from each individual impactor run. These fitted curves have forced continuity in coordinates and slopes. Averages of size distributions for multiple runs are made using the fitted curves to provide interpolation values at a consistent set of particle diameters, irrespective of the diameters at which the data points fall in the original individual run data sets. Statistical analyses are performed to locate and remove outliers from the data being averaged, following which averages, variances, standard deviations and confidence intervals are calculated. The averages and statistical information are available in tabular and graphical form in several size distribution formats (cumulative mass loading, cumulative percentage by mass, differential mass, differential number). The averaged data are stored in disk files for subsequent manipulation. Additional programs permit data sets from control device inlet and outlet measurements to be combined to determine fractional collection efficiencies and confidence limits of the calculated efficiencies.

These results are available in graphical form with a choice of log-probability or log-log presentations and as tabular output. The program is set up to handle all commercially available round jet cascade impactors, including common modifications, which are in current use in stack sampling. Other round jet impactors can be easily substituted and slot type impactors could be accommodated with slight program revision.

This report was submitted in partial fulfillment of Contract No. 68-02-2131, Technical Directive No. 10101, by Southern Research Institute under the sponsorship of the U.S. Environmental Protection Agency. This report covers the period March 1, 1978, to May 1, 1978, and work was completed as of May 12, 1978.

CONTENTS

FIGURES

FIGURES (Cont'd)

TABLES

ACKNOWLEDGEMENTS

This work was conducted under U.S. Environmental Protection Agency contract number 68-02-2131. The continued support and assistance of Mr. D. B. Harris, the EPA project officer on this contract is gratefully acknowledged. Additional appreciation is also expressed to Mr. B. Gaston and Ms. A. Henry for their assistance with the programs; to Mr. K. Cushing whose earlier program was used as a basis for program MPPROG; and Dr. W. B. Smith for his assistance and guidance during the program development.

SECTION 1

INTRODUCTION

Cascade impactors have gained wide acceptance as a practical means of making particle size distribution measurements. These devices are regularly used in a wide variety of environments, ranging from ambient conditions to flue gas streams at 500°C (950°F). Specially fabricated impactors can be used for more extreme conditions.

Because of their usefulness, the U.S. Environmental Protection Agency has funded research which has explored the theoretical and practical aspects of impactor operation. As part of this research, an effort has been made to design a comprehensive data reduction system which will make full use of cascade impactor measurements.

The cascade impactor data reduction system (CIDRS) described here is designed to automatically reduce data taken with any one of four commercially available round jet cascade impactors: the Andersen Mark III Stack Sampler, the Brink Model BMS-11 (as supplied and with extra stages), the University of Washington Mark III Source Test Cascade Impactor, and the Meteorology Research Incorporated Model 1502 Inertial Cascade Impactor. Provision is not made in this system for reducing data taken with slotted jet impactors. With modification the computer programs can accommodate any round jet impactor with an arbitrary number of stages, and with more extensive revision, can be made to handle data from slotted jet impactors.

The computer programs which comprise this data reduction system are written in the FORTRAN IV language. The plotting subroutines used were written specifically for the Digital Equipment Corporation (DEC) PDP-15/76 computer, and these programs are not compatible with other plotting systems. However, these programs can be used as a guide when revision is made for use with another operating system.

A broad outline of the program fundamentals is given here with sufficient detail for anyone without a specialized knowledge of computers to understand the methods and rationale of the program. The program comprises two major blocks. The first block treats data from individual impactor runs while the second treats data from groups of runs, providing averages, statistical information and fractional penetration (efficiency) results. The six mainline programs which make up the data reduction system are described in the overall program flow shown in Table 1.

TABLE 1. PROGRAM FLOW

BLOCK 1. SINGLE RUN ANALYSIS

I. Impactor Program (MPPROG)

Takes testing conditions and stage weights to produce stage D_{50}'s, cumulative and cumulative % mass concentrations <D_{50}, geometric mean diameters, and mass number size distributions. Executed for each run.

II. Fitting Program (SPLIN1)

Uses modified spline technique to fit cumulative mass loading points for each plot. Stores fitting coefficients and boundary points on file. Executed for each run.

III. Graphing Program (GRAPH)

Produces individual run graphs with points based on stage weights and impactor D_{50}'s. Also superimposes plot based on fitted data, if desired. Graphs include cumulative mass loading, cumulative % mass loading, and mass and number size distributions. Can be executed as desired for each run.

BLOCK 2. GROUPED DATA

IV. Statistical Program (STATIS)

Recalls cumulative mass loading fitting coefficients to produce average cumulative mass loading, average % cumulative mass loading, average mass size distribution, and average number size distribution plots each with 50% or 90% confidence bars. Executed for each group or data to be averaged.

Programs I-IV are used for both inlet and outlet data sets.

V. Efficiency Program (PENTRA) or (PENLOG)

Recalls average mass size distribution values along with 50% confidence limits for inlet and outlet to plot percent penetration and efficiency with 50% confidence bars. Executed once for each pair or groups and used to define a fractional efficiency curve.

SECTION 2

SUMMARY AND CONCLUSION

CIDRS represents a powerful, versatile tool for reducing and managing data obtained with cascade impactors. It provides the capability for single and multiple run data analysis with varying degrees of smoothing of single run data available and averaging and statistical analysis of multiple runs. Results from the program are not biased by forced fits to arbitrary distribution forms. Finally, the program makes possible a very significant time saving in handling and processing field data obtained with cascade impactors.

SECTION 3

INDIVIDUAL RUN DATA ANALYSIS

This portion of the impactor data reduction package utilizes impactor hardware information, particulate catch information, and sampling conditions from single impactor runs to calculate size distributions. The overall distributions are available in several forms. The run analysis and output presentation are accomplished by three main programs, MPPROG, SPLIN1, and GRAPH. MPPROG and SPLIN1 perform analysis and manipulation while GRAPH is totally devoted to various forms of graphical presentation of the calculated distributions. The routines used in GRAPH are specifically for use on a PDP-15/76 computer and are not compatible with most other computers without modification. However, the general structure of GRAPH should serve as a useful base for programming to achieve similar graphical output from other computing systems.

PROGRAM MPPROG

In MPPROG, sampling hardware information, sampling conditions and particulate catch information are used to determine the effective cut sizes of the various impactor stages and the concentrations of particles caught on these stages. The output is organized into several tabular forms and stored on a disk file for later use.

Input Data to MPPROG

Because individual impactors, even of the same type, do not necessarily have precisely the same operational characteristics, the program calculates stage cut diameters on an impactor specific basis. Hardware data are stored within the program which include, for each impactor to be used, the number of stages, the number of jets per stage, the jet diameters, the stage calibration constants, and flow-pressure drop relations for each stage.

Run specific input data to MPPROG are listed in Table 2. The maximum particle diameter must be measured by microscopic examination of the particles collected on the first stage. Gas analysis must be made at the same time the impactor is run.

TABLE 2. INPUT DATA TO MPPROG

1. Impactor identification (required to call up hardware information)
2. Gas composition (CO_2, CO, N_2, O_2, H_2O)
3. Impactor flow rate (ACFM at stack conditions)
4. Stack pressure
5. Stack temperature
6. Gas temperature within impactor
7. Duration of sampling
8. True density of particles
9. Maximum particle diameter present in sample
10. Masses of catches by stage

Output from MPPROG

Both input information pertaining to individual impactor runs and calculated results are listed on the line printer. A sample of output from MPPROG generated to the line printer is shown in Table 3.

Input information for each run is given in the first five lines of the printout. (Impactor pressure drop in inches of mercury is calculated within the program; gas composition is input as dry fractions and output as wet gas composition by percent.) The mass collected on each stage in milligrams is listed in the tenth line of the Table.

The mass loading is calculated from the total mass of particles collected by the impactor and the total gas volume sampled, and it is listed in four different units after the heading CALC. MASS LOADING. The units are defined as:

GR/ACF - grains per actual cubic foot of gas at stack conditions of temperature, pressure, and water content.

GR/DSCF - grains per dry standard cubic foot of gas at engineering standard conditions of gas. Engineering dry standard conditions in the English system are defined as 0% water content, 70°F, and 29.92 inches of Hg.

MG/ACM - milligrams per actual cubic meter of gas at stack conditions of temperature, pressure, and water content.

MG/DNCM - milligrams per dry normal cubic meter of gas at engineering normal conditons of the gas. Engineering dry normal conditions in the metric system are defined as 0% water content, 21°C and 760 mm of Hg (Torr).

Below the run condition data summary, the information pertinent to each stage is summarized in columnar form in order of decreasing particle size from left to right. Thus S1 is the first stage, S8 is the last stage, and FILTER is the back-up filter. If a precollector cyclone was used, a column labeled CYC would appear to the left of the S1 column and information relevant to the cyclone would be listed in this column. Beneath each impactor stage number is listed the corresponding stage index number, which also serves as identification for the stage.

Directly beneath these listings are the effective stage cut diameters. The effective stage cut diameter is assumed to be equal to the particle diameter for which the stage collection efficiency is 50%. This diameter, D_{50}, is calculated from an equation of the form

$$D_{50} = k_s \left[\frac{\mu d}{\rho_p c v}\right]^{\frac{1}{2}} \tag{1}$$

TABLE 3. SAMPLE CALCULATIONS

INPUT INFO

HYPOTHETICAL ANDERSEN

IMPACTOR FLOWRATE = 0,500 ACFM — IMPACTOR TEMPERATURE = 400,0 F = 204,4 C — SAMPLING DURATION = 20,00 MIN

IMPACTOR PRESSURE DROP = 0,3 IN, OF HG — STACK TEMPERATURE = 400,0 F = 204,4 C

ASSUMED PARTICLE DENSITY = 1,35 GM/CU,CM, — STACK PRESSURE = 26,50 IN, OF HG — MAX, PARTICLE DIAMETER = 100,0 MICROMETERS

GAS COMPOSITION (PERCENT) — CO2 = 1,94 — CO = 0,00 — N2 = 76,53 — O2 = 20,53 — H2O = 1,00

CALC, MASS LOADING = 8,0711E-03 GR/ACF — 1,4748E-02 GR/DNCF — 1,8470E+01 MG/ACM — 3,3748E+01 MG/DNCM

		S1	S2	S3	S4	S5	S6	S7	S8	FILTER
	IMPACTOR STAGE	S1	S2	S3	S4	S5	S6	S7	S8	FILTER
	STAGE INDEX NUMBER	1	2	3	4	5	6	7	8	9
	D50 (MICROMETERS)	10,72	9,93	6,35	4,18	2,21	1,28	0,67	0,33	
INPUT INFO	MASS (MILLIGRAMS)	0,72	0,40	0,53	0,09	0,38	1,43	1,25	0,04	0,39
	MG/DNCM/STAGE	4,71E+00	2,62E+00	3,47E+00	5,89E-01	2,49E+00	9,35E+00	8,18E+00	2,62E-01	2,55E+00
CUMULATIVE	CUM, PERCENT OF MASS SMALLER THAN D50	86,23	78,59	68,45	66,73	59,46	32,12	8,22	7,46	
CUMULATIVE	CUM, (MG/ACM) SMALLER THAN D50	1,59E+01	1,45E+01	1,26E+01	1,23E+01	1,10E+01	5,93E+00	1,52E+00	1,38E+00	
CUMULATIVE	CUM, (MG/DNCM) SMALLER THAN D50	2,91E+01	2,65E+01	2,31E+01	2,25E+01	2,01E+01	1,08E+01	2,77E+00	2,52E+00	
CUMULATIVE	CUM, (GR/ACF) SMALLER THAN D50	6,96E-03	6,34E-03	5,52E-03	5,39E-03	4,80E-03	2,59E-03	6,64E-04	6,02E-04	
CUMULATIVE	CUM, (GR/DNCF) SMALLER THAN D50	1,27E-02	1,16E-02	1,01E-02	9,84E-03	8,77E-03	4,74E-03	1,21E-03	1,10E-03	
DIFF	GEO, MEAN DIA, (MICROMETERS)	3,27E+01	1,03E+01	7,94E+00	5,15E+00	3,04E+00	1,68E+00	9,30E-01	4,75E-01	2,36E-01
DIFF	DM/DLOGD (MG/DNCM)	4,86E+00	7,93E+01	1,78E+01	3,24E+00	8,96E+00	3,95E+01	2,94E+01	8,56E-01	8,47E+00
DIFF	DN/DLOGD (NO, PARTICLES/DNCM)	1,96E+05	1,02E+08	5,03E+07	3,35E+07	4,52E+08	1,18E+10	5,18E+10	1,13E+10	9,12E+11

NORMAL (ENGINEERING STANDARD) CONDITIONS ARE 21 DEG C AND 760MM HG,

where D_{50} = effective cut size,
k_S = stage calibration constant,
μ = gas viscosity,
d = jet diameter,
ρ_p = particle density,
C = Cunningham slip correction factor, and
v = jet velocity.

