INTERNATIONAL STANDARD

ISO 9276-3

First edition
2008-07-01

Representation of results of particle size analysis —

Part 3: Adjustment of an experimental curve to a reference model

Représentation de données obtenues par analyse granulométrique —

Partie 3: Ajustement d'une courbe expérimentale à un modèle de référence

Reference number
ISO 9276-3:2008(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 9276-3 was prepared by Technical Committee ISO/TC 24, *Sieves, sieving and other sizing methods*, Subcommittee SC 4, *Sizing by methods other than sieving*.

ISO 9276 consists of the following parts, under the general title *Representation of results of particle size analysis*:

— *Part 1: Graphical representation*

— *Part 2: Calculation of average particle sizes/diameters and moments from particle size distributions*

— *Part 3: Adjustment of an experimental curve to a reference model*

— *Part 4: Characterization of a classification process*

— *Part 5: Methods of calculation relating to particle size analyses using logarithmic normal probability distribution*

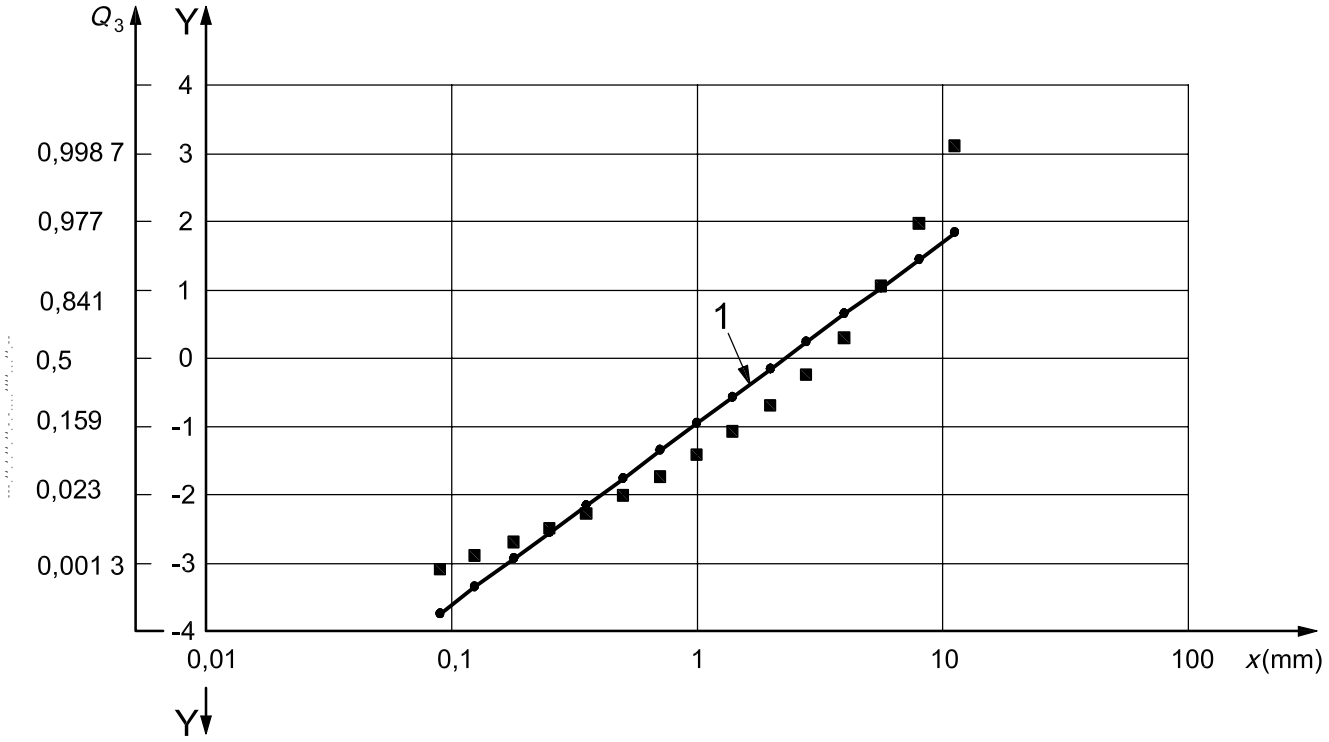
The following part is under preparation:

— *Part 6: Descriptive and quantitative representation of particle shape and morphology*

Introduction

Cumulative curves of particle size distributions are sigmoids, therefore fitting to a model distribution function or rendering statistical intercomparison is difficult. These disadvantages can, however, be remedied by transforming these sigmoids into straight lines by means of appropriate coordinate systems, e.g. log-normal, Rosin-Rammler or Gates-Gaudin-Schuhmann (log-log). Target size distributions in particle technology industries can also be described in terms of distribution models.

In such systems, a classic linear regression assumes that the squares of the deviations between the experimental points and the theoretical straight line are, on average, equal. This is only valid in the transformed cumulative distribution value system, but not in their linear representation, and therefore named a quasilinear regression. In particular, the scale extension makes the values of the squares of the deviations at the extremities of the graph vary by several orders of magnitude. In addition, the sum of the squares of the deviations obtained by this method is not related to any simple distribution and does not allow any statistical test.



Key

$Q_3(x)$	cumulative distribution by volume or mass
x	particle size
Y	quantiles of the standard normal distribution
1	quasilinear regression full line
•	quasilinear fit point
▪	Q_3 (x) data point

Figure 1 — Example of a functional paper with log-normal plot (cumulative distribution values plotted on a normal ordinate against particle size on a logarithmic abscissa with inverse standard normal distribution transformed) and quasilinear regression full line

The experimental data in Figure 1 are taken from ISO 9276-1:1998[1], Annex A and represent a sieve-measuring result example between 90 µm and 11,2 mm.

The mathematical treatment, corresponding to non-linear coordinate systems, mentioned above, agrees with a quasilinear regression. Here the non-linear transformation of the Y-axis results in a non-linear transformation of the Y-deviations, e.g. another consideration of deviations at the tails of a distribution than at their centre.

One possibility to compensate for the non-linear transformation of the Y-differences, in the result of the non-linear transformation of the Y values, is the introduction of weighting factors in the quasilinear regression (see Annex E).

Moreover, a non-linear regression delivers the best adjustment and allows the most flexibility, such as statistical tests on number distributions, the adjustment of truncated or multimodal distributions or any other arbitrary models, but it requires a start approximation and a numerical mathematical procedure.

The standard deviation of residuals between experimental points and the model in the non-transformed scale allows the quantification of the degree of alignment and the statistical comparison of experimental distributions. A value of greater than e.g. 0,05 indicates a non-adequate reference model.

Representation of results of particle size analysis —

Part 3:
Adjustment of an experimental curve to a reference model

1 Scope

This part of ISO 9276 specifies methods for the adjustment of an experimental curve to a reference model with respect to a statistical background. Furthermore, the evaluation of the residual deviations, after the adjustment, is also specified. The reference model can also serve as a target size distribution for maintaining product quality.

This part of ISO 9276 specifies procedures that are applicable to the following reference models:

a) normal distribution (Laplace-Gauss): powders obtained by precipitation, condensation or natural products (pollens);

b) log-normal distribution (Galton MacAlister): powders obtained by grinding or crushing;

c) Gates-Gaudin-Schuhmann distribution (bilogarithmic): analysis of the extreme values of the fine particle distributions;

d) Rosin-Rammler distribution: analysis of the extreme values of the coarse particle distributions;

e) any other model or combination of models, if a non-linear fit method is used (see bimodal example in Annex C).

This part of ISO 9276 can substantially support product quality assurance or process optimization related to particle size distribution analysis.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 9276-2, *Representation of results of particle size analysis — Part 2: Calculation of average particle sizes/diameters and of moments from particle size distributions*

ISO 9276-5, *Representation of results of particle size analysis — Part 5: Methods of calculation relating to particle size analyses using logarithmic normal probability distribution*

3 Symbols and abbreviated terms

a	straight line intercept (equation of a straight line)
b	slope (gradient) of the straight regression line (equation of a straight line)
d'	intercept parameter of RRSB distribution
GGS	(Gates-) Gaudin-Schuhmann distribution
LND	logarithmic normal probability distribution, defined in ISO 9276-5
n	number of size classes
n_F	degrees of freedom, which is the number of data points, n, minus the number of fit model parameters
N	number of particles in the measured sample
$\boldsymbol{p}$	set of model parameters, vector
q	density of particle size distribution
$Q(x)$	observed cumulative distribution, total of the particles finer than x, between 0 and 1
$Q^*(x; \boldsymbol{p})$	model estimation, theoretical cumulative distribution depending on the reference model with parameters, $\boldsymbol{p}$
r	type of quantity of a size distribution, $r = 0$: number, $r = 3$: volume or mass
RRSB	Rosin-Rammler (Sperling and Bennet) distribution (derived from Weibull-distribution)
s	standard deviation of LND, logarithm of geometric standard deviation [ISO 9276-5]
s_{ql}	mean square deviation of the quasilinear regression in the transformed scale
s_{res}	standard deviation of the residuals, square root from residual variance
x	particle size
$x_{50,r}$	median particle size of distribution with type of quantity, r, intercept parameter of LND
$x_{max,r}$	intercept parameter of GGS distribution with type of quantity, r
$X(x)$	transform of x plotted on the x-axis [$X = x$ for a normal distribution and $X = \ln x$ or $\lg x$ for a log-normal, Rosin-Rammler or bilogarithmic (log-log) distribution], X is equivalent to ξ in ISO 9276-1 and ISO 9276-5
$Y(Q)$	transform of Q plotted on the y-axis ($Y =$ inverse of standard normal distribution for a normal distribution, see Table 1 for other model types)
$Y^* = a + bX$	general expression of the equation for the straight regression line of a model cumulative particle size distribution
z	dimensionless normalization variable in LND [ISO 9276-5]
α	slope parameter of GGS distribution
ζ	integration variable, based on z, in LND
ν	exponent of RRSB distribution
ω	weighting coefficient

4 Adjustment of an experimental curve to a reference model

4.1 General

The estimation of parameters to be used in the regression equations appearing in this part of ISO 9276 are calculated from either particle size distribution values, Q, fractions of these particle size values, dQ, or density values, q. These particle size distribution parameters may also be used as parameters for other regression equations.

