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*

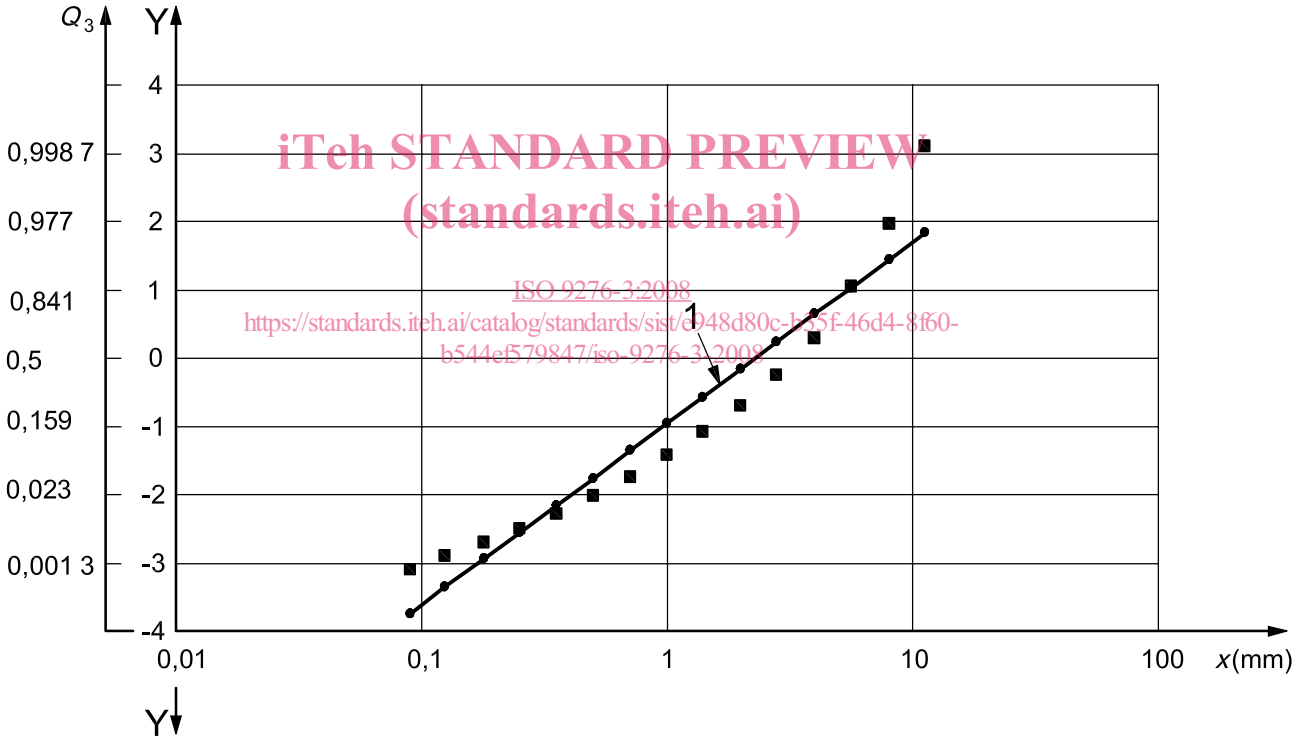
The following part is under preparation:

— *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: $[x_i, Q_i = Q(x_i)]$ $i = 1,\ldots, n$

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):

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

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 $z=(\ln x-\ln x_{50,r})/s$	$Q(x;d',\nu)=1-\exp\left[-\left(\frac{x}{d'}\right)^{\nu}\right]$	$Q_r(x)=\begin{cases}\left(\frac{x}{x_{\max,r}}\right)^{\alpha} & \text{for } x\leqslant x_{\max,r}\\ 1 & \text{for } x>x_{\max,r}\end{cases}$
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).

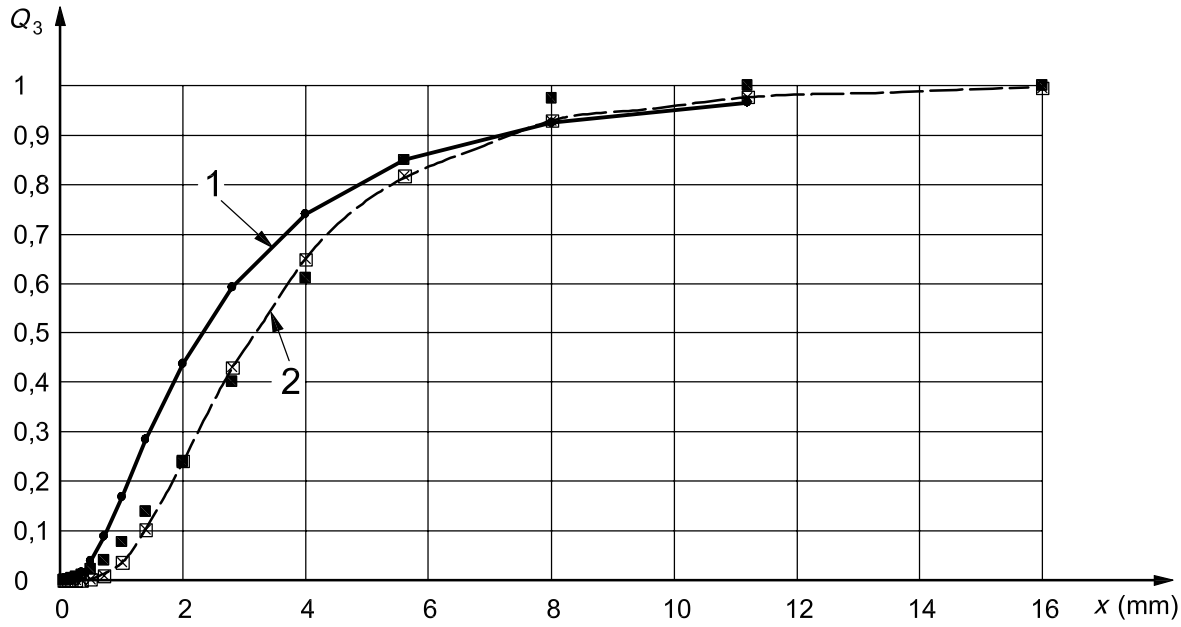
The numerical procedure may be based for instance on the Levenberg-Marquardt method, which is a popular alternative to the Gauss-Newton method (References [7], [8]). Some spreadsheet programs include a non-linear regression tool (add-in) for easy numerical optimization, for instance based on a code from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