Because the particle diameter, D_{50}, enters the equation for C, the solution of Equation 1 is done by an iterative process. If the particle density, ρ_p, is set equal to the true density of the particles, as is the case in the sample output in Table 3, the resulting diameter calculated from Equation 1 is the Stokes diameter, D_S. If ρ_p is set equal to 1.0 the resulting diameter is the aerodynamic diameter D_A as defined by the Task Group on Lung Dynamics.[2] If both ρ_p and C are set equal to 1.0, the resulting diameter is the aerodynamic impaction diameter, D_{AI}, as defined by Mercer.[3] Unless otherwise specified, MPPROG will automatically provide parallel output in terms of D_S and D_A. Parallel results in terms of D_S and D_{AI} or in terms of D_A and D_{AI} are available if called for.

The stage weights, in units of milligrams as input, are likewise listed for the respective stages on the line labeled MASS. The mass loadings from each stage follow and are labeled MG/DNCM/STAGE (milligrams per dry normal cubic meter per stage).

The percentage of the total mass sampled contained in particles with diameters smaller than a particular D_{50} is called the CUMULATIVE PERCENT OF MASS SMALLER THAN D_{50}. It is the cumulative percentage of total mass accumulated to the stage j.

The cumulative mass loading of particles smaller in diameter than the corresponding D_{50} is listed in four different units: milligrams per actual cubic meter, milligrams per dry normal cubic meter, grains per actual cubic foot, and grains per dry normal cubic foot. Note that these are the same units used for calculating the total mass loading. They represent both metric and English units and both stack conditions and engineering dry standard conditions of temperature, pressure, and water content.

The geometric mean diameter for the particles collected on each stage is then listed in micrometers. The geometric mean diameter of a given stage may be expressed as the square root of the product of the D_{50} of the given stage and the D_{50} of the previous stage. In calculating the geometric mean diameter of the first stage (or cyclone if applicable), the maximum particle diameter is used instead of the "D_{50} of the previous stage." In calculating the geometric mean diameter of the filter, one-half the D_{50} of the last stage (stage eight, here) is used instead of the "D_{50} of the given stage." (There is no D_{50} for the back-up filter since all remaining particles are captured by this filter.)

Finally, an approximate differential particle-size distribution is listed as DM/DLOGD, in milligrams per dry normal cubic meter, and as number concentration, DN/DLOGD, in number of particles per dry normal cubic meter.

Differential size distributions may be derived two ways:

1. Finite difference methods may be used based on the D_{50}'s (abscissa) and the particulate masses on each stage (ordinate). This technique was used to generate the differential size distribution data in Table 3.

2. Curves may be fitted to the cumulative mass distribution from which the differential curves (slope) for each test can be calculated. This method is preferred and is described in the following paragraphs.

PROGRAM SPLIN1

In many, if not most, sampling programs, a number of impactor runs will be made. Frequently, these runs will be made using several impactors, having different performance characteristics. The latter may be true even if the same type of impactor is used throughout a sampling program. This behavior results both from manufacturing variations which cause calibration differences and run-to-run variations in sampling rates, which cause shifts in the D_{50}'s. Averaging results from such testing to obtain a representative composite size distribution requires that the distributions be broken down into like size intervals for all the runs to be averaged. The same requirement for like size intervals also holds for using inlet and outlet data from control device sampling programs to obtain fractional efficiencies. This requires curves to be fit to the data for each run to permit interpolation to obtain values at common diameters for all runs to be compared or averaged.

Before making the final selection of the spline technique for fitting curves to the size distribution data, consideration was given to a number of alternate fitting methods, and several of them were tried. It was concluded that any attempt to fit a predetermined functional form (e.g. log-normal) to the data was generally not proper. Because the slope of the cumulative distribution curve, the differential distribution, is the required quantity for calculating fractional efficiencies, consideration was also given to fitting curves to the $\Delta M/\Delta \log D$ approximations of the true differential distribution, which are estimated directly from the stage loadings and D_{50}'s However, the magnitude of the steps in D_{50} are large enough in most impactors as to frequently make $\Delta M/\Delta \log D$ a poor approximation to $dM/d\log D$. Moreover, the boundary conditions are more difficult to handle in fitting curves to $\Delta M/\Delta \log D$ than in fitting to the cumulative distributions. It was ultimately concluded, after many trial fitting methods were tested that the best use was made of the data if the fitting was done to the cumulative distribution curve by means of a "SPLINE" method.

SPLIN1 operates by fitting a curve which is continuous in X and Y and the first derivative of Y with respect to X to the cumulative mass concentration size distribution data. The resulting fitted curve is similar to that which one would draw through the data points using a "French curve" or mechanical spline. This fitted curve invokes no _a priori_ assumptions as to the shape of the distribution (i.e., power law, log-normal, etc.). The manner in which the spline fits are made is described below.

Initial attempts at using the spline technique on the set of points defining the cumulative distribution curve obtained directly from the D_{50}'s were not satisfactory. The difficulty occurred as a result of the inability of the method to generate sufficiently rapid changes in curvature when the curve to be generated was defined by a small number of points. A satisfactory fit could be obtained by adding a set of interpolated points between the original data points of the measured cumulative curve. These points are generated by means of a series of parabolas through consecutive sets of three adjacent data points of the original cumulative curve defined by the impactor stage data. The fitting is done using log (concentration) and log (particle diameter) as variables and begins with the segment containing the smallest D_{50} in the data set.

The sequence of operations by which the interpolated points are generated is shown in Figures 1. A series of parabolas are fit through consecutive sets of three data points beginning at the smallest D_{50} as shown in Figures 1a and 1b. In this description, three interpolation points between each pair of D_{50}'s will be assumed. However, the program will accommodate up to five interpolation points. The use of more points will improve the accuracy of the fitting, but will require more storage capacity. The interpolation points are located along the parabolas, between the lower pair of the three D_{50} points used to generate the parabola. The interpolated points are spaced evenly in log diameter between the pair of original points. A similar process is used to generate interpolated points between consecutive pairs of D_{50}'s up to the segment which terminates at the D_{50} of the first collection stage as illustrated in Figures 1c to 1e. A slightly different procedure which will be described later, is used for segments which include the first collection stage D_{50}.

Since the fitting is for a cumulative curve, negative slopes are not allowed. Therefore, a check is made for negative first derivatives of the interpolation parabola at the bounds of

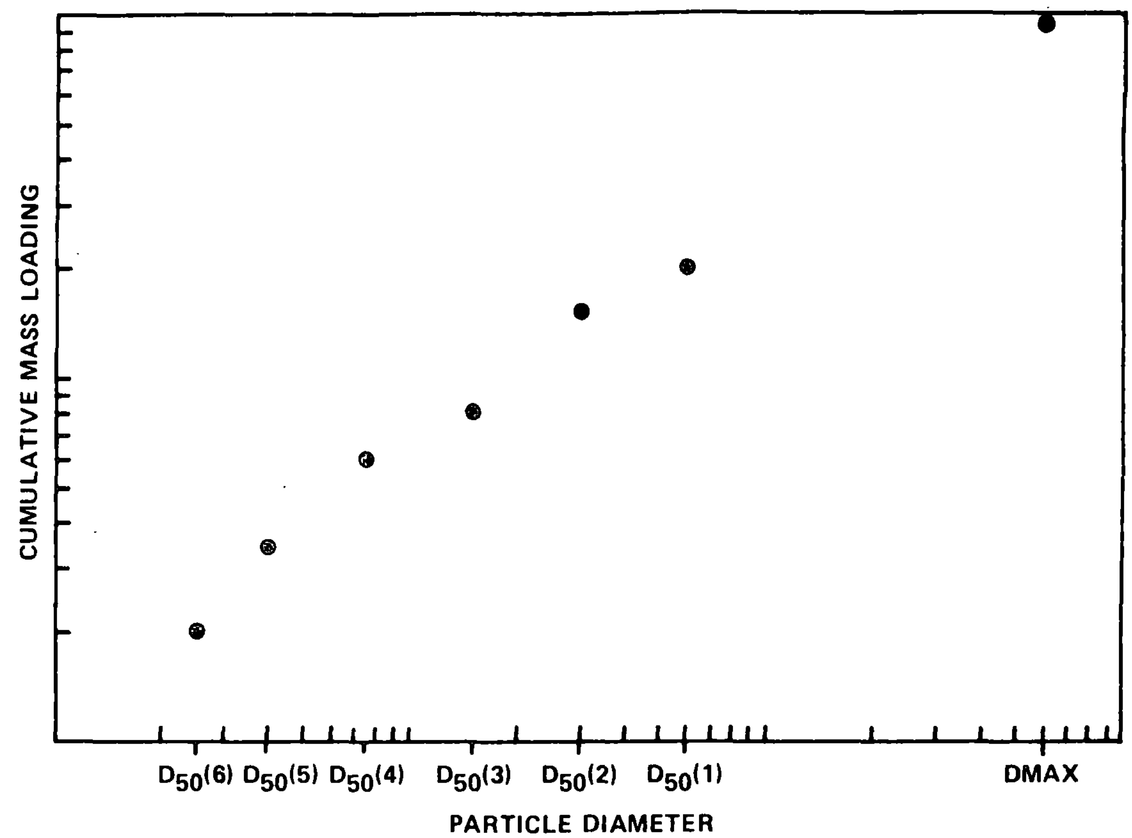

Figure 1a. Cumulative size distribution from raw impactor data.

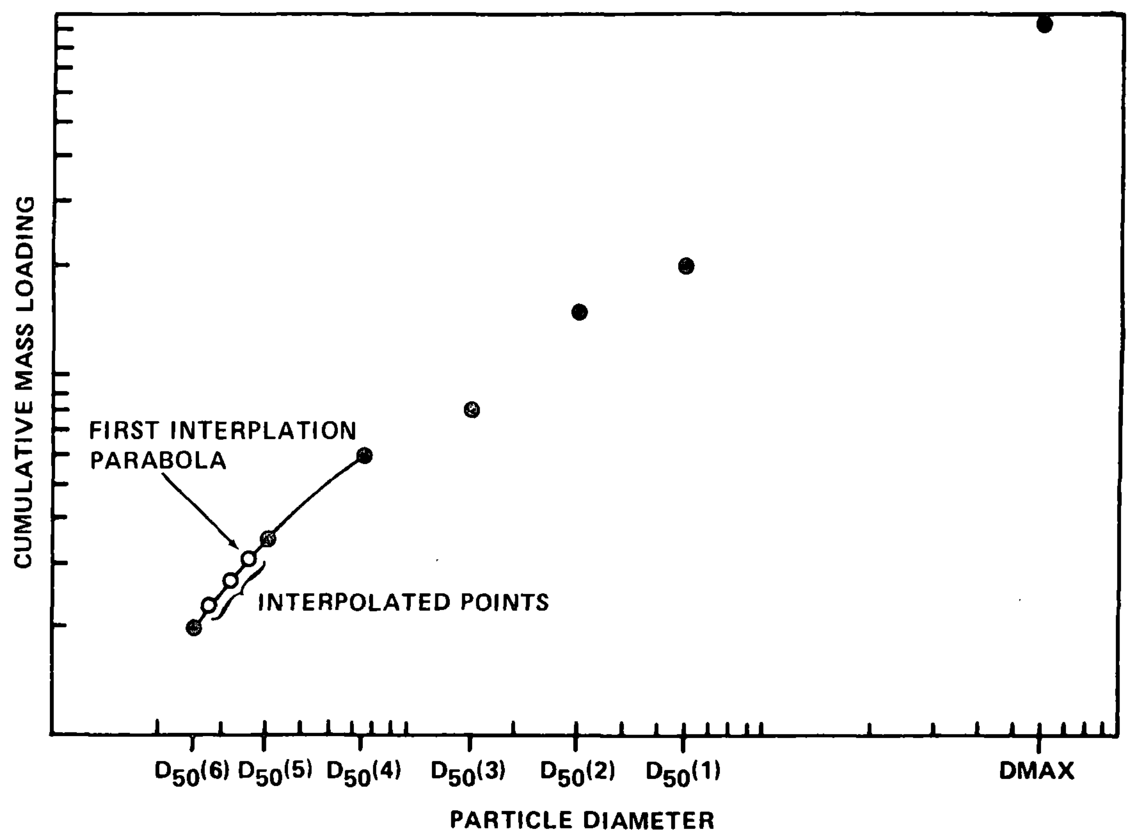

Figure 1b. Start of development of interpolated points between first and last D_{50}.