Generally a certain distribution model $Q^*(x; \boldsymbol{p}) = Q^*(x; a,b\ldots)$

should be adjusted to measuring data: $[x_i, Q_i = Q(x_i)]$ $i = 1,..., n$

The intention and capability of the regression equation is to find the optimum parameters $p = a$, b... such that the mean square deviation between measured Q values, $Q(x)$, and the model, $Q^*(x; p)$, will be minimized:

$$s^2(\boldsymbol{p}) = \frac{1}{n}\sum_{i=1}^{n}\left[Q^*(x_i; \boldsymbol{p}) - Q(x_i)\right]^2 \xrightarrow[\boldsymbol{p}]{} \min \tag{1}$$

4.2 Quasilinear regression method

The non-linear (or rather non-linear) optimization problem in Equation (1) can be transformed by Y to a linear Equation (2) for the various statistical models used in this part of ISO 9276. The values of X are the transformed particle size values obtained from any particle size distribution.

$$Y^* = Y^*(Q^*) = a + bX \tag{2}$$

The solution and optimization using a linear regression with Equation (2) in the transformed state, delivers an approximation for Equation (1), which can be replaced with the following quasilinear regression Equation (3):

$$s_{ql}^2(\boldsymbol{p}) = \frac{1}{n}\sum_{i=1}^{n}\left[bX + a - Q(x_i)\right]^2 \xrightarrow[\boldsymbol{p}]{} \min \tag{3}$$

The solution of Equation (3) minimizes the absolute deviations in the transformed format (see Figure 1).

This quasilinear regression can also be used for all standardized particle size distributions using the various transformation equations listed in Table 1 (Reference [3]).

The ordinates, designated Y, are the transforms of the $Q(x)$ cumulative distribution values obtained by the formula of the relevant reference model.

The quasilinear regression is an analytical method, it requires no start approximation. But the non-linear transformation of the Y-axis results in a non-linear transformation of the Y-deviations, e.g. percentage deviations have to be considered differently at the tails of a distribution compared to at their centre.

The extension of this method to a weighted quasilinear regression method also does not deliver the optimum adjustment, see Annex E.

4.3 Non-linear regression method

4.3.1 General

Finding the general optimum model parameters in the linear scale according to Equation (1) is not possible with analytical equations; a numerical optimization procedure, known as non-linear regression, is required.

A non-linear regression requires a start approximation and a numerical mathematical procedure (Reference [4]). If, however, this non-linear regression approach is used, an optimum adjustment and a flexibility may be conveyed to statistical tests of number distributions or to the adjustment of truncated, multimodal distributions or any other arbitrary models.

The estimation of parameters, for use with various types of standardized distribution used as reference models (e.g. normal, LND, RRSB or GGS), is based on different strategies, when either a number or a mass (or volume) distribution is considered (Reference [5]). The star symbol in Equations (4) and (5) indicates the model estimation while the emboldened symbol $\boldsymbol{p}$ represents the model parameters to be optimized.

Table 1 — Equations used for three statistical models

Quantity	**Model**		
	LND (see also ISO 9276-5)	**RRSB**	**GGS**
Distribution model	$Q(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{\zeta^2}{2}\right) \mathrm{d}\zeta$ with $z = (\ln x - \ln x_{50,r})/s$	$Q(x; d', \nu) = 1 - \exp\left[-\left(\frac{x}{d'}\right)^{\nu}\right]$	$Q_r(x) = \begin{cases} \left(\frac{x}{x_{\max,r}}\right)^{\alpha} & \text{for} \quad x \leqslant x_{\max,r} \\ 1 & \text{for} \quad x > x_{\max,r} \end{cases}$
Intercept, a	$x_{50,r}$	d'	$x_{\max}$
Slope, b	$1/s$	ν	α
$Y(Q)$	$Q(Y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Y} \exp\left(-\frac{\zeta^2}{2}\right) \mathrm{d}\zeta$ with the standard normal distribution, $Y = \Phi^{-1}(Q)$	$Y = \ln [-\ln (1-Q)]$	$Y = \ln Q$
$X(x)$	$\ln x$	$\ln x$	$\ln x$
Linear model	$Y = \frac{1}{s} X - \frac{\ln x_{50,r}}{s}$	$Y = nX - n \ln d'$	$Y = \alpha X - \alpha \ln x_{\max}$

All the non-linear (numerical) estimation strategies need a first estimate of the adjustment parameters before starting the numerical procedure. The best starting estimate may be obtained from the quasilinear regression with Equation (3).

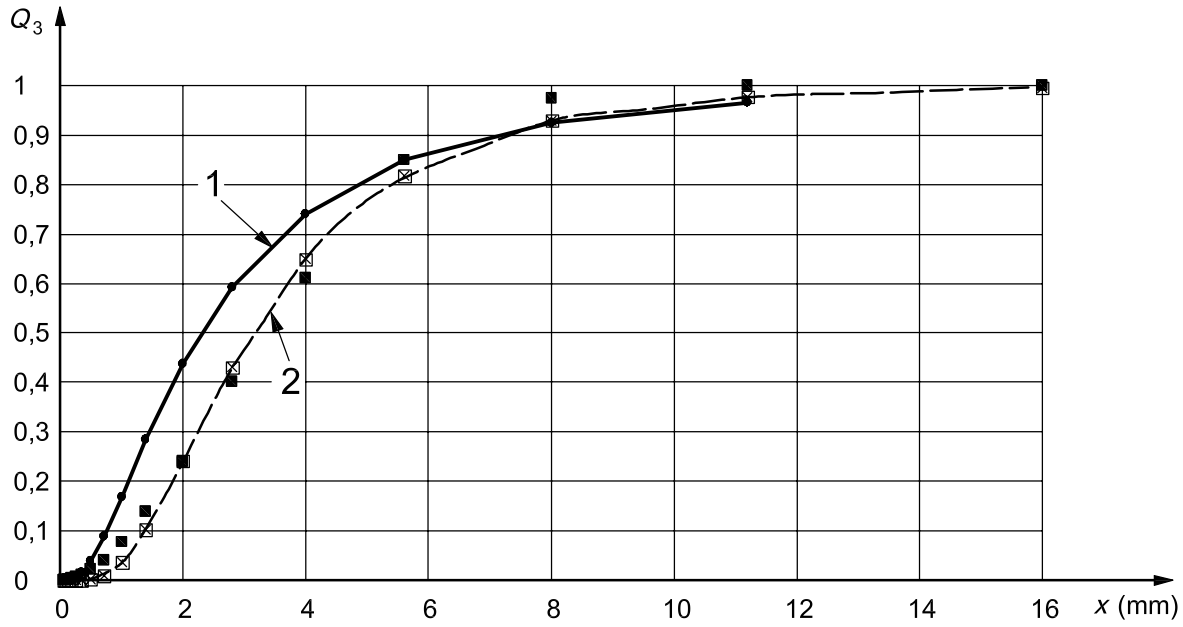
The numerical procedure may be based for instance on the Levenberg-Marquardt method, which is a popular alternative to the Gauss-Newton method (References [7], [8]). Some spreadsheet programs include a non-linear regression tool (add-in) for easy numerical optimization, for instance based on a code from Reference [9].

4.3.2 Estimation criterion for both a mass (or volume) distribution and a number distribution

The minimum sum of the squares of the deviations (the least squares) between measured Q values, $Q(x)$, and the model, $Q(x, \boldsymbol{p})$, is written for the example of a mass-related distribution as

$$s^2 = \sum_{i=1}^{n} \left[Q_3^*(x_i; \boldsymbol{p}) - Q_3(x_i) \right]^2 \xrightarrow[\boldsymbol{p}]{} \min \tag{4}$$

Figure 2 shows the quasilinear regression line from Figure 1 as a curve in linear scales and the non-linear regression from the least squares of the same data, which obviously represents a better adjustment of the experimental data. The quantification of the goodness of fit is given in Clause 5.



Key

Q_3 cumulative distribution by volume or mass

x particle size

1 quasilinear regression full line

2 non-linear regression — least squares

● quasilinear fit point

☒ least squares fit point

■ $Q_{3,i}$ measured

Figure 2 — Log-normal distribution: the quasilinear regression from Figure 1 in linear scales and the non-linear regression from the least squares of the same data, $(Q_3 - Q_3^*)$

Examples of how the experimental sieve-analysis data obtained from ISO 9276-1:1998[1], Annex A, can be approximated and transformed by either RRSB or GGS state models are shown in Annex A.

The influence of the type of quantity of the distribution on the goodness of fit is shown in Annex B. Different types of quantity place emphasis of adjustment on different size ranges.

Annex C shows the spreadsheet example calculations for the numerical procedure of the non-linear fit in Figure 2. Furthermore, an example for a bimodal distribution with five model parameters to be optimized is shown, using the same algorithm.