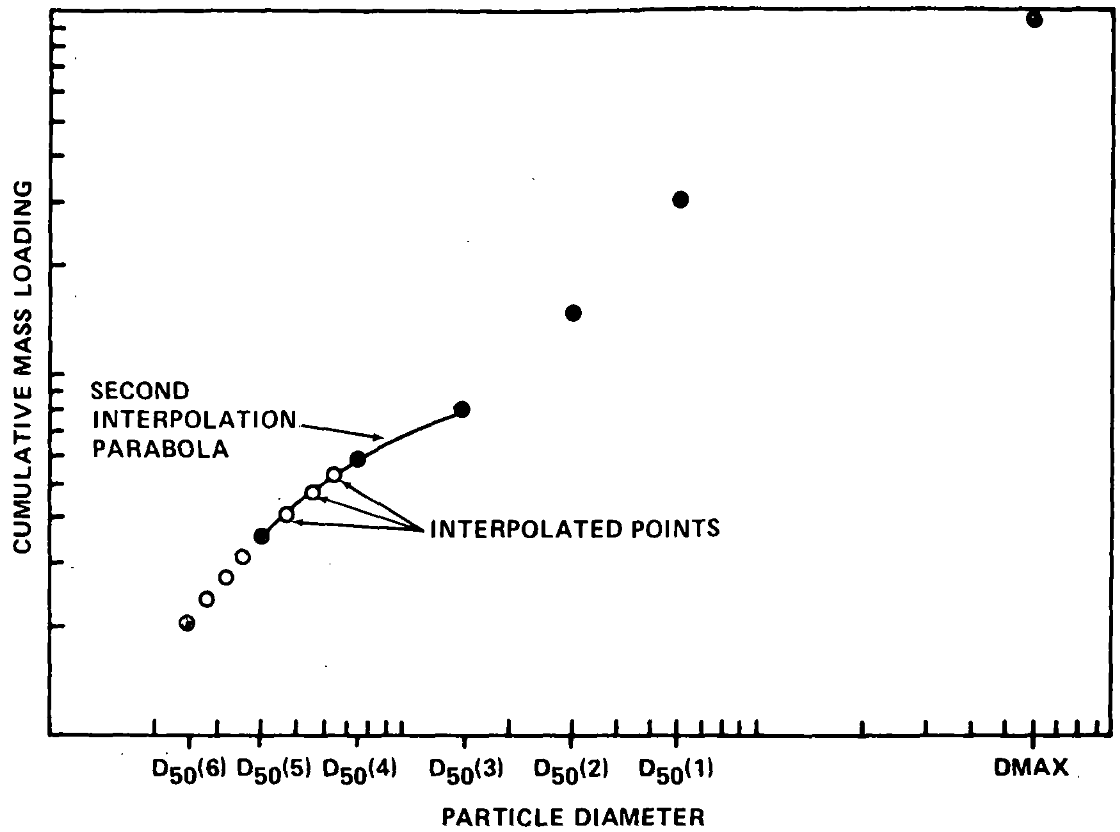

Figure 1c. Continued generation of interpolated points

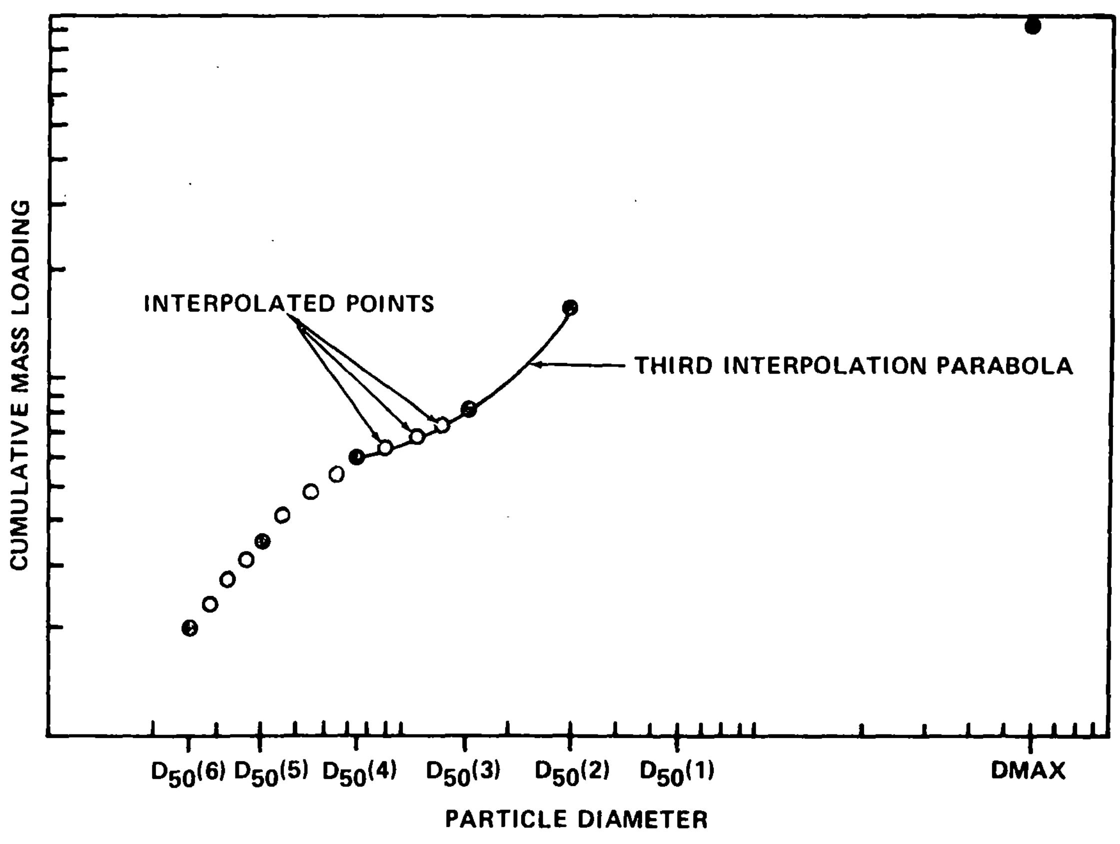

Figure 1d. Continued generation of interpolated points

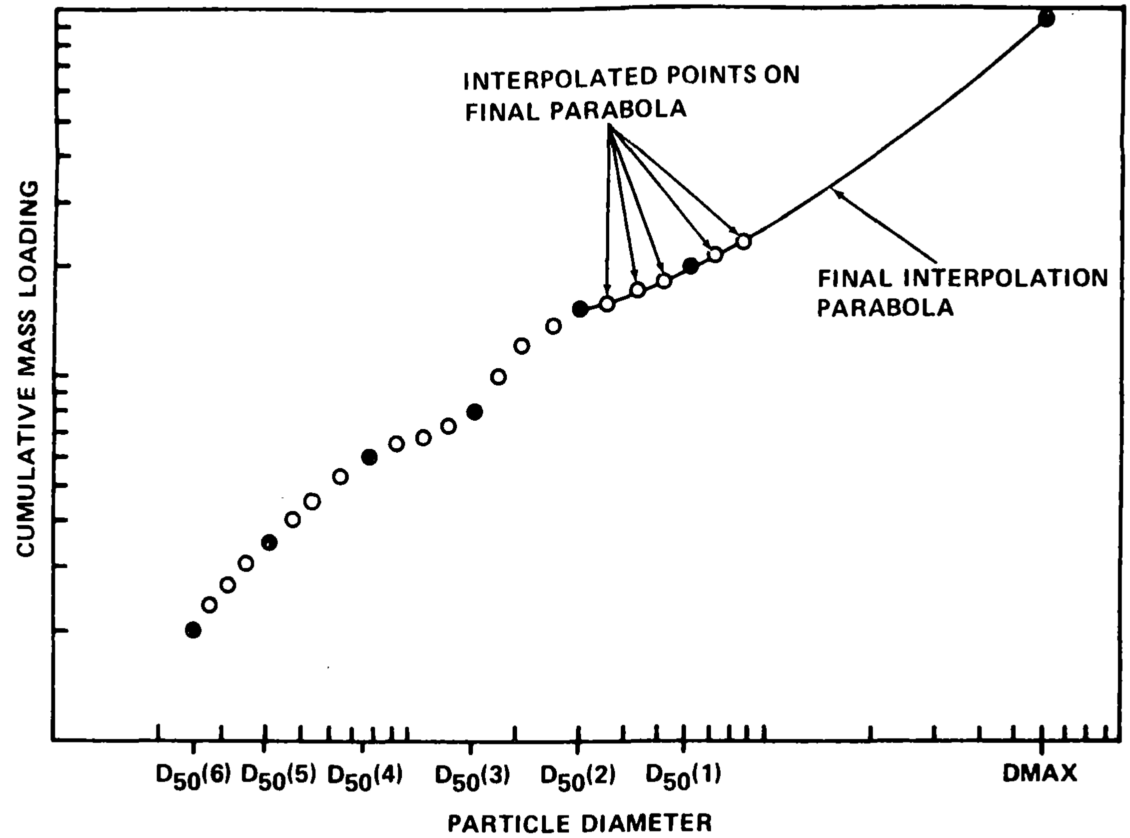

Figure 1e. Generation of interpolated points on parabola which includes DMAX.

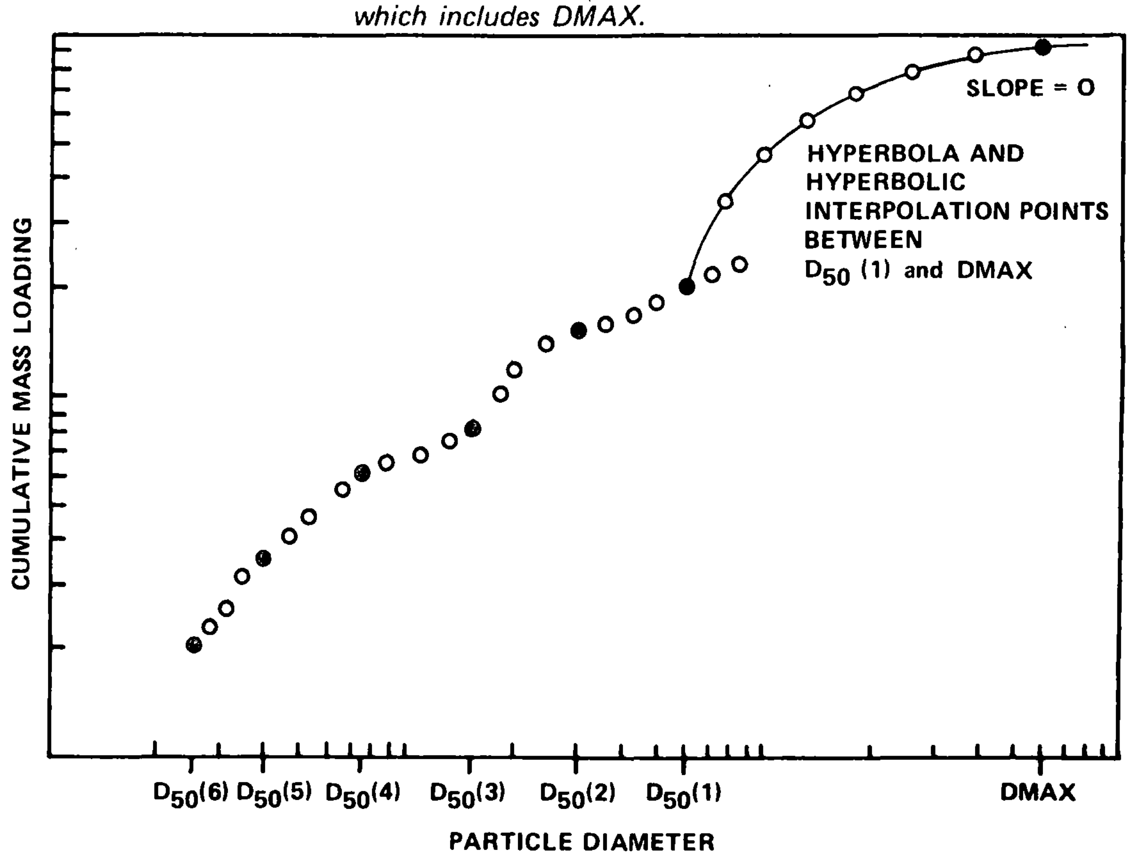

Figure 1f. Generation of interpolated points on hyperbola through $D_{50}(1)$ and DMAX

each segment within which the interpolated points are to be generated. If a negative derivative is found in any segment other than the first (the segment including the smallest D_{50}) a straight line interpolation between the segment bounds is used rather than parabolic interpolation. If a negative first derivative is found in the first segment to be fitted, a fictitious point is generated and used to form a parabola which has no negative derivatives in this segment. This fictitious point has the same concentration value as that of the first point on the cumulative curve and has a diameter defined by

$$D_{fictitious} = \frac{(D_{50} \text{ of last stage})^2}{(D \quad \text{of next to last stage})} \qquad (2)$$

The interpolated values for the segment between the last two D_{50}'s on the cumulative curve are then generated from the parabola which passes through this fictitious point, and the points for the last two stages on the cumulative distribution curve.

In the region about the first stage D_{50}, three sets of interpolated points are generated. The first are generated by parabolic interpolation using a parabola through DMAX, D_{50} (stage 1) and D_{50} (stage 2) as was done in the case of the previous segments. However, in addition to these, two more points are generated along the parabola above the first stage D_{50}. These additional points are spaced evenly in log (diameter) at the same intervals in log (diameter) as the interpolated points between D_{50} (stage 1) and D_{50} (stage 2) as shown in Figure 1e. These points are used in generating the final curve fit up to the point on the cumulative distribution curve defined by the first stage D_{50}. The third set of points is illustrated in Figure 1f.

Note that the cumulative mass distribution used in the illustrations of Figures 1 is one in which a large step in concentration occurs between D_{50} (stage 1) and DMAX. This is typical of a cumulative curve for a bimodal distribution in which one mode has a median diameter substantially greater than first stage D_{50}. The interpolation parabola through DMAX, D_{50} (stage 1) and D_{50} (stage 2) does not properly represent the shape of the true distribution curve in this region. In particular, the true curve must have zero slope at DMAX. It was empirically determined that a hyperbolic interpolation equation fit in terms of linear concentration and linear diameter between DMAX and D_{50} (stage 1) with the hyperbola asymptotic to the total loading at infinite particle size resulted in acceptable results for the final spline fits. Therefore a seven point hyperbolic interpolation is used in addition to the previously described parabolic interpolation over this segment of the curve. This hyperbolic interpolation is illustrated in Figure 1f. The use of the two sets of interpolated points in the final interval will be discussed later.

Generation of the Final Spline Fit

The original data points, defined by the D_{50}'s, together with the interpolated points just generated, form a set of points along a continuous curve (if one disregards the two sets of points in the final segment) which has no negative slopes. However, the derivative of the curve in most cases will not be continuous at the D_{50} points. The spline fit to be described is a smoothing technique which generates a series of parabolic segments that approximates a continuous curve through the complete set of points defining the cumulative distribution. The segments to be generated now will pass near or through those points and will have forced continuity in both coordinates and first derivatives. The technique is applied first to cover the interval between the first and last stage D_{50}'s and then a second time to cover the interval between the first stage D_{50} and DMAX. From this point on in the discussion, no distinction is made between the original points defined by the D_{50}'s and the interpolated values located between them.

The spline fit is generated by joining successive parabolas at locations determined by the x (or log diameter) coordinates of the points which now represent the cumulative distribution curve (original points at the D_{50}'s plus the interpolated points). These parabolas have continuity in slope forced by the fitting procedure and are generated in such a fashion as to pass near or through the points on the cumulative distribution curve.

The procedure is illustrated in Figures 2. The spline fit is begun at the lowest point, O, on the distribution curve (at the D_{50} of the last stage). The parabola used to generate the interpolated points between the last two stages is assumed to be the fitted curve up to the first interpolated point. (Point 1 in Figure 2a.) This parabola, a, is followed until the x-coordinate at point 1 is reached. At the point A, located on this parabola by the x-coordinate of point 1, a new parabola is fitted as shown in Figure 2b. The parabola, b, is forced to pass through point A with the same slope at A as the parabola used to define point A, and is forced to pass through the third point above point 1 in the set of points defining the cumulative curve, i.e. point 4. The parabola, b, is followed to the point defined by the x-coordinate of point 2, thus locating a point B. At B a new parabola is fit with forced slope continuity with b passing through the third point ahead of point 2, i.e., point 5, as shown in Figure 2c. From C this process is repeated using point C and 6 to generate a new parabola, d, and termination point D, e, and E, etc., until a termination point at the D_{50} of the first collection stage is reached. The last three points obtained by parabolic interpolation are used in generating the spline fit parabolas up to the first

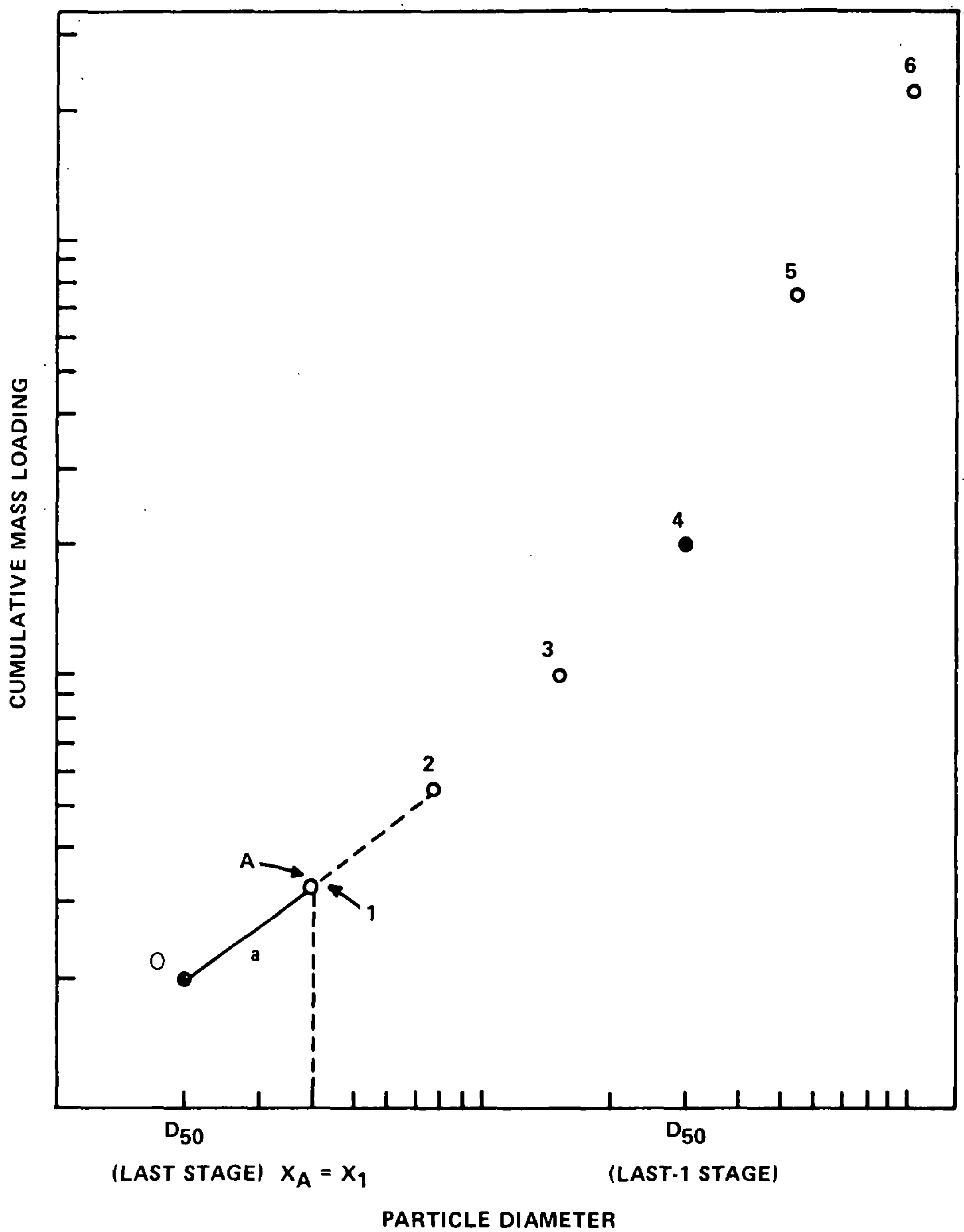

Figure 2a. *Start of the curve fitting procedure. Cumulative mass loadings derived from stage catches are represented by solid circles. Interpolated values are shown with open circles.*

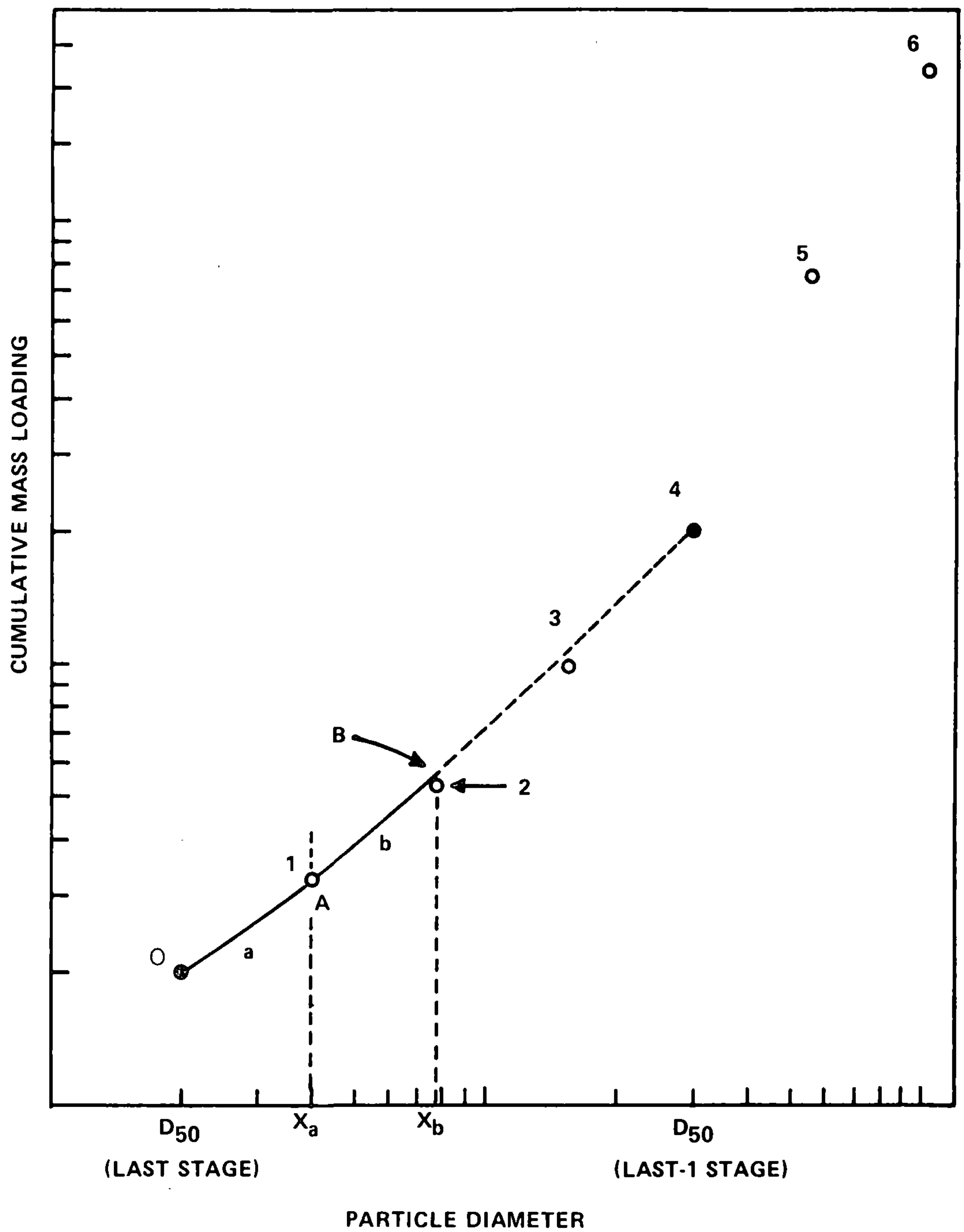

Figure 2b. *Second step in the curve fitting procedure. Cumulative mass loadings derived from stage catches are represented by solid circles. Interpolated values are shown with open circles.*