4.3.3 Estimation criterion for number distributions only and known sample size as particle number, N

Another estimation criterion for non-linear fit, which can be used only for number distributions and known sample size, N, is the χ^2-minimum criterion:

$$\chi^2 = N\sum_{i=1}^{n} \frac{\left\{\left[Q_0(x_i) - Q_0(x_{i-1})\right] - \left[Q_0^*(x_i;\, \boldsymbol{p}) - Q_0^*(x_{i-1};\, \boldsymbol{p})\right]\right\}^2}{Q_0^*(x_i;\, \boldsymbol{p}) - Q_0^*(x_{i-1};\, \boldsymbol{p})} \xrightarrow[\boldsymbol{p}]{} \min \tag{5}$$

It can quantify the improvement of accuracy by the measurement of larger particle numbers. This criterion compares the observed particle number variance in the numerator of Equation (5) with that predicted by Poisson statistics in the denominator of each size class.

Annex D shows the application for a χ^2-test of number distributions of known sample size, which quantifies the importance of large sample sizes for the analysis data interpretation.

5 Goodness of fit, standard deviation of residuals and exploratory data analysis

The basic regression, Equation (1), is used to find the optimum parameters, $\boldsymbol{p} = a, b...$, in such a way that the mean square deviation between measured Q values and the model Q^* is minimized.

Therefore, the sum of the error (the difference between the model and the measured values, squared), which is also known as the mean residual variance, is an adequate quality parameter for an adjustment. The square root, the standard deviation of the residuals, s_{res}, can be expressed by Equation (6) (Reference [5]):

$$s_{\text{res}}(\boldsymbol{p}) = \sqrt{\frac{1}{n_{\text{F}}} \sum_{i=1}^{n} \left(Q^*(x_i; \boldsymbol{p}) - Q(x_i) \right)^2} \tag{6}$$

where n_{F} represents the number of degrees of freedom, equal to the number of data points, n, minus the number of fit parameters (2 parameters, a and b, in the models in Table 1).

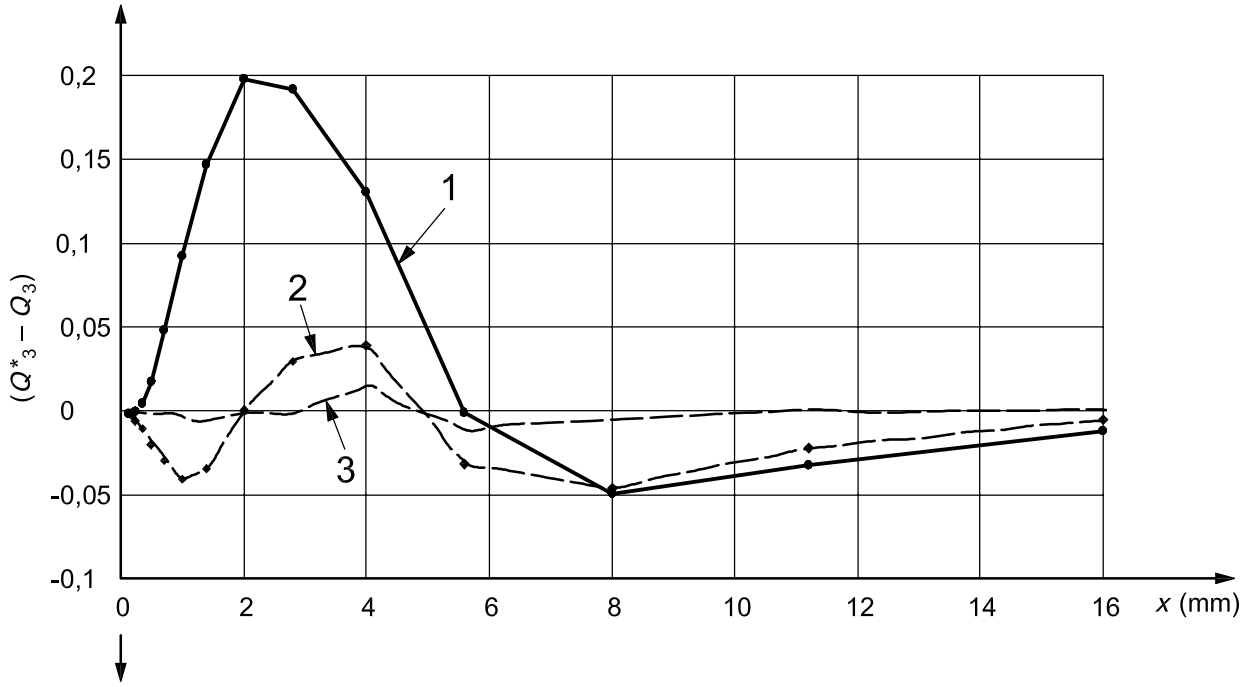
Equation (6) corresponds to the well-known and defined equation for the standard deviation of a distribution. Equation (6) may be used in the choice of the best model distribution, but can also be used to quantify the quality of fit of the several regression methods within one distribution model.

Pure statistical interpretation of the standard deviation of the residuals requires independent measurements between the different size classes. This is fulfilled exactly in the case of sieve-analysis by independent (parallel) sieving of each single sieve with the same split input sample. Already arranging the sieves in a tower will result in error propagation. But even for not exactly independent (uncensored) data points in the different size classes, Equation (6) represents the best measure for relative comparison of regression results.

For instance, the non-linear regression to log-normal distribution in Figure 2 reduces the standard deviation of Q_3-residuals, s_{res}, from 0,099 6 for the quasilinear regression to 0,026 2. The RRSB model gives the lowest standard deviation of 0,005 4 after application of the non-linear regression method. The corresponding residuals themselves are plotted in Figure 3.

For exploratory data analysis, the focus is on the data — their structure, outliers and models suggested by the data. Exploratory data analysis techniques are generally graphical.

Graphical representation of the residual deviations between measured values, Q_3, and model values, Q_3^*, is recommended to check systematic trends and randomness of scatter (Figure 3).



Key

$Q_3^* - Q_3$	regression value minus experimental value for cumulative distribution by volume or mass
x	particle size
1	residuals of LND, quasilinear fit
2	residuals of LND, non-linear fit
3	residuals of RRSB, curvilinear fit
●	quasilinear fit point
◆	non-linear fit point
■	curvilinear fit point

Figure 3 — Residual deviations of experimental values, $Q_3(x)$, to quasilinear regression and non-linear regression (log-normal distribution) values, $Q_3^*(x)$, from Figure 2; in addition, the residual deviation of non-linear regression (RRSB) is plotted (see Annex A)

6 Conclusions

The influence of the distribution model type can be seen for log-normal distribution, RRSB, and GGS state models in Annex A. Furthermore, the comparison in Annex A illustrates that the choice of the right model distribution is more important than the different methods of regression.

In summary, the fit of particle size data to a distribution model is dependent upon:

a) the optimization method (e.g. linear, quasilinear, non-linear);

b) the optimization criterion;

c) the type of data — density (q), cumulative (Q) or differential distribution (dQ);

d) the type of quantity, number distribution or mass/volume distribution;

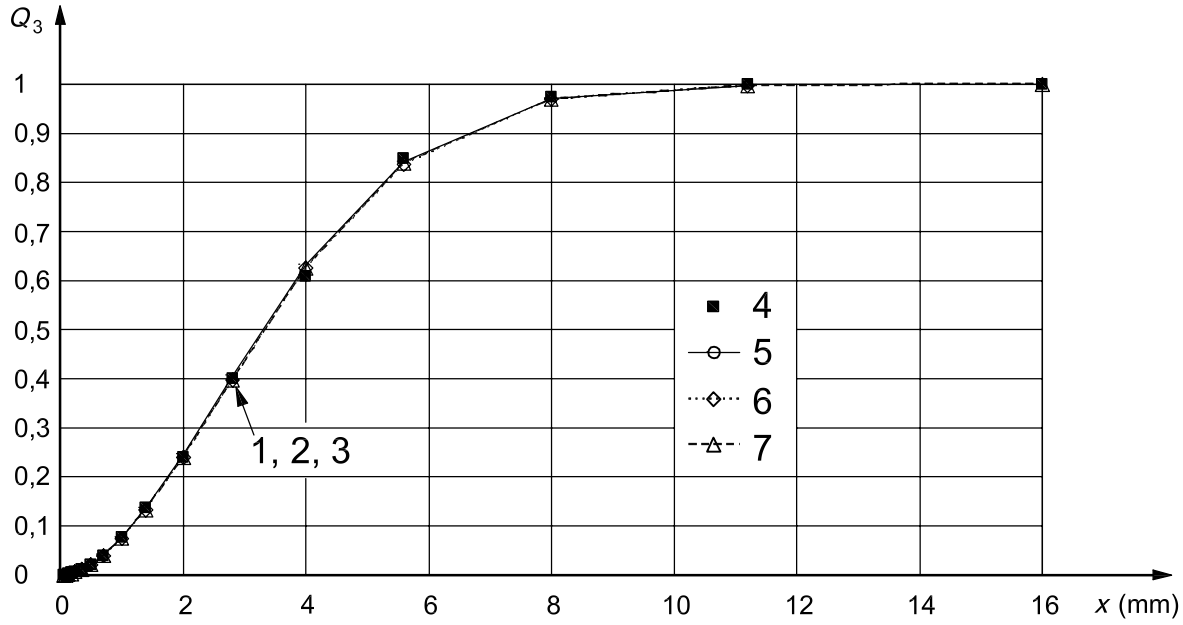
e) the smallest and the largest particle sizes within the distribution.

It is advisable to produce a graphical presentation of the application of any data transformation and the consequences of the calculations involved.

The goodness of fit should not be considered in isolation, but with correspondence to the reference model. For instance the cubic spline interpolation method would produce better goodness of fit, but with model parameters not including physical significance.

Annex A
(informative)

Influence of the model on the regression goodness of fit



Key

Q_3 cumulative distribution by volume or mass

x particle size

1 quasilinear regression

2 weighted quasilinear regression

3 non-linear regression — least squares

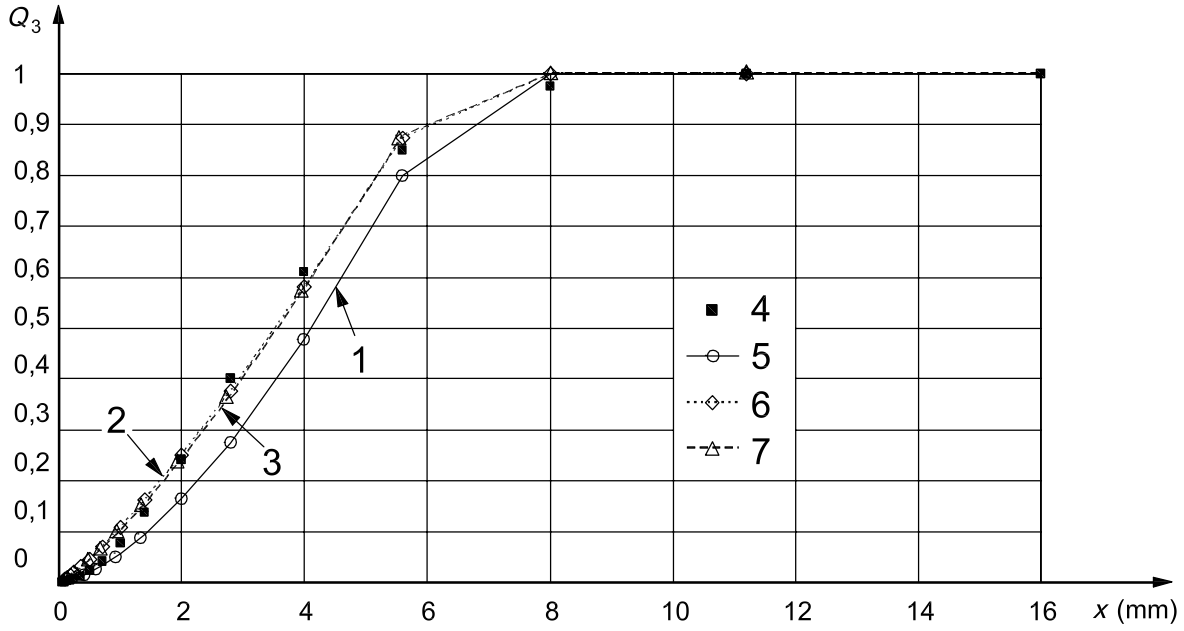
4 $Q_{3,i}$ measured

5 quasilinear fit point

6 weighted quasilinear fit point

7 least squares fit point

Figure A.1 — Example of the RRSB approximation for sieve-analysis data from ISO 9276-1:1998[1], Annex A: a quasilinear regression (full line) (standard deviation of Q_3-residuals 0,006 96); a weighted quasilinear regression (0,005 36) (see Annex E); and a non-linear regression (0,005 35)



Key

Q_3 cumulative distribution by volume or mass

x particle size

1 quasilinear regression

2 weighted quasilinear regression

3 non-linear regression — least squares

4 $Q_{3,i}$ measured

5 quasilinear fit point

6 weighted quasilinear fit point

7 least squares fit point

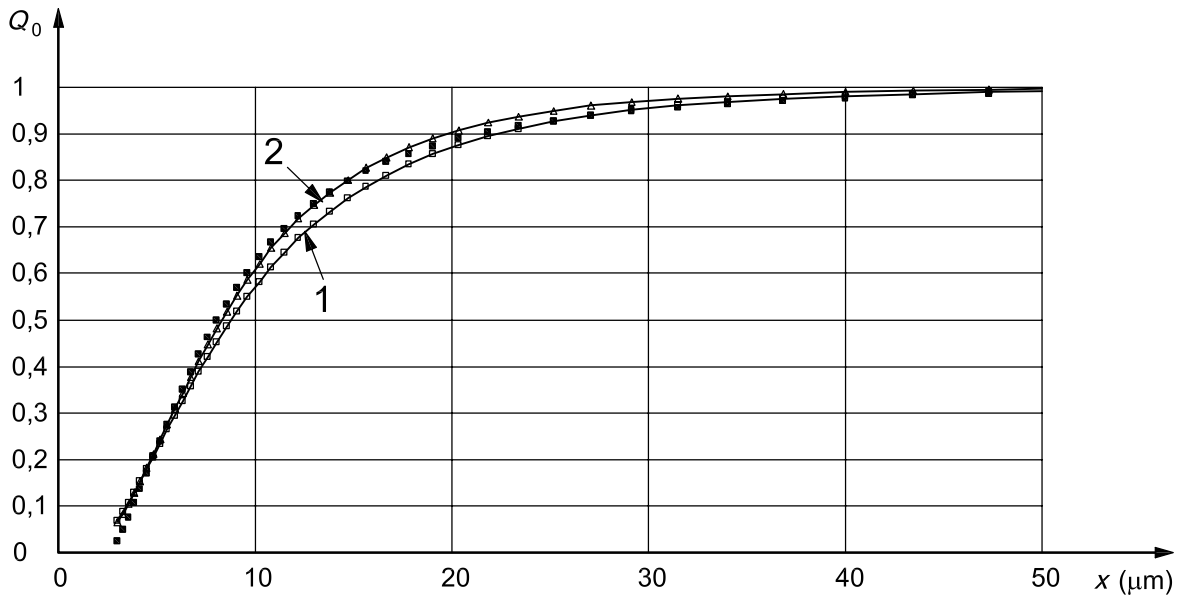
Figure A.2 — Example of the GGS approximation for sieve-analysis data from ISO 9276-1:1998[1], Annex A: a quasilinear regression (full line) (standard deviation of Q_3-residuals 0,054); a weighted quasilinear regression (0,02); and a non-linear regression (0,005 35)

Annex B
(informative)

Influence of the type of distribution quantity on the regression result

This example shows the measurement of 123 807 single milled lactose particles using an extinction counter. The primary result, Q_0, was fitted to a LND (Figure B.1) and additionally converted into a volume distribution, Q_3, using numerical moment equations from ISO 9276-2 (volume of particle-size class calculated from the particle number times third power of particle size). After this model-free numerical conversion, the LND regression of Q_3 (Figure B.2) leads to a slightly different standard deviation.

Figure B.3 demonstrates the relation of a number and a volume LND according to ISO 9276-5 as a horizontal parallel shift in the transformed scales. The different preferences of the fit methods can be seen in Figure B.4 with a log-scaled particle size. The number distribution fit prefers, in this example, the sizes smaller than 20 µm while the volume-distribution fit prefers sizes larger than 20 µm which corresponds to the greatest number of data points. The parallel shift, based on the special LND-property, enables different results in the middle (at the median sizes 8,3 µm and 51 µm) or at the tails of the distribution.



Key

Q_0 cumulative distribution by number

x particle size

1 quasilinear regression

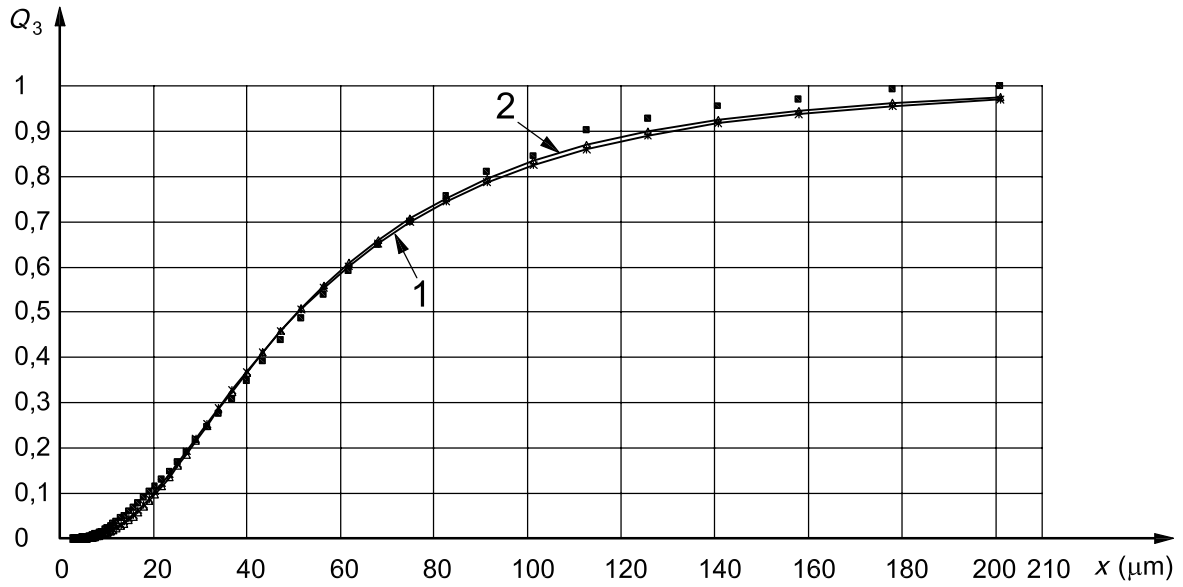
2 non-linear regression

■ $Q_{0,i}$ measured

□ quasilinear fit point

△ non-linear fit point

Figure B.1 — Log-normal distribution: number distribution procedures as in Figure 2, a non-linear regression gives a standard deviation of s = 0,68 and a median size, $x_{50,0}$ = 8,3 µm