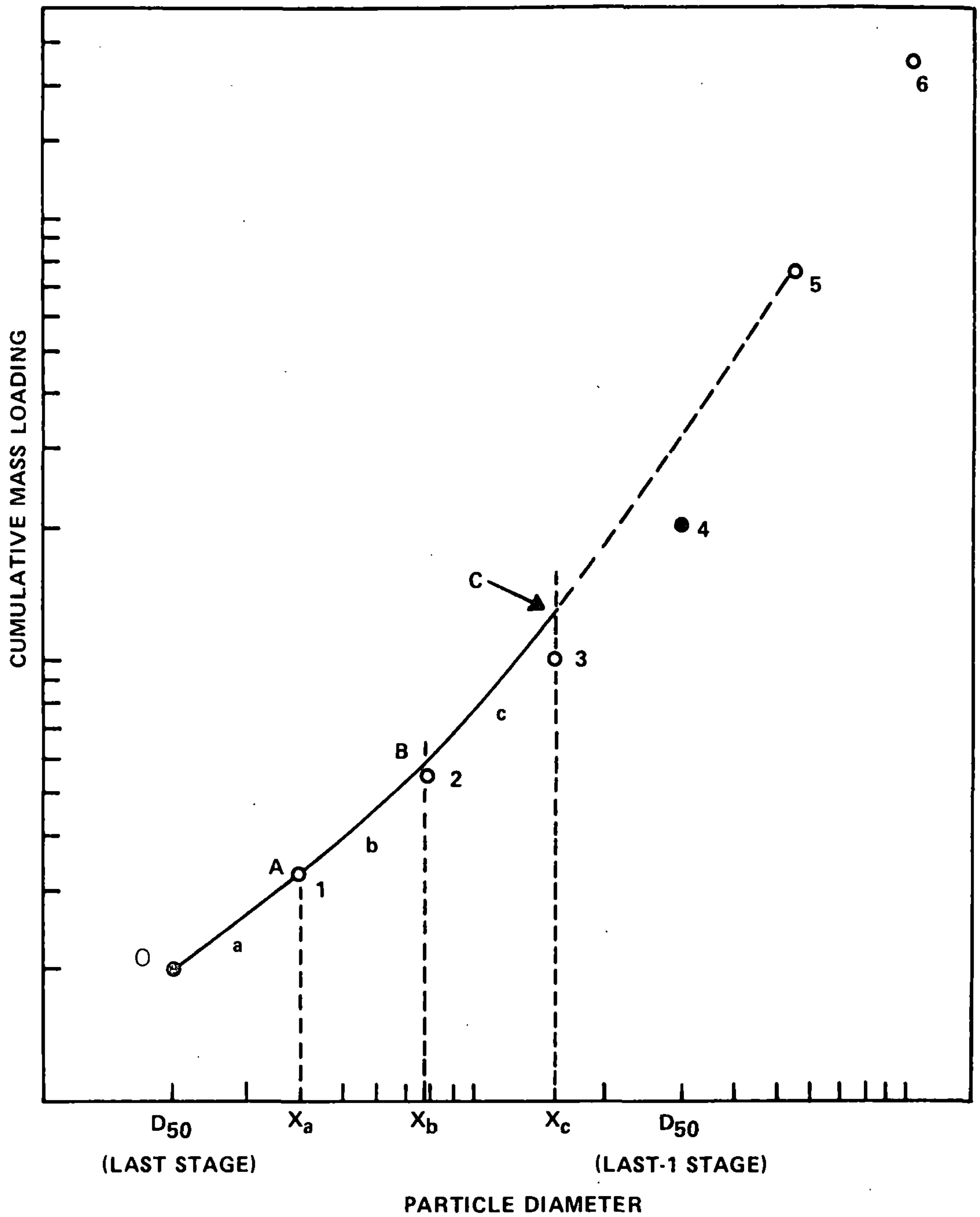

Figure 2c. Third step in the curve fitting procedure. Cumulative mass loadings derived from stage catches are represented by solid circles. Interpolated values are shown with open circles.

collection stage D_{50}. The coefficients of the fitting spline fit parabolas for the segments a, b, c, d, . . etc., are saved for future use. These now represent the smoothed curve and will be used henceforth to define the cumulative curve for that run.

The final spline fit starts by picking up at the point on the fitting parabola which terminated at the D_{50} of the first stage. The same procedure as before is followed, except that the third point ahead determined by the hyperbolic interpolation is now used for fitting, and the fitting parabolas are followed to x-coordinates defined by the hyperbolic interpolation points. The curve generated in this second zone of the spline fit [i.e., between D_{50} (stage 1)and DMAX] is an extrapolation which has been found to be reasonably good to diameters equal to about 2 to 3 times the first stage D_{50}. By using the second, third (as illustrated), fourth, etc., point ahead in generating the final parabola segments, one can influence the amount of smoothing provided by the program.

The cumulative concentration and slope of the cumulative curve, dm/dlogD, can be calculated for any arbitrary particle size by locating the fitting coefficients for the spline segment containing that size. The boundary locations of each of the parabolic segments, O, A, B, C, . . ., and the fitting coefficients for each segment are stored in a disk file for subsequent use by other programs (e.g., GRAPH, STATIS, etc.).

Problems Resulting from Extremely Close Stage Cut Diameters (D_{50}'s)

When two stages are used on an impactor which differ only slightly in D_{50}, the second of the two will collect too much material because of the finite slope of real impactor stage collection characteristics. The simplest example of this effect would be obtained if two identical stages are used sequentially. If that were the case, in an ideal impactor the second stage should collect no material; however, because of the finite slope of the real stage collection efficiency curve, it will collect some particles. This would lead to the formation of a step increase (infinite slope) in the cumulative concentration curve. The severity of the effect is reduced as the spacing between the D_{50}'s increases but can be sufficiently severe so as to cause significant errors in the size distribution curves if it is not properly accounted for. Calibrations indicate that the effective cut diameters, or D_{50}'s, at the first two stages of several impactors suffer from this problem. The program MPPROG, because of this, ignores the presence of the second stage of Andersen, MRI, and University of Washington impactors in generating the cumulative mass concentration curve from which the fitted curves will be made by SPLIN1. This procedure effectively nullifies the problem. However, if calibrations of future versions of these impactors do not show the small spacing in D_{50}, MPPROG should be modified appropriately so as not to lose good information when the curve fits are made.

PROGRAM GRAPH

Program GRAPH is dedicated entirely to presenting data from single impactor runs. The output forms available on call are cumulative mass loading versus D_{50}, $\Delta M/\Delta \log D$ versus geometric mean diameter, and $\Delta N/\Delta \log D$ versus geometric mean diameter as calculated in MPPROG. The latter are available on both Stokes, aerodynamic and aerodynamic impaction diameter bases. As an option, up to ten runs can be superimposed on a single plot. Plots and tabular output of the fitted curves from SPLIN1 are also available. The fitted curves from SPLIN1 are plotted superimposed on the data points from MPPROG, but only as single run plots. The plots are all made on log-log grids.

The tabular output includes cumulative percent mass loading less than particle diameter generated from the SPLIN1 fitted curves, dM/dlogD versus particle diameter, and dN/dlogD versus particle diameter generated by differentiation of the SPLIN1 fitted curves.

SECTION 4

ANALYSIS OF GROUPED DATA

PROGRAM STATIS

STATIS is a program for combining data from multiple impactor runs under a common condition. The program tests data from a series of runs (specified by the user) for outliers, flags and removes outliers from the set, and then provides output in the form of average size distributions with confidence intervals as desired in both tabular and graphical form. The program is set up to provide 50% confidence intervals; however, changes can be made for the calculation of other confidence intervals as desired (e.g., 90% or 95%).

The input data to STATIS are the fitted polynomial segments generated from MPPROG by SPLIN1 which now define the cumulative mass loadings for each run. The particle diameter basis for averaging (i.e., aerodynamic, aerodynamic impaction, Stokes) is user specified on control cards used to execute STATIS.

The fitting equations from SPLIN1 are differentiated at preselected particle diameters to obtain the quantity $(dM/d\log D_i)_j$ where i refers to particle diameter and j refers to the sequence number of a particular run in the set to be averaged. The values, at each particle diameter, D_i, are subjected to an outlier analysis based on the deviations of the values of dM/dlogD for individual runs from the mean for all runs.

The outlier test used is that for the "Upper 5% Significance Level".[4] A curve fitted to the tabular list of critical values for excluding an outlier is used to generate the table. The value X_i is excluded from statistical analysis based on the following condition:

$$\frac{|X_i - \overline{X}|}{S} > C_n$$

where X_i = individual value

$\overline{X}$ = mean of all values

S = standard deviation of the data set

C_n = critical value = function of number of values in the data set, n.

The application of this test requires that there be three or more runs in the sequence to be averaged. This outlier test is repeated after discarding any outliers already identified, provided there are at least three runs remaining in the set of retained points.

After discarding outliers, a final average, standard deviation, and confidence interval are calcuated for each $(dM/dlogD_i)$. These values are output on the line printer and are plotted on call by the user.

Cumulative size distributions on a mass basis or percentage basis are derived from the averaged dM/dlogD values by integration of these values. The choice of integrating the dM/dlogD curve rather than direct computation of the cumulative averages from the individual cumulative distributions is based on the fact that an error in a single stage weight is propagated forward throughout the cumulative curve for all stages subsequent to the one on which the error occurred. This would cause substantial quantities of good data from other stages to be discarded by the outlier analysis. Integration of the averaged differential distribution, on the other hand, allows the data from the remaining, error free, stages to have their proper influence on the averaged cumulative distributions. These cumulative distributions are again output in tabular form and, on call, in graphical form.

The cumulative distributions can be obtained either including or excluding particles smaller than 0.25 μm in diameter. The option of excluding the particles smaller than 0.25 μm is made available because of the fact that in a significant percentage of sampling situations, impactor back up filter catches can be dominated by oversize particles because of bounce and/or reentrainment. This results in a filter weight gain which can be many times higher than the weight of the fine particles which, ideally, should be the only material present. In those cases, omission of the material which is nominally smaller than 0.25 μm from the cumulative distributions will make the result a much better representation of the true size distribution. This, of course, is true only when the D_{50} of the last impactor stage is about 0.25 to 0.5 μm as is usually the case with the commercially available impactors.

Standard deviations and confidence limits for the cumulative distributions are calculated from the approximation that the variance (and square of a confidence interval) for a sum, A + B, is given by the sums of the variances (and squares of the confidence intervals) for A and B separately, i.e.,

$$\text{Variance}_{A+B} = \text{variance}_A + \text{variance}_B \qquad (3)$$

and

$$(\text{confidence interval})^2_{A+B} = (\text{confidence interval})^2_A + (\text{confidence interval})^2_B. \quad (4)$$

The averaged differential size distributions generated by STATIS are stored in a disk file for use by the programs PENTRA or PENLOG in calculating control device fractional efficiency curves.

Tabular and graphical output from STATIS includes cumulative mass loading versus diameter, cumulative percentage on a mass basis versus diameter, dM/dlogD versus diameter, and dN/dlogD versus diameter. The graphical presentations are made on log-log grids with the exception of the cumulative percentage plot which is made on a log-probability grid. All output forms, graphical and tabular, include confidence limits. The choice of diameter definition used is left to the user. An index of runs which were rejected through the outlier analysis before averaging is also printed. Rejection at any one particle size does not result in the run being excluded at all particle sizes.

PROGRAMS PENTRA/PENLOG

These two programs are virtually identical and provide tabular and graphical output of control device penetration and/or efficiency versus particle size for a preselected series of particle sizes from about 0.25 to 20.0 µm. The only difference between the two programs is in the form of the graphical output. In the case of PENTRA, the fractional efficiency curves are presented on a log-probability grid while in PENLOG they are presented on a log-log grid.

The calculations are made from averaged sets of inlet and outlet data developed by STATIS. The user identifies the pair of averaged data sets from which the efficiency is to be calculated together with the diameter basis required (i.e., Stokes, aerodynamic, aerodynamic impaction). The program retrieves the appropriate averaged data sets and calculates the fractional efficiency as

$$\text{efficiency}_i(\%) = 1.0 - \frac{(\text{dm/dlogD}_i)_{\text{outlet}}}{(\text{dM/DlogD}_i)_{\text{inlet}}} \times 100.0 \quad (5)$$

where i refers to the i^{th} particle diameter in the preselected diameter sequence. Simultaneously, if both the inlet and outlet data sets included two or more runs, confidence limits are calculated based on a method described by Y. Beers.[5] The confidence level associated with the limits generated by the program as provided are 50% levels; however, other levels can be generated by simply changing values of three constants used to generate the appropriate t-table.