Key

Q_3 cumulative distribution by volume or mass

x particle size

1 weighted quasilinear regression

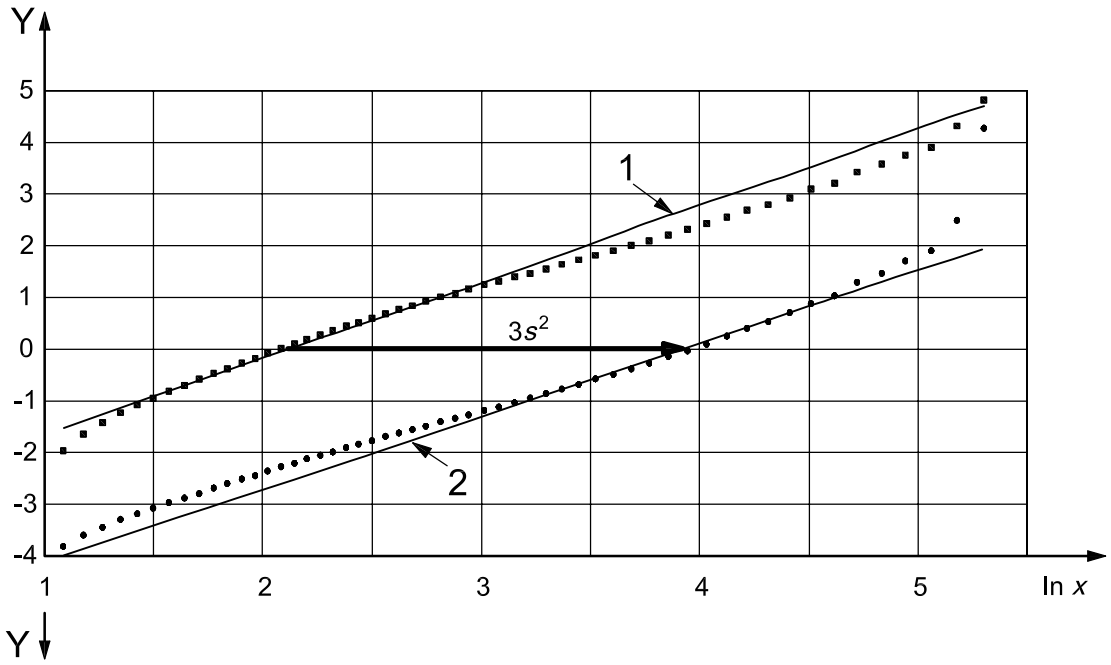
2 non-linear regression

■ $Q_{3,i}$ calculated from $Q_{0,i}$ measured

★ weighted quasilinear fit point

△ non-linear fit point

Figure B.2 — Log-normal distribution: volume distribution procedures as in Figure 2, a non-linear regression gives a standard deviation of s = 0,71 and a median size $x_{50,3}$ = 51 µm



Key

x particle size

Y quantiles of standard normal distribution

1 non-linear regression — Q_0

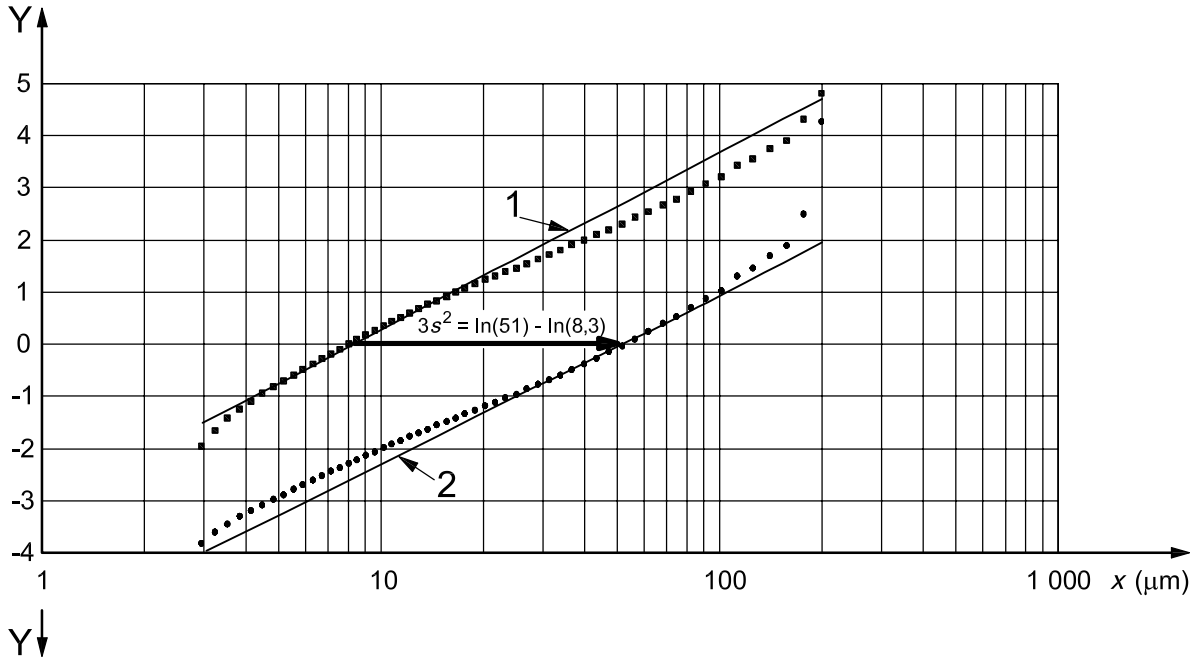
2 non-linear regression — Q_3

■ $Q_{0,i}$ measured

● $Q_{3,i}$ calculated

Figure B.3 — Log-normal distribution: measured number distribution and for each class numerically calculated, volume distribution

Figure B.3 shows no completely exact parallel shift of the regression lines which would result from analytical conversion according to ISO 9276-5. These lines are the result of non-linear regressions in the original x versus Q scale of the number data and separately for the volume data. The log-normal distribution model gives a shift of $3s^2$ for quantity conversion from number to volume according to ISO 9276-5. Here a shift of 1,81 can be seen, that means $s = 0{,}77$ (compared to 0,68 in Figure B.1 and 0,71 in Figure B.2).



Key

x particle size

Y quantiles of standard normal distribution

1 non-linear regression — Q_0

2 non-linear regression — Q_3

■ $Q_{0,i}$ measured

● $Q_{3,i}$ calculated

Figure B.4 —Log-normal distribution from figure B.3 with log-scaled particle size abscissa

Annex C
(informative)

Examples for non-linear regression

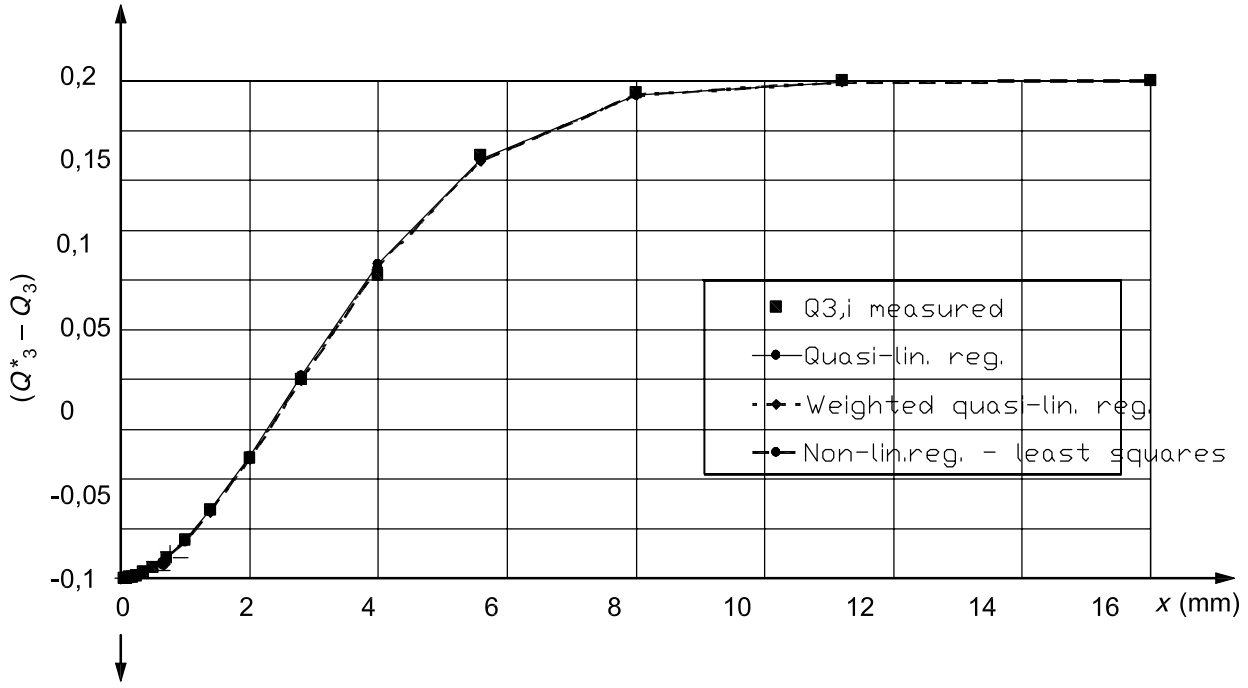
C.1 Spreadsheet example for non-linear regression to log-normal distribution

All the numerical estimation strategies need a first estimate of the adjustment parameters before starting the numerical procedure. As starting estimate, the quasilinear regression in Equation (3) was used.

The numerical procedure is based for instance on the Levenberg-Marquardt method, which is a popular alternative to the Gauss-Newton method (References [7], [8]). For this example, a spreadsheet programme was used, which includes a non-linear regression tool (add-in), based on a code from Reference [9].

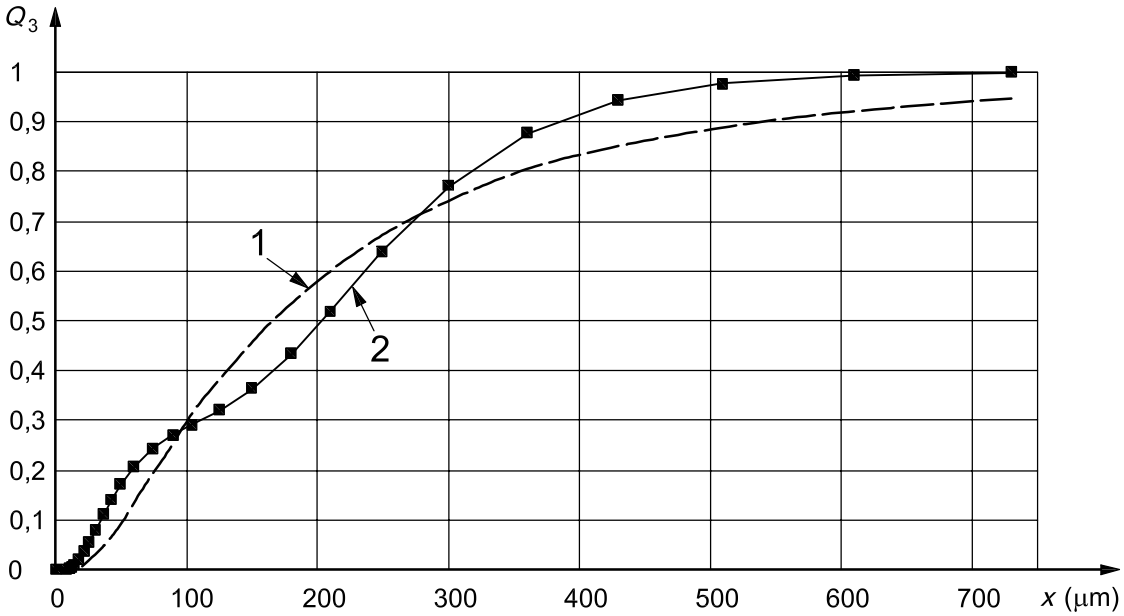
The start parameters from the example in Figure 1 were calculated by Equation (3) to a mean of 2,29 mm and standard deviation of 0,865. Spreadsheet C.1 shows the optimized solver parameters of 3,13 mm and standard deviation of 0,638. This result is plotted in Figure 2.

Spreadsheet C.1

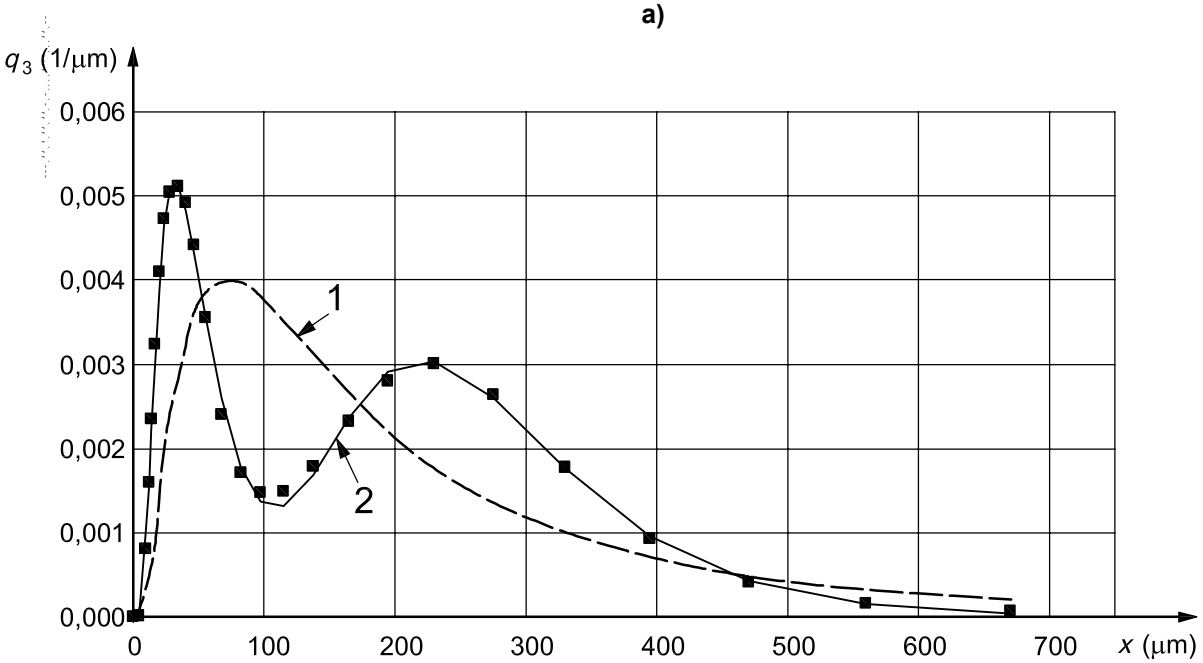


C.2 Example of bimodal distribution with five parameters

In the example of a typical laser diffraction measuring result of coffee powder, the spreadsheet in Clause C.1 was extended to a second log-normal distribution and a mixing ratio. For estimating start parameters, the (monomodal) quasilinear regression with Equation (3) delivered a mean of 134 µm and a standard deviation of 0,874. The following start parameters were used: both standard deviations 0,874, the x_{10} and x_{90} quantiles with 35 µm and 370 µm as mean values and 0,5 as mixing ratio.



a)



b)

Key

Q_3 cumulative distribution by volume or mass

q_3 density distribution by volume or mass

x particle size

1 LND all: $s_{res} = 0{,}059\ 1$

2 LND1+LND2: $s_{res} = 0{,}001\ 7$

■ $Q_{3,i}$ measured

Figure C.1 — The extended spreadsheet calculated the following results for the coffee: log standard deviations 0,652 and 0,377, mean values 47,9 µm and 256,3 µm and a coarse volume fraction of 0,675

Annex D
(informative)

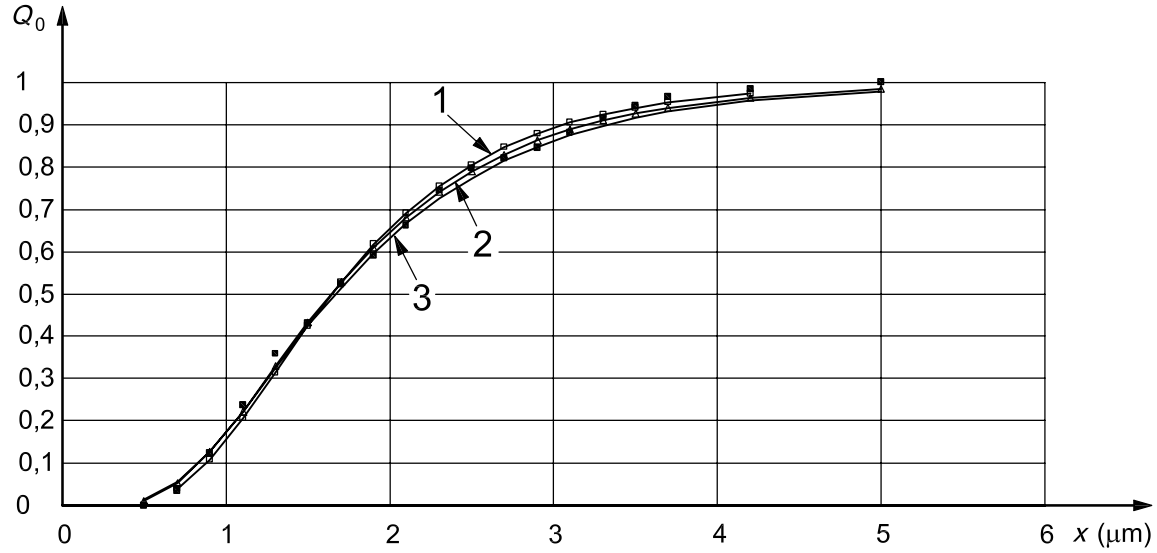
χ^2-Test of number distributions of known sample size

The measuring result of a manual microscope-count of 421 glass spheres was fitted with the non-linear regression of least squares of the Q_0 values in Figure D.1 with a reasonable goodness of fit of 0,016 5.

Another estimation criterion for non-linear fit, which can be used only for number distributions and known sample size, N, is the χ^2-minimum criterion of Equation (5).

Figures D.2 and D.3 show examples for a statistical test, based on the χ^2-criterion. It can quantify the improvement of accuracy when analysing larger particle numbers. This criterion compares the observed particle number variance in the numerator on the right-hand side of Equation (5) with that predicted by Poisson statistics in the denominator of each size class.

The χ^2-minimum criterion is represented by minimization of differences of counted and estimated absolute numbers in the classes, derived from the differences of Q_0 values in Equation (5). Therefore the fit in Figures D.2 and D.3 should be shown for the density values, q_0.