SECTION 5

TEST CASES AND EXAMPLES OF RESULTS

Tests of the final fitting process were made by generating fictitious impactor runs having known size distributions. These runs were generated by calculating the D_{50}'s associated with several sets of sampling conditions. The stage weights required to produce exact unimodal and bimodal log normal distributions were then generated for these sets of particle diameters. The program was exercised on these artificially constructed runs and results obtained from the fitting procedure were then compared to the original distributions. Figures 3 and 4 show examples of two such tests in the form of differential size distributions. Figure 3 illustrates the input distribution and recovered distribution for an aerosol having a mass median diameter of 4.0 μm and geometric standard deviation of 3.0. Recovered distributions from the spline fit and the approximation results, $\Delta M/\Delta \log D$, taken directly from the stage weights both show excellent agreement with the input distribution. The fitted results for diameters larger than 7.0 μm represent an extrapolation to sizes larger than the first stage D_{50}. Figure 4 illustrates a similar test for a bimodal distribution having equal amplitude modes, mass median diameters of 2.0 μm and 10.0 μm, and geometric standard deviations of 1.5. Again, beyond about 7.0 μm, the fitted points represent extrapolations. Note that the $\Delta M/\Delta \log D$ approximations derived directly from the stage weights lie very close to the input curve in regions where the slopes are not large but fall significantly above the true curve in regions of high slope. Errors expressed as percentage deviations from true values are shown for two cases each for unimodal MMD = 4.0, σg = 2.0, and bimodal MMD = 2.0, σg = 1.5, and MMD = 10.0, σg = 1.5 distributions in Figures 5 and 6. Note that the results from the fitted curves generally fall within ±10% or better of the true values in the size interval covered by the impactor stage D_{50}'s and are for the most part within ±50% of the true values in the extrapolation region above the first stage D_{50}. Much larger errors occur with the $\Delta M/\Delta \log D$ approximations to the differential distributions obtained directly from the stage weights. The errors shown in Figures 5 and 6 result only from the fitting procedure and do not include any effects from non-ideal behavior in the impactors. Errors arising from the latter can be much greater, as described by McCain and McCormack (1978).[6]

Examples of some of the graphical output formats available from the program are shown in Figures 7 through 10. Figure 7

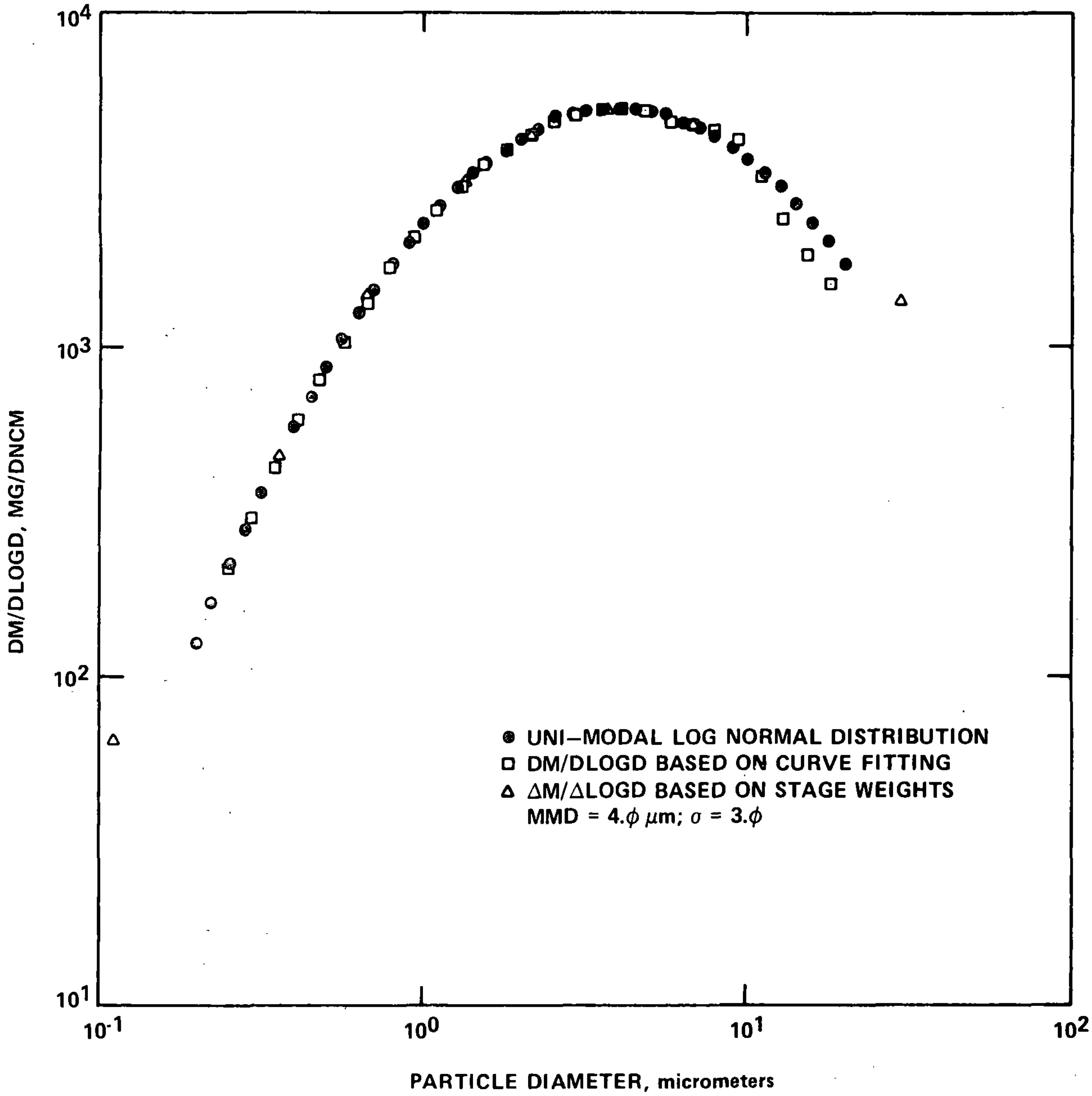

Figure 3. Approximate differential size distribution based on stage weights and same distribution based on spline fitting are compared to a true unimodal log normal distribution.

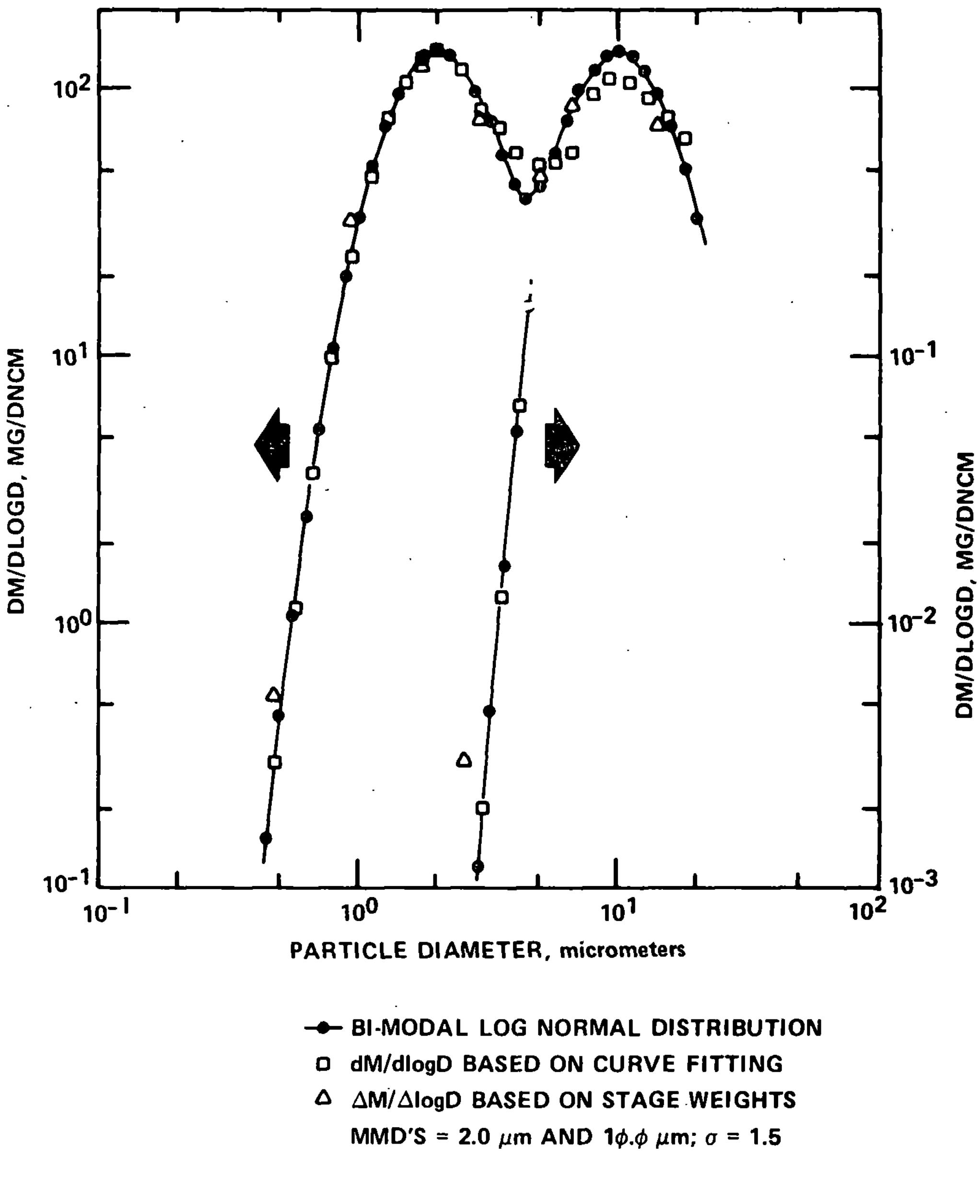

Figure 4. Approximate differential size distribution based on stage weights and same distribution based on spline fitting are compared to a true bimodal log normal distribution.

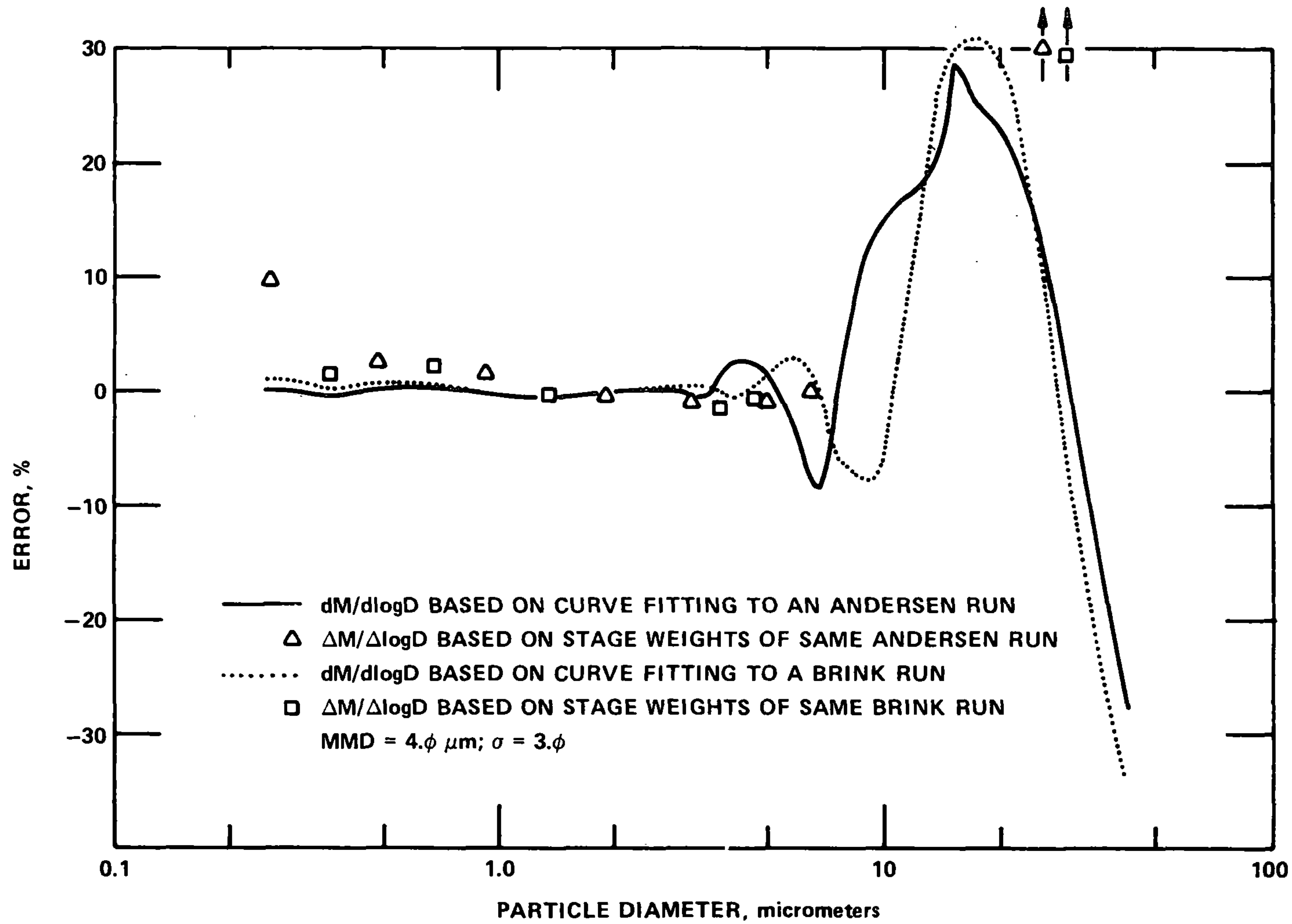

Figure 5. Percent error of approximate differential size distributions based on stage weights and the same distributions based on spline fitting from a unimodal log normal distribution.