Key

Q_0 cumulative distribution by number

x particle size

1 quasilinear regression — Q_0

2 non-linear regression — least squares — Q_0

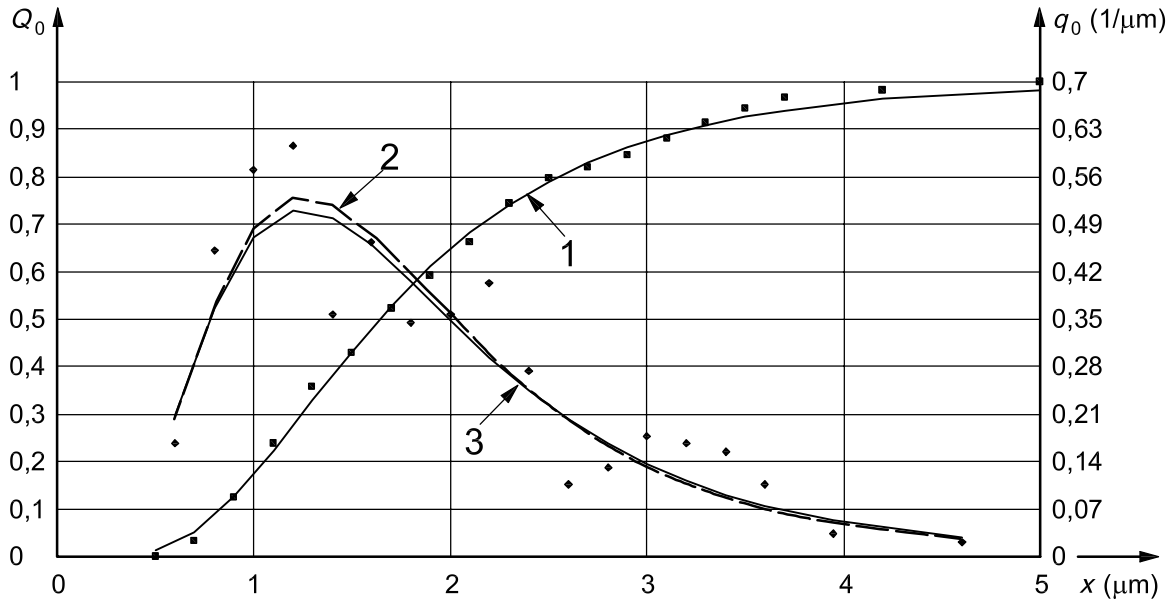
3 non-linear regression — χ^2 — Q_0

■ $Q_{0,i}$ measured

□ quasilinear fit point

△ least squares fit point

Figure D.1 — Log-normal distribution: procedures used are similar to those shown in Figure 2, the standard deviation of ($Q_0 - Q_0^*$) residuals is 0,016 5



Key

Q_0 cumulative distribution by number

q_0 density distribution by number

x particle size

1 non-linear regression — least squares — Q_0

2 q_0 calculated from dQ_0

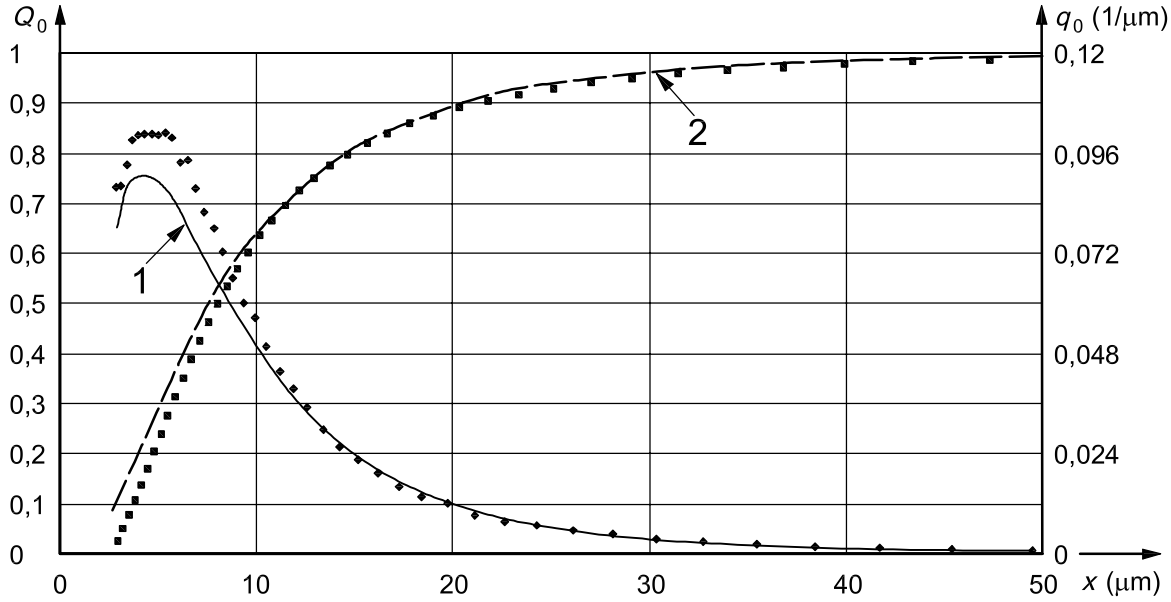
3 non-linear regression — χ^2 — q_0

■ $Q_{0,i}$ measured

◆ $q_{0,i}$ measured

Figure D.2 — Log-normal distribution of $Q_0(x)$ from Figure D.1 and the calculated density distribution (dotted line) in comparison to χ^2-regression of particle numbers from $q_0(x)$.

The non-linear regression — least squares — Q_0 plots in Figures D.1 and D.2 have been derived by minimization of differences of counted and estimated Q_0 values. Differences of LND model-estimated Q_0 values are divided by the class width to derive the q_0 values plotted in Figure D.2, line 2, these are comparable to the χ^2-regression (line 3) of the directly measured q_0 data points (diamonds).



Key

Q_0 cumulative distribution by number

q_0 density distribution by number

x particle size

1 non-linear regression — χ^2 — q_0

2 Q_0 calculated from χ^2 — q_0

■ $Q_{0,i}$ measured

◆ $q_{0,i}$ measured

Figure D.3 — Another example for log-normal distribution: The χ^2-criterion divides the square of the number difference by the estimated number, therefore high numbers are allowed to deviate more

The χ^2-statistical test compares the observed χ^2 (which is equal to 28,6 in the example in Figure D.2, calculated for a sum of 18 classes, i.e. 15 degrees of freedom) with a theoretical χ^2 value of 25 for 5 % error or 30,6 for 1 % error. These errors express the risk of a wrong rejection of the right hypothesis — a log-normal distribution. With a 1 % risk for a wrong rejection, we cannot exclude the log-normal distribution in that example in Figure D.2. This evaluation represents a higher quality of information than the reasonable goodness of the least squares fit of Q_0 of 0,016 5.

With a large number of measured particles (the example in Figure D.3), we can with certainty exclude the log-normal distribution. The χ^2-criterion divides the square of the number difference by the estimated number, therefore high numbers are allowed to deviate more. But $\chi^2 = 209\ 2$ calculated as a sum of 56 classes, i.e. 53 degrees of freedom, must be compared to a theoretical χ^2 value of 71 for a 5 % error or of 80 for a 1 % error.

The χ^2-test quantifies the increase in certainty of fulfilling a model distribution or a quality target distribution resulting from increasing sample size. It should be used with care since such a test can support the interpretation of measurement data, although it cannot replace it automatically.

Other statistical tests may require preconditions, which are not fulfilled by particle size analysis data. The F-test, which may be used to check that two results come from the same source, is very sensitive to deviations from the normal distribution of errors.

ISO/TR 13425[3] gives guidance on the selection and an overview of all the referenced standards, guides, technical reports, and DIS developed by ISO/TC 69 from a user perspective.

ISO 11453[2] includes procedures for tests and confidence intervals relating to proportions, for instance single Q values instead of a distribution.

Annex E
(informative)

Weighted quasilinear regression

E.1 Weighted quasilinear regression

Linear regression of transformed experimental data does not allow an adequate representation of the deviation. Therefore the deviations between data and corresponding model points must be weighted. The weighting factors, ω, represent the mathematical compensation of the non-linear transformation. They can be calculated by inversion of the transformation $Y(Q)$.

When the same data, as shown in Figure 1, together with the quasilinear regression line of Figure 1 is plotted in a non-transformed format, the original data reveals a disproportionality between the deviations which occur in the middle and at the ends of the distribution function (Figure E.1).

The mean square of the vertical deviations between measured and model values in Figure E.1 is reduced obviously in the original scale by the introduction of the weighting factors into the regression. This result of the analytical regression procedure has a standard deviation of Q_3 residuals of 0,032 1, but the numerical procedure of the non-linear regression in Figure 2 results in a better value of 0,026 2. Nevertheless, weighted quasilinear regression is a widely used method because of its analytical procedure. If applied, it should be specified.