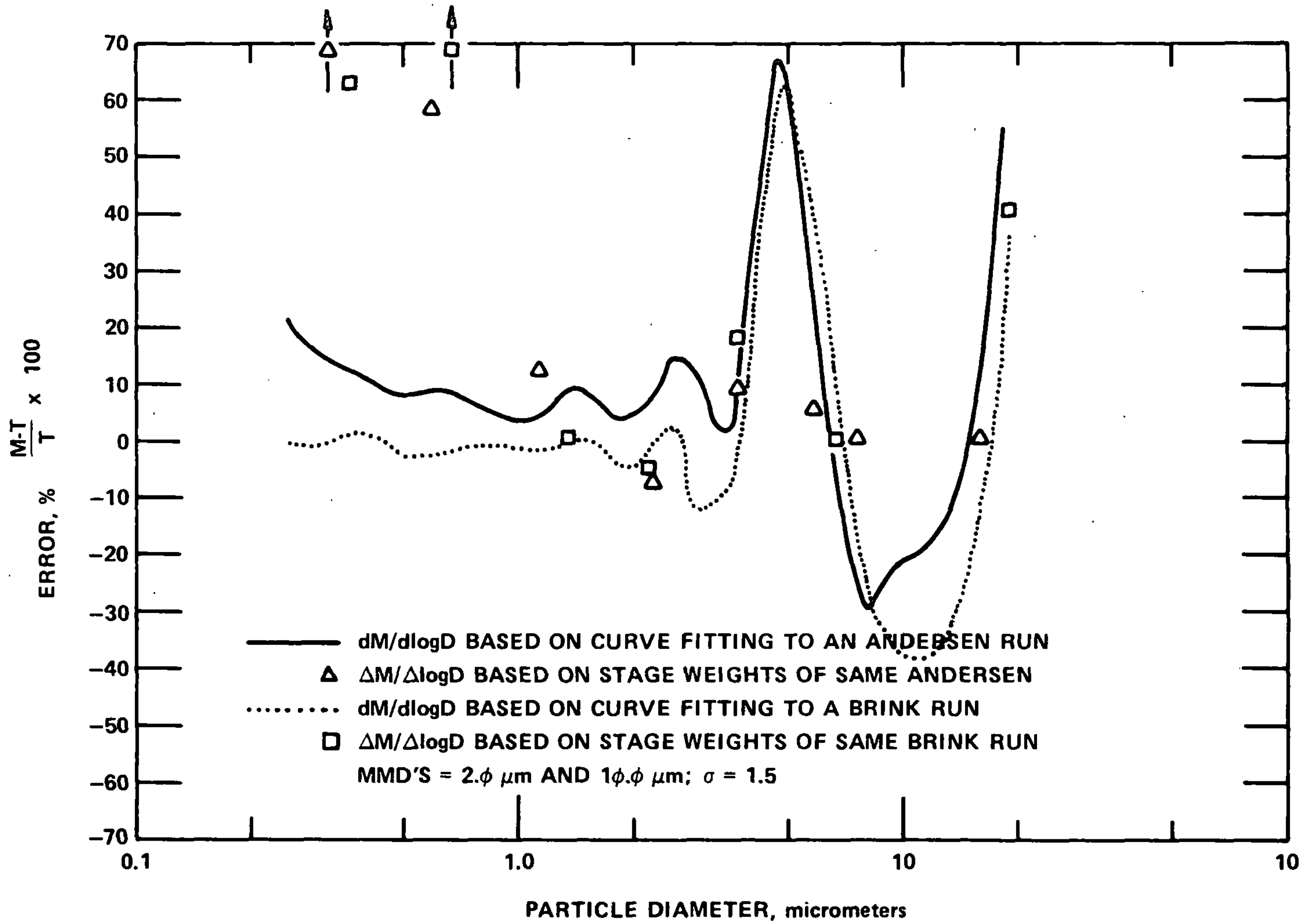

Figure 6. Percent error of approximate differential size distributions based on stage weights and the same distributions based on spline fitting from a bimodal log normal distribution.

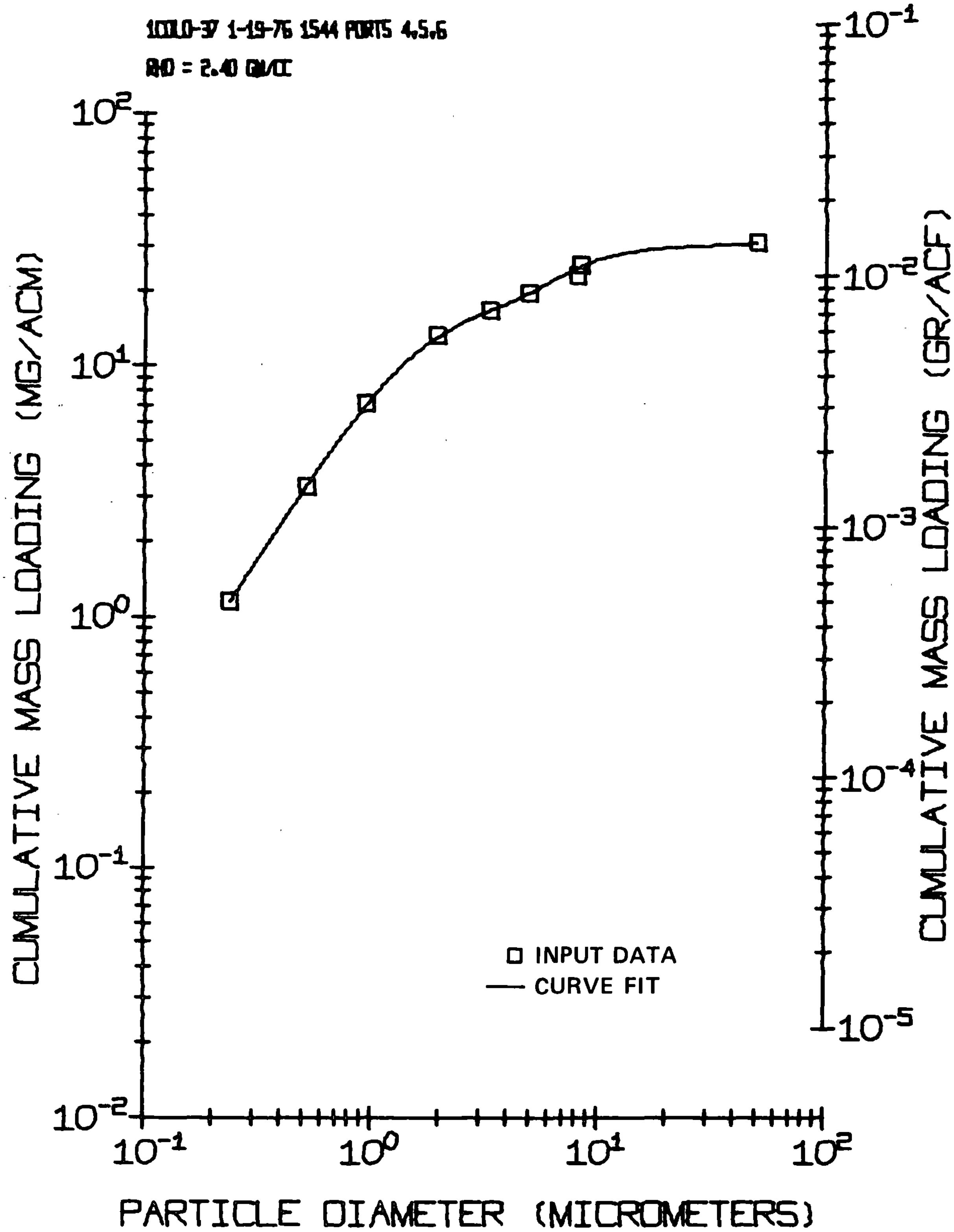

Figure 7. Single run cumulative mass distribution with original data points based on stage weights and fitted curve from SPLIN1.

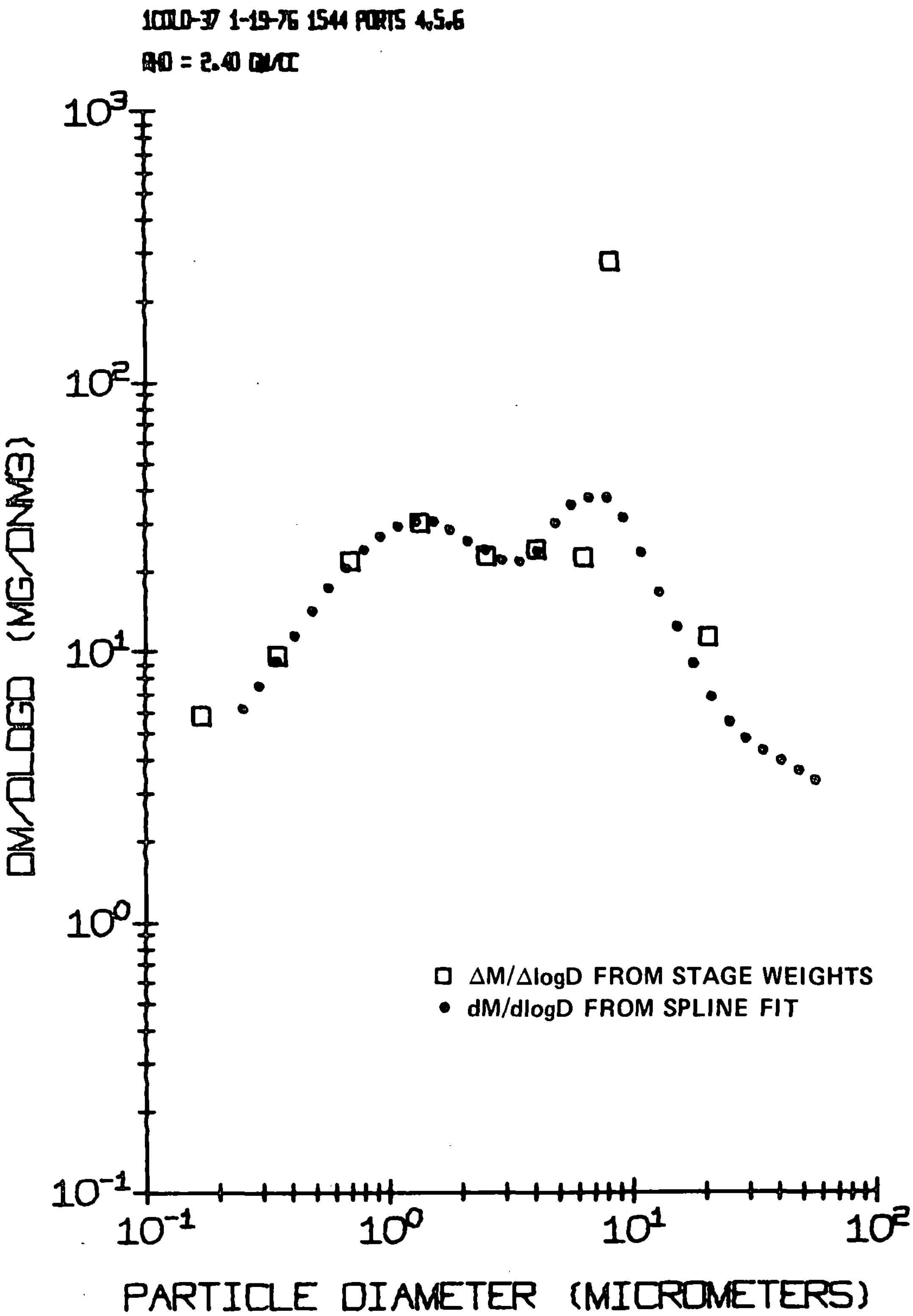

Figure 8. Differential size distribution obtained for the run in Figure 7 based on stage weights and based on curve fitting.