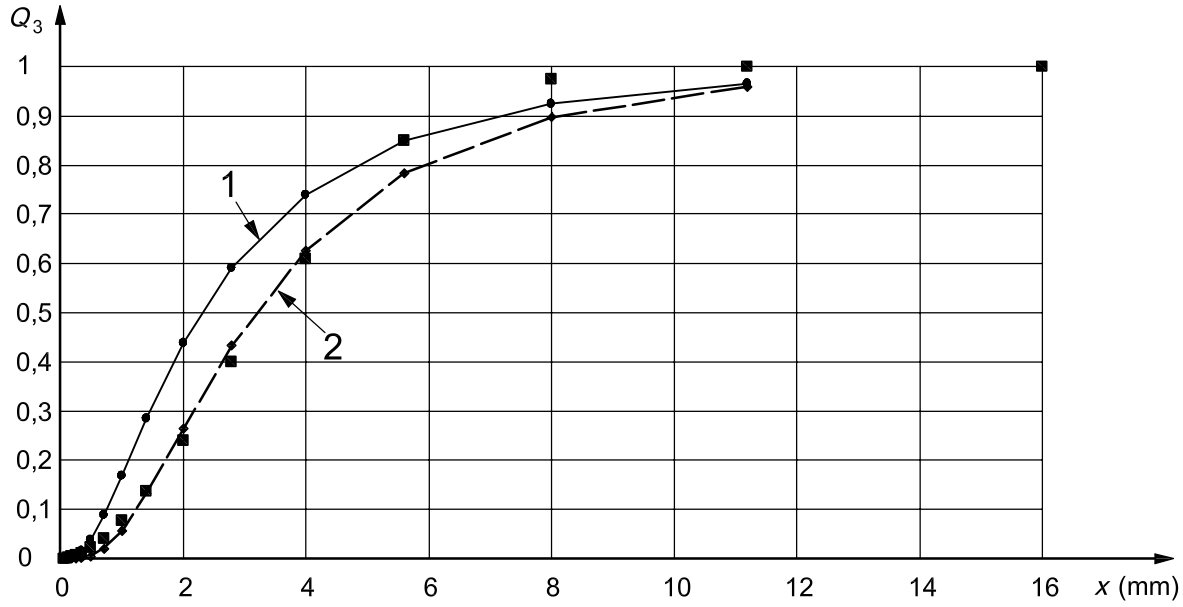
The weighting factors, ω, are needed for the quasilinear regression to compensate for non-linear transformation of the Y differences (which is done unintentionally using special coordinate systems), which result from the non-linear transformation of the Y values.

Equation (E.1) represents an adjusted approximation to the disproportions, which occur, by the transformation. This weighted adjustment, or weighting coefficient, is symbolized by ω.

$$s_{\mathrm{ql}}^2(\boldsymbol{p}) = \frac{1}{n}\sum_{i=1}^{n}\left\{\omega_i\left[bX + a - Q(x_i)\right]\right\}^2 \xrightarrow[\boldsymbol{p}]{} \min \tag{E.1}$$

Equation (E.2) can be used with all standardized distribution models. The different transformation equations are listed in Table E.1.

$$\omega \propto \left[\frac{\mathrm{d}Y(Q)}{\mathrm{d}Q}\right]^{-1} = \left[\frac{\mathrm{d}Q(Y)}{\mathrm{d}Y}\right] \tag{E.2}$$



Key

Q_3 cumulative distribution by volume or mass

x particle size

1 quasilinear regression

2 weighted quasilinear regression

■ $Q_{3,i}$ measured

● quasilinear fit point

◆ weighted quasilinear fit point

Figure E.1 — Log-normal distribution in linear axes: experimental values $Q_3(x)$ from Figure 1 together with quasilinear regression (full line), both plotted on a non-transformed coordinate grid of Q versus x. In addition, the weighted quasilinear regression is shown as a dotted line

Table E.1 — Weighting factors for the three statistical models in Table 1

Quantity	**Model**		
	LND (see also ISO 9276-5)	**RRSB**	**GGS**
$Y(Q)$	$Q(Y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Y} \exp\left(-\frac{\zeta^2}{2}\right) \mathrm{d}\zeta$ with the standard normal distribution $Y = \Phi^{-1}(Y)$	$Y = \ln [-\ln (1-Q)]$	$Y = \ln Q$
Weighting factor $\omega \propto \frac{\mathrm{d}Q(Y)}{\mathrm{d}Y}$	$\exp\left(-\frac{Y^2}{2}\right)$	$\ln (1-Q) \times (1-Q)$	Q

E.2 Formulae of the parameters a and b in the weighted quasilinear regression

The parameters, a and b, in Equations (E.3) and (E.4) respectively, are calculated from the observed values, X and Y, from all the experimental points within a particle size distribution, together with the weighting coefficient, ω, calculated for each point (Reference [10]).

$$a = \frac{\sum \omega Y - b \sum \omega X}{\Sigma \omega} \tag{E.3}$$

$$b = \frac{\sum \omega XY - \left(\sum \omega X \times \sum \omega Y\right) / \sum \omega}{\sum \omega X^2 - \left(\sum \omega X \times \sum \omega X\right) / \sum \omega} \tag{E.4}$$

E.3 Spreadsheet example for Equations (E.3) and (E.4)

Spreadsheet E.1

x mm	**Q3,i**	X	Y		w =	wX	wX²	wY	wY²	wXY
Measured values		ln x	Transf. Q	Weights	Weights²					
0,063	**0**			exp(-Y²/2)						
0,09	**0,001**	-2,41E+00	-3,09E+00	8,44E-03	7,12E-05	-1,72E-04	4,13E-04	-2,20E-04	6,80E-04	5,30E-04
0,125	**0,0019**	-2,08E+00	-2,89E+00	1,52E-02	2,30E-04	-4,78E-04	9,95E-04	-6,66E-04	1,93E-03	1,38E-03
0,18	**0,0035**	-1,71E+00	-2,70E+00	2,63E-02	6,94E-04	-1,19E-03	2,04E-03	-1,87E-03	5,05E-03	3,21E-03
0,25	**0,006**	-1,39E+00	-2,51E+00	4,26E-02	1,82E-03	-2,52E-03	3,49E-03	-4,56E-03	1,15E-02	6,33E-03
0,355	**0,011**	-1,04E+00	-2,29E+00	7,26E-02	5,27E-03	-5,46E-03	5,65E-03	-1,21E-02	2,76E-02	1,25E-02
0,5	**0,022**	-6,93E-01	-2,01E+00	1,32E-01	1,73E-02	-1,20E-02	8,32E-03	-3,49E-02	7,02E-02	2,42E-02
0,71	**0,04**	-3,42E-01	-1,75E+00	2,16E-01	4,67E-02	-1,60E-02	5,47E-03	-8,17E-02	1,43E-01	2,80E-02
1	**0,077**	0,00E+00	-1,43E+00	3,62E-01	1,31E-01	0,00E+00	0,00E+00	-1,87E-01	2,66E-01	0,00E+00
1,4	**0,138**	3,36E-01	-1,09E+00	5,52E-01	3,05E-01	1,03E-01	3,46E-02	-3,33E-01	3,62E-01	-1,12E-01
2	**0,24**	6,93E-01	-7,06E-01	7,79E-01	6,07E-01	4,21E-01	2,92E-01	-4,29E-01	3,03E-01	-2,97E-01
2,8	**0,4**	1,03E+00	-2,53E-01	9,68E-01	9,38E-01	9,66E-01	9,94E-01	-2,38E-01	6,02E-02	-2,45E-01
4	**0,61**	1,39E+00	2,79E-01	9,62E-01	9,25E-01	1,28E+00	1,78E+00	2,58E-01	7,22E-02	3,58E-01
5,6	**0,85**	1,72E+00	1,04E+00	5,84E-01	3,42E-01	5,88E-01	1,01E+00	3,54E-01	3,67E-01	6,10E-01
8	**0,975**	2,08E+00	1,96E+00	1,47E-01	2,15E-02	4,46E-02	9,28E-02	4,21E-02	8,24E-02	8,75E-02
11,2	**0,999**	2,42E+00	3,09E+00	8,44E-03	7,12E-05	1,72E-04	4,16E-04	2,20E-04	6,80E-04	5,32E-04
16	**1**		Y=STANDNORMINV(Q3,i)							
			Sum:		3,34E+00	3,37E+00	4,23E+00	-6,67E-01	1,77E+00	4,78E-01

Bibliography

[1] ISO 9276-1:1998, *Representation of results of particle size analysis — Part 1: Graphical representation*

[2] ISO 11453, *Statistical interpretation of data — Tests and confidence intervals relating to proportions*

[3] ISO/TR 13425, *Guidelines for the selection of statistical methods in standardization and specification*

[4] STINTZ, M., REINSCH, V. Comments on ISO/CD 9276-3. Short communication in ISO/TC 24/SC 4/WG 1, 2004-03-14

[5] NIST/SEMATECH. *e-Handbook of statistical methods*. Available (2008-01-25) at: http://www.itl.nist.gov/div898/handbook/

[6] HEIDENREICH, E., KLÖDEN, W. Zur Auswertung granulometrischer Analysen auf EDV-Anlagen [On the evaluation of granulometric analyses on electronic data-processing systems]. *Chem. Techn*. 1981, **33**, pp. 130-134

[7] BATES, D.M., WATTS, D.G. *Nonlinear regression analysis and its applications*. Wiley, New York, 1988. 365 p

[8] GILL, P.E., MURRAY, W., WRIGHT, M.H. The Levenberg-Marquardt method. In: *Practical optimization*, pp. 136-137. Academic Press, London, 1981

[9] LASDON, L.S., WARREN, A.D., JAIN, A., RATNER, M. Design and testing of a generalized reduced gradient code for nonlinear programming. *ACM Trans. Math. Softw*. 1978, **4**, pp. 34-50

[10] LAFAYE, A., BESANÇON, P. Interprétation mathématique des résultats d'analyses granulométriques. Ajustage d'une courbe cumulative expérimentale à une distribution de référence [Mathematical interpretation of granulometric results. Fitting of an experimental cumulative curve to a reference distribution]. *Sci. Techn. Pharm*. 1984, **13** (3), pp. 159-167

ICS 19.120

Price based on 23 pages