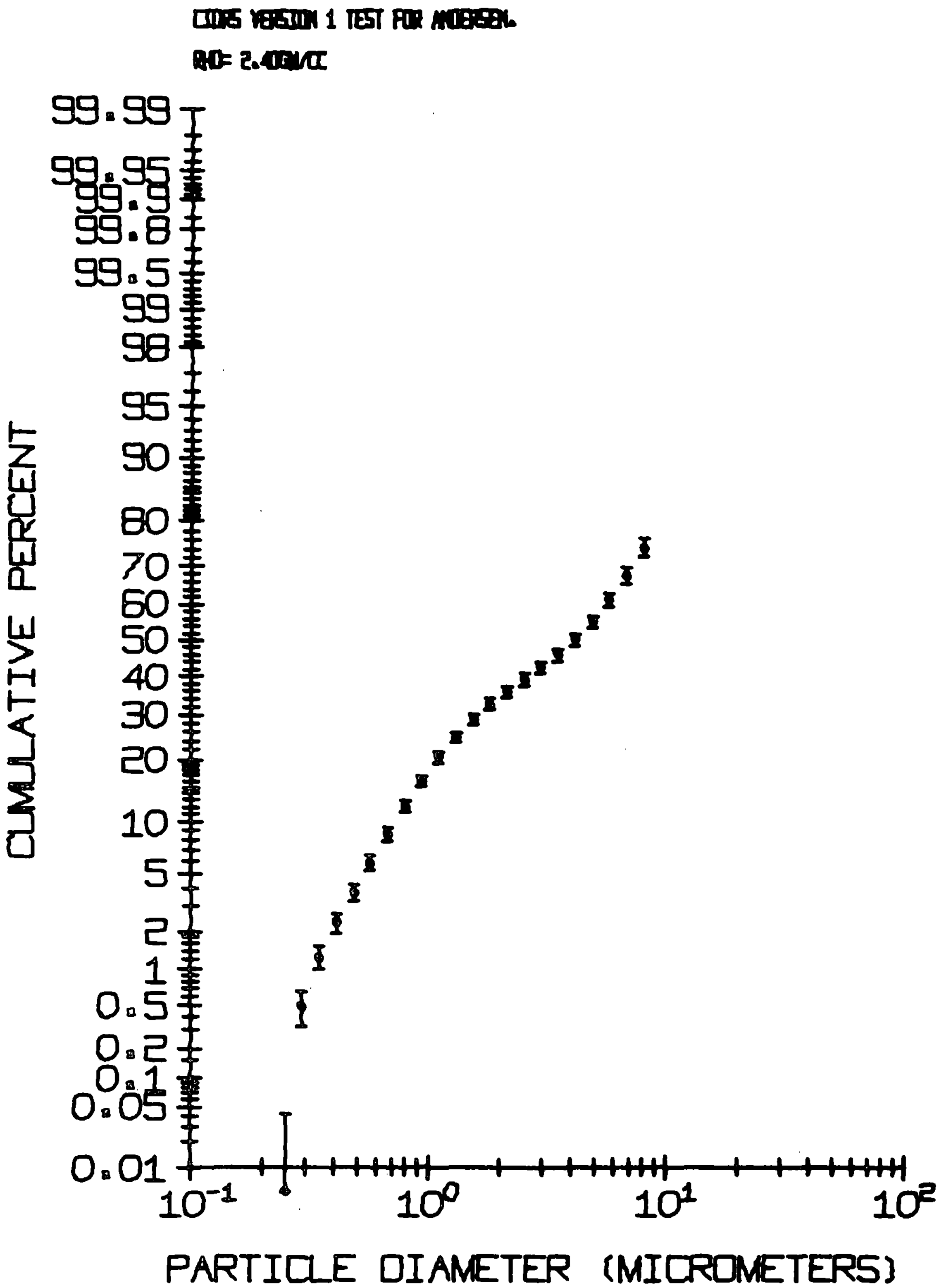

Figure 9. Cumulative mass distribution on a percentage basis with confidence limits obtained from the average of several runs similar to that shown in Figures 7 and 8.

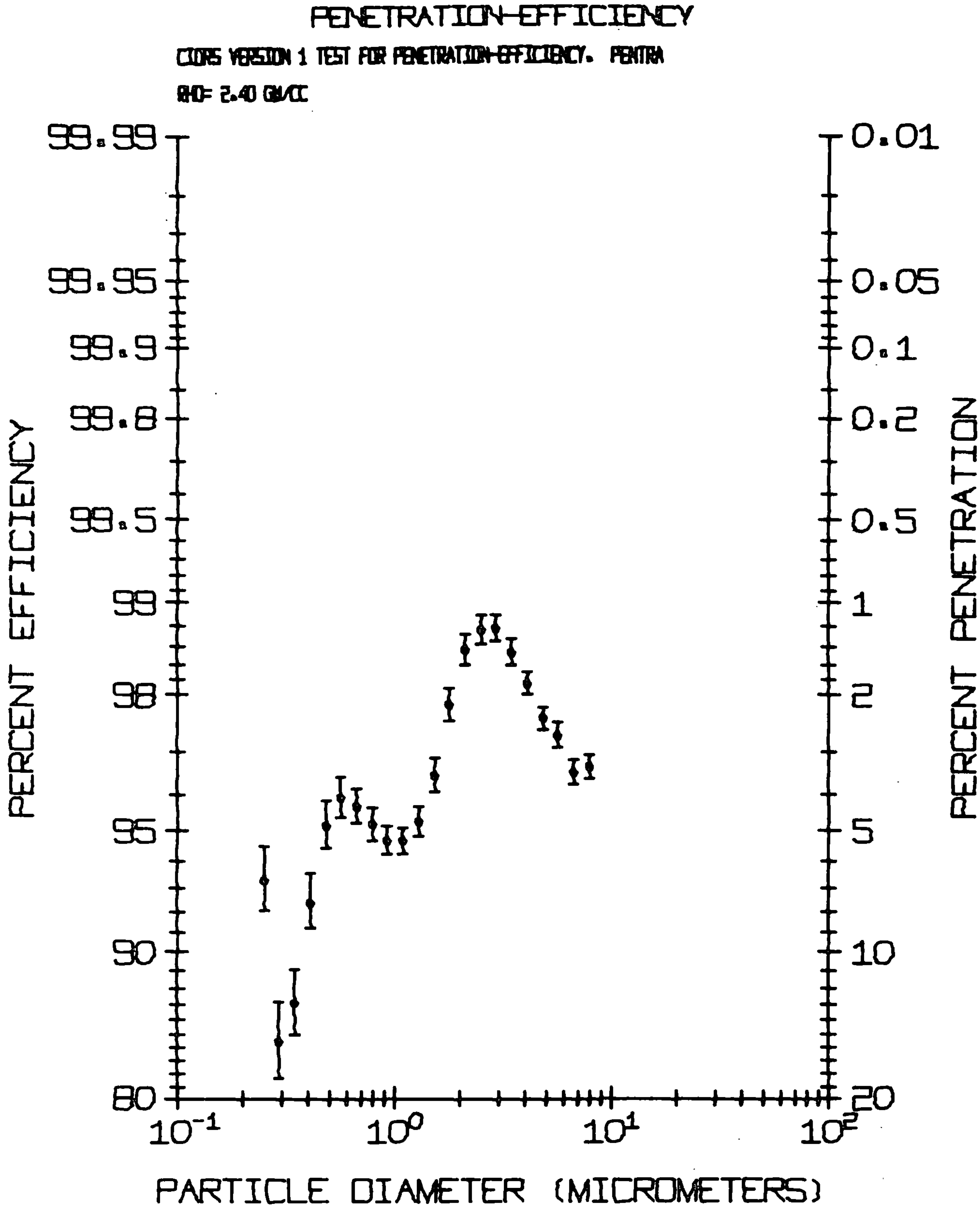

Figure 10. A control device penetration curve with confidence limits obtained from sets of averaged inlet and outlet runs.

illustrates a single run cumulative mass distribution with the original data points and fitted curve from SPLIN1. Figure 8 shows the differential distribution obtained for the run shown in Figure 7. Figure 9 illustrates a cumulative mass distribution on a percentage basis with confidence limits obtained from the average of several runs similar to that shown in Figures 7 and 8. Figure 10 illustrates a control device penetration curve with confidence limits obtained from sets at averaged inlet and outlet runs.

REFERENCES

1. Johnson, J. W., G. I. Clinard, L. G. Felix, and J. D. McCain. A Computer-Based Cascade Impactor Data Reduction System. EPA-600/7-78-042, March, 1978. 601 pp.

2. Morrow, P. E., (Chairman, Task Group on Lung Dynamics). Deposition and Retention Models for Internal Dosimetry of the Human Respiratory Tract. Health Physics, 12:173-208. 1966.

3. Mercer, T. T., M. I. Tillery, and H. Y. Chow. Operating Characteristics of Some Compressed Air Nebulizers. Am. Ind. Hyg. Assoc. J. 29:66-78. 1968.

4. Quality Assurance Handbook for Air Pollution Measurement Systems, Vol. 1. Principles. EPA-600/9-76-005, 1976.

5. Beers, Y. Introduction to the Theory of Error. 2nd Edition. Addison-Wesley, Reading, Mass. 1957.

6. McCain, J. D., and J. E. McCormack. Non-Ideal Behavior in Cascade Impactors. 70th Annual Meeting, APCA, Toronto, 1977. Paper 77-35.3.

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-600/7-78-132a	2.	3. RECIPIENT'S ACCESSION NO.
4. TITLE AND SUBTITLE A Data Reduction System for Cascade Impactors		5. REPORT DATE July 1978
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) J.D. McCain, G.I. Clinard, L.G. Felix, and J.W. Johnson		8. PERFORMING ORGANIZATION REPORT NO. SoRI-EAS-78-331
9. PERFORMING ORGANIZATION NAME AND ADDRESS Southern Research Institute 2000 Ninth Avenue, South Birmingham, Alabama 35205		10. PROGRAM ELEMENT NO. EHE624
		11. CONTRACT/GRANT NO. 68-02-2131, T.D. 10101
12. SPONSORING AGENCY NAME AND ADDRESS EPA, Office of Research and Development Industrial Environmental Research Laboratory Research Triangle Park, NC 27711		13. TYPE OF REPORT AND PERIOD COVERED Task Final; 3-5/78
		14. SPONSORING AGENCY CODE EPA/600/13

15. SUPPLEMENTARY NOTES IERL-RTP project officer is D. Bruce Harris, Mail Drop 62, 919/541-2557.

16. ABSTRACT The report describes a computer-based data reduction system for cascade impactors. The system utilizes impactor-specific calibration information, together with operating conditions and other pertinent information (e.g., stage weights, sampling duration), to determine particle size distributions in several forms for individual runs. The program can handle all commercial round-jet cascade impactors, including common modifications, which are in current use in stack sampling. Other round-jet impactors can be easily substituted. Slotted impactors could be accommodated with slight program revision. A spline technique is applied to fit a curve in the cumulative size distribution obtained from each individual impactor run. The fitted curves have forced continuity in coordinates and slopes. Size distribution averages for multiple runs are made using the fitted curves for interpolation at consistent particle diameters, regardless of the diameters at which the data points fall in the original individual run data sets. After statistical analyses to locate and remove outliers from the data being averaged, averages, variances, standard deviations, and confidence intervals are calculated. The averages and statistical information are available in tables and graphs in several size distribution formats. Averaged data are stored on disks for subsequent use.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Air Pollution; Sampling	Air Pollution Control	13B; 14B
Dust	Stationary Sources	11G
Impactors	Particulate	13I
Data Reduction	Cascade Impactors	09B
Size Separation	Size Distribution	07A,13H
Flue Gases		21B

18. DISTRIBUTION STATEMENT Unlimited	19. SECURITY CLASS *(This Report)* Unclassified	21. NO. OF PAGES 44
	20. SECURITY CLASS *(This page)* Unclassified	22. PRICE

EPA Form 2220-1 (9-